جامعة وهران2
كلية العلوم الاجتماعية
قسم علم النفس والأرطوفونيا

أطروحـــة

للحصول على شهـــادة دكتوراه في العلوم في علـــم النفـــس

تخصص: القيـــاس والتقويـــم

استخدام نمـوذج راش اللوغاريتمي أحـادي البارامتر ونظرية القيـاس الكلاسيكيـة في تحليل وتدريج بنود اختبار مستويات التفكير الهندســي مبني وفق نموذج فان هيل

ـ دراسة سيكومترية مقارنة ـ

مقدمة ومناقشة علنا من طرف الطالب:

زيـــاد رشيـــد

أمام لجنـــــة المناقشـــــة المكونة من:

د. بن طـــاهر بشير	أستاذ محاضر –أ-	جامعة وهـــران2	رئيـــــساً
د. بوقصـــاره منصور	أستاذ محاضر –أ-	جامعة وهـــران2	مشرفاً ومقرراً
أ.د. منصوري مصطفى	أستاذ التعليم العالي	جامعة مســـتغانم	منـــــاقشاً
د. زقـــاوة أحمد	أستاذ محاضر –أ-	المركز الجامعي غليزان	منـــــاقشاً
د. بن موســـى سمير	أستاذ محاضر –أ-	جامعة تيـــارت	منـــــاقشاً
د. قمـــراوي محمـــد	أستاذ محاضر –أ-	جامعة وهـــران2	منـــــاقشاً

السنة الجامعيـــة: 2018-2019

بسم الله الرحمن الرحيم

الإهداء

إلى روح أبي الطاهرة وإلى أمي الكريمة رمز العطاء والتضحية أدامها اللّه.

إلى رفيقة دربي ومن شاركتني هموم الحياة وساعدتني على تحمل عناء البحث، إلى زوجتي الغالية (غانية أم إلياس) وفقها اللّه وسدد على دروب الخير خطاها ومتعها بالصحة والعافية.

إلى فلذات أكبادي أبنائي وبناتي الأعزاء، الذين ضحوا بالكثير من أوقاتهم وعطلهم من أجلي كتب اللّه لهم الصلاح والفلاح، ووفقهم اللّه إلى كل خير، ويسر أمورهم في الدنيا والأخرة إلى (إلياس، مريم، إسحاق، أسماء).

إلى إخواني وأخواتي الأعزاء.

شكر وعرفان

الحمد اللّه رب العالمين، القائل في كتابه العزيز: " رَبِّ أَوْزِعْنِي أَنْ أَشْكُرَ نِعْمَتَكَ الَّتِي أَنْعَمْتَ عَلَيَّ وَعَلَى وَالِدَيَّ وَأَنْ أَعْمَلَ صَالِحًا تَرْضَاهُ وَأَصْلِحْ لِي فِي ذُرِّيَّتِي إِنِّي تُبْتُ إِلَيْكَ وَإِنِّي مِنَ الْمُسْلِمِينَ" آية (15) سورة الاحقاف، **والصلاة والسلام على سيدنا محمد ﷺ سيد المرسلين، والمبعوث رحمة للعالمين، القائل في الحديث الذي رواه سعيد الخدري:** " **من لا يشكر الناس لا يشكر اللّه** "، **وعلى آله وصحبه الطيبين الطاهرين، ومن أتبعهم إلى يوم الدين، فبعد شكر اللّه العلي القدير على توفيقه في اتمام هذا البحث، أقدم و أخص بالشكر الجزيل أستاذي الفاضل الدكتور: منصور بوقصاره، المشرف على الأطروحة الذي شاركني الجهد في توجيهي وإرشادي، وتذليل الصعوبات في طريق بحثي بسعة علمه، ورحابة صدره، فجزاه اللّه عنا خير الجزاء.**

كما أتقدم بأسمى عبارات الشكر والتقدير لأعضاء لجنة المناقشة الذين تفضلوا بقبول مناقشة هذه الاطروحة.

كما أوجه شكري الجزيل لجميع الأساتذة المحكمين للاختبار، والشكر موصلاً كذلك لمديرية التربية لولاية الوادي للسماح لي بإجراء هذا البحث، ومن خلالهم إلى مديري الثانويات التي طبق فيها الاختبار، و عينة البحث من التلامذة، على تعاونهم معنا.

كما لا يفوتني أن أتقدم بالشكر الجزيل إلى الدكتور: أيمن القهوجي، والدكتور: جهاد شويخ، و الدكتور: زلمان يوسيسكين (Usiskin,Z) مدير مشروع جامعة شيكاغو (1982)، والدكتور: مايك لياناكر (Linacre,M) مطور برنامج (Winsteps)، على نصائحهم وإجابتهم على كل تسؤلاتي.

وختاماً أسأل المولى عز وجل أن يجعل عملي هذا خالصاً لوجهه الكريم، وأن يلهمني التوفيق والسداد في القول والعمل، إنه ولي ذلك والقادر عليه، والحمد اللّه رب العالمين والصلاة والسلام على أشرف الأنبياء والمرسلين.

الباحث: زياد رشيد

استخدام نموذج راش اللوغاريتمي أحادي البارامتر و نظرية القياس الكلاسيكية في تحليل وتدريج بنود اختبار مستويات التفكير الهندســـي مبني وفق نموذج (فان هيل) – دراسة سيكومترية مقارنة –

المــلخص:

هدف البحث الحالـــي الى المقارنـــة بين النظرية الكلاسيكية للقيـــاس ونموذج (راش) في تحليل وتدريج اختبار محكي المرجع لقياس مستويات التفكير الهندسي مبني وفق نظرية (فان هيل). تكون الاختبار من (28) مفردة من نوع الاختيار من متعدد. طبق الاختبار على عينتين استطلاعية وأساسية مكونتين من (681) تلميذا وتلميذة من تلامذة المرحلة الثانوية. أظهرت نتائج تحليل الاختبار وفق النظرية الكلاسيكية تمتعه بدلالات ثبات وصدق، وكذلك لمؤشرات صعوبة وتمييز مقبولة. كما أظهرت نتائج تدريج الاختبار بعد استبعاد (6) مفردات و(11) فردا، بأن الاختبار ثابت، حيث بلغ معامل الثبات للمفردات (0.99) و معامل الثبات للأفراد (0.65)، كما أن مفردات الاختبار غطت مدى الصعوبة بشكل جيد، فتراوحت قيم الصعوبة للمفردات ما بين (−1.61− 1.46) لوجيت، وقيم قدرات الأفراد ما بين (−3.44− 4.63) لوجيت، ووقعت قيم جميع المفردات (22) مفردة ضمن قيمة احصائي الملائمة لنموذج (راش) (MNSQ) بين (0.7−1.3) ، وقيمة احصائي الملائمة (ZSTD) بين (−0.2،0.2).

أظهرت النتائج وبشكل عام، أن المؤشرات الإحصائية عند تحليل وتدريـــج مفردات اختبار مستويات التفكير الهندســـي كانت مقبولة في كلا النظريتين، إلا أن معالـــم الاختبار وفق نموذج راش كانت أكثر استقراراً ودقة منها في النظرية الكلاسيكية للقياس، والتي من شأنها تحقيق الموضوعية في تفسير النتائج.

كلمـــات مفتاحيـــة: نموذج راش؛ النظرية الكلاسيكية في القياس؛ التفكير الهندســـي؛ نموذج (فان هيل)؛ الاختبار المحكي؛ الأداء التفاضلـــي للمفردة.

Using Rasch one-parameter logistic model and classical measurement theory in analysis In calibrated test items of geometric thinking levels constructed based on Van Hiele model.

-A psychometric Comparative Study –

Abstract :

The objective of the current research is to compare Classical Test Theory (CTT) with Rasch model in Analyzing and calibrating a Criterion Referenced Test (CRT) to measure levels of geometric thinking based on Van Hiele Theory. The test of geometric thinking levels is a multiple choice test consisted of (28) items. The test was administered on an exploratory sample and basic one, both of them composed of (681 males and females) students of secondary school.

The results of the test analysis according to the Classical Test Theory show good reliability and validity indices, as well as an acceptable discrimination and difficulty index. In addition, the results of the calibrated test, after excluding (6) item and (11) person, showed that the test was reliable. The reliability coefficient of the items is (0.99) and the reliability coefficient of persons is (0.65), and also the items of test covered the difficulty range well, between (-1.61 - 1.46) Logits, and values of the ability of persons between (3.44 - 4.63) Logits, and all the 22 items lies within an acceptable range of MNSQ value (0.7-1.3) and ZSTD value (-0.2,0.2).

In general, the results showed that the statistical indicators in the Analyzing and calibrating of the test subjects of the levels of geometric thinking were acceptable in both methods. However, the parameters of the test according to Rasch model were more stable and accurate than in the classical theory of measurement. Which would achieve objectivity in interpreting the results.

Key words: Rasch model, Classical theory of measurement, geometric thinking, Van Hiele model, Criterion Test, Differential Item Functioning (DIF).

L'Utilisation du modèle logistique à un seul paramètre de Rasch et la théorie classique de la mesure pour analyser et calibrage des items de l'échelle les Niveaux de La pensée géométrique qui est construit selon le modèle de Van Hiele.

-Etude psychométrique Comparative–

Résumé:

La recherche actuelle a pour but de faire la comparaison entre la théorie classique de la mesure et le modèle de Rash dans l'analyse et la graduation d'un test de référence pour mesurer les niveaux de pensée géométrique basés sur la théorie de Van Hiele.

l'échelle est appliqué sur deux échantillons exploratoire et fondamentale, composés de (681) élèves du cycle secondaire. Les résultats de l'analyse du test selon la théorie classique ont montré qu'il se caractérise par des significations de validité et de fiabilité, et des indices de difficulté et de discrimination acceptables.

Les résultats de l'échelle après le calibrage, et la suppression de 6 items et 11 personnes, ont montré que l'échelle est fiable, le coefficient de fiabilité des items a atteint 0.99 et des personnes a atteint 0.65, les items de l'échelle couvraient bien l'ampleur de la difficulté, les valeurs de difficulté pour les items allaient enter (-1.61 à 1.46) Logit et les valeurs des capacités des individus entre (-3.34 à 4.63) Logit, ainsi que les indices de (Mean Square Fit Statistics Outfit-Infit, MNSQ), pour 22 items étaient limités enter (0.84 à 1.13) logit, et des indices appropriés de (The Standardized Fit Statistics Outfit- Infit, ZSTD) entre (-0.2 à 0.2).

Les résultats ont montré, et de manière générale, que les indicateurs statistiques lors de l'analyse et du calibrage des items de l'échelle des niveaux de la pensée géométrique, était acceptable selon les deux méthodes, sauf que les paramètres de l'échelle selon le modèle de Rasch étaient plus stable et plus précis que celle de la théorie classique de la mesure. Ce qui permettrait d'obtenir une objectivité dans l'interprétation des résultats.

Mots clés: Modèle de Rasch , la théorie classique de la mesure, La pensée de géométrique, Modèle de Van Hiele, l'échelle critérielle, le fonctionnement différentiel des items (DIF).

قائمة المحتويـــــات

المحتويـــــــات	الصفحة
-البسملة	أ
-الإهداء	ب
-شكر وتقدير.....	ج
-ملخص البحث باللغة العربية.....	د
-ملخص البحث باللغة الإنجليزية.....	هـ
-ملخص البحث باللغة الفرنسية.....	و
-قائمة المحتويات.....	ز
-قائمة الجداول.....	م
-قائمة الأشكال.....	ف
-قائمة الملاحق.....	ر
-مقدمة.....	01
الفصـــل الأول	
مدخل لتقديـــم موضوع البحث	
1- إشكالية البحث.....	08
2- فرضيات البحث.....	13
3- التحديد والضبط الإجرائي للمفاهيم الأساسية للبحث.....	14
4- حدود البحث.....	18
5- أسباب اختيار موضوع البحث.....	19
6- أهميـــة البحث.....	20
7- أهـــداف البحث.....	21
الفصـــل الثانـــي	
القيـــاس النفسي والتربوي وبناء الاختبارات وفق النظرية الكلاسيكية للقياس	
- تمهيد.....	24

1- التطور التاريخي للقياس النفسي والتربوي..... 24
2- مفهوم القياس النفسي والتربوي..... 27
3- مستويات القياس النفسي والتربوي..... 27
4- نظريات القياس النفسي والتربوي..... 31
5- بناء الاختبارات التحصيلية وفق النظرية الكلاسيكة للقياس..... 36
6- الاختبارات مرجعية المعيار (المرجعية الجماعة) NRT أو الاختبارات السيكومترية..... 36
7- الاختبارات مرجعية المحك (المرجعة للهدف) CRT أو الاختبارات الإيديومترية..... 44
- ملخص الفصل..... 60

الفصـــل الثالـــث

توظيف نظرية الاستجابـــة للمفردة الاختبـــارية في بنـــاء الاختبارات التحصيلية

- تمهيد..... 63
1- التطور التاريخي لنظرية الاستجابة للمفردة (IRT) ونماذجها..... 63
2- افتراضات نماذج نظرية الاستجابة للمفردة الاختبارية أحادية البعد..... 65
3- أنواع نماذج الاستجابة للمفردة أحادية البعد..... 74
4- نموذج راش أحادي المعلم (البارامتر)..... 78
4-1- معنى الموضوعية في نموذج راش..... 80
4-2- الجانب السيكولوجي في نموذج راش..... 81
4-3- وحدات القياس المستخدمة في نموذج راش..... 82
5- تقدير بارامترات النماذج أحادية البعد..... 83
6- صدق الاختبار وثباته في ضوء نموذج راش اللوغاريتمي الاحتمالي..... 86
7- تدريج الاختبار وفق نموذج راش..... 87
8- ملاءمة البيانات للنموذج..... 88
8-1- كيفية التحقق من حسن مطابقة البيانات الاختبارية الافتراضات نموذج راش..... 88
9- دالة المعلومات..... 92
10- معادلـــة أو تكافؤ الاختبارات..... 95
11- الأداء التفاضلي (التمايزي) للمفردة الاختبارية (DIF)..... 97

- ملخص الفصل..... 109

الفصـــل الرابـــع

التفكيـــر الهندســـــي

- تمهيـــد..... 111
1- التفكيـــر..... 111
1-1- ماهية التفكير..... 111
1-2- تعريف التفكير..... 112
1-3- نظريات التفكير..... 113
1-4- تصنيفات التفكير..... 115
1-5- أدوات أو عناصر التفكير..... 116
1-6- دور المفاهيم في التفكير..... 116
2- الهندســـة..... 117
2-1- ماهية الهندسة وتعريفها..... 120
2-2- البنية الهندسية..... 121
2-3- المفاهيم الهندسية وطبيعتها..... 121
2-4- النظرة الحديثة للهندسة..... 123
2-5- الهندسة المدرسية..... 123
2-6- أهداف تدريس الهندسة..... 125
2-7- أهمية تدريس الهندسة..... 126
2-8- مهارات التفكير الهندسي..... 129
3- التفكيـــر الهندســـي ونموذج (فان هيل)..... 130
3-1- نموذج (فان هيل) لتنمية التفكير في الهندسة..... 131
3-2- مستويات نموذج (فان هيل) للتفكير الهندســـي..... 131
3-3- خصائص نموذج (فان هيل) لمستويات التفكير الهندسي..... 139
3-4- التعديلات على مستويات التفكير الهندسي لفان هيل..... 143

3-5- النشاط والتفكير من وجهة نظر (بياجيه)..... 146
3-6- الاستنتاج حول أفكار (بياجيه) وأفكار (فان هيل)..... 148
- ملخص الفصل..... 149

الفصـــل الخـــامس

الدراســـات والبحـــوث السابقـــة

-تمهيد..... 151
1- دراســـات تناولت المقارنـــة بين نموذج راش اللوغاريتمـــي والنظرية الكلاسيكيـــة في القياس(CCT) ، والتعليق عليها..... 151
2- دراســـات تناولت استخدام النظرية الكلاسيكيـــة في القياس في تطوير اختبارات لمستويات التفكير الهندسي وفق نموذج فان هيل، والتعليق عليها..... 160
3- دراسات تناولت استخدام نظرية الاستجابة للمفردة الاختبارية (IRT) ونموذج راش في بناء وتدريج اختبارات مستويات التفكير الهندسي وفق نموذج فان هيل، والتعليق عليها..... 171
4- دراسات تناولت الأداء التفاضلـــي للمفردة الاختباريـــة (DIF) باستخدام نظرية الاستجابة للمفردة الاختبارية، والتعليق عليها..... 180
5- التعليق العام على الدراسات السابقة و موقع الدراسة الحالية منها..... 186

الفصـــل الســـادس

الإجـــراءات الميـــدانية للبحث

- تمهيد..... 192
1- المنهج المتبع في البحث..... 192
2- المجتمع وعينة البحث..... 193
3- أداة البحث..... 201
4- خطوات تطبيق اختبار مستويات التفكير الهندسي على أفراد عينة البحث..... 201
5- تجهيز بيانات اختبار مستويات التفكير الهندسي للتحليل الإحصائـــي..... 204
6- الأساليب والبرامج الإحصائية المستخدمة في تحليل بيانات البحث..... 205
- ملخص الفصل..... 209

الفصـــل السابـــع

إجراءات بنـــاء اختبار مستويات التفكير الهندســـي وفقاً لنموذج (فان هيل)

-تمهيد..... 211

1-الخطوات الإجرائية لبناء اختبار مستويات التفكير الهندســـي..... 211

2-التحليل الإحصائي لمفردات اختبار مستويات التفكير الهندســـي..... 224

2-1- مؤشر الصعوبة لمفردات الاختبار..... 225

2-2- مؤشر التمييز أو حساسية مفردات الاختبار..... 227

2-3- تحليل فعالية البدائل (مموهات) الاختبار..... 228

2-4- تقويم الصدق التقاربي والتباعدي (التمييزي) لمفردات الاختبار..... 230

3-تقييم فاعلية اختبار مستويات التفكير الهندســـي..... 232

3-1- تقويم صدق درجات الاختبار..... 232

3-2- تقويم ثبات درجات الاختبار..... 234

-ملخص الفصل..... 236

الفصـــل الثـــامن

عرض وتحليل نتائـــج البحث

- تمهيد..... 238

1-عرض وتحليل نتائج فرضيات البحث..... 239

1-1- عرض وتحليل نتائج الفرضية الأولى..... 239

1-2- عرض وتحليل نتائج الفرضية الثانية..... 247

1-3- عرض وتحليل نتائج الفرضية الثالثة..... 255

1-4- عرض وتحليل نتائج الفرضية الرابعة..... 265

1-5- عرض وتحليل نتائج الفرضية الخامسة..... 271

1-6- عرض وتحليل نتائج الفرضية السادسة..... 281

1-7- عرض وتحليل نتائج الفرضية السابعة..... 286

1-8- عرض وتحليل نتائج الفرضية الثامنة..... 291

1-9- عرض وتحليل نتائج الفرضية التاسعة..... 295

الفصـــل التاســـع

تفسير ومناقشـــة نتائـــج البحث

- تمهيد..... 299

1-تفسير ومناقشة نتائج فرضيات البحث..... 299

1-1- تفسير ومناقشة نتائج الفرضية الأولى..... 299

1-2- تفسير ومناقشة نتائج الفرضية الثانية..... 305

1-3- تفسير ومناقشة نتائج الفرضية الثالثة..... 309

1-4- تفسير ومناقشة نتائج الفرضية الرابعة..... 316

1-5- تفسير ومناقشة نتائج الفرضية الخامسة..... 319

1-6- تفسير ومناقشة نتائج الفرضية السادسة..... 322

1-7- تفسير ومناقشة نتائج الفرضية السابعة..... 324

1-8- تفسير ومناقشة نتائج الفرضية الثامنة..... 329

1-9- تفسير ومناقشة نتائج الفرضية التاسعة..... 332

-خلاصة نتائج البحث..... 337

-توصيات البحث 340

-مقترحات البحث..... 341

-خاتمة البحث..... 342

-قائمة المراجع..... 344

-قائمة الملاحق..... 357

قائمـــة الجداول

الرقم	عنوان الجـــدول	الصفحة
01	يوضح طريقة تسجيل البيانات الخاصة لحساب معامل التمييز بطريقة معادلة برينان.....	52
02	يوضح توافق ثنائي البعد (2*2) لنوع المجموعة (مرجعية ومستهدفة) وطبيعة الإجابـــــة (صح وخطأ).....	101
03	يوضح وصف ومثال لكل مستوى من مستويات التفكير الهندسي الخمسة لـــ(فان هيل)...	137
04	يوضح قيم الارتباطات بين المحكات المستخدمة في طريقة التصحيح التحليلية المطبق على اختبار (CPT) و اختبار (NPT) مع الاختبـــارات الفرعيـــة لاختبار فان هيل لمستويات التفكير الهندسي المعدل.....	162
05	يوضح معاملات الثبـــات للاختبارات المستخدمة بكيودر- ريتشاردسون(KR_{20}) أو ألفا كرونباخ.....	163
06	يوضح توزيع أفراد المجتمع الأصلي وفقا لمتغيري الجنس والمستوى الدراسي.....	193
07	يوضح توزيع عينة البحث على الثانويات الستة والمستويات الدراسية الثلاثة.....	194
08	يوضح عدد أفراد عينة البحث الاستطلاعية ن=104 ونسبتهم المئوية موزعين بحسب المستوى الدراسي والجنس.....	196
09	يوضح عدد أفراد عينة البحث الأساسيـــة ن=577 موزعين بحسب متغيرى الجنس و المستوى الدراسي.....	197
10	يوضح توزيع أفراد عينة البحث بحسب متغير السن وتكراره.....	198
11	يوضح توزيع عينة البحث حسب الشعب العلمية والمستويات الدراسية على الثانويات الستة.....	199
12	يوضح تقويم المحتوى والأهداف السلوكية لكل مفردة من مفردات اختبار مستويات التفكير الهندسي في ضوء أراء المحكمين باستخدام معادلة لاوش (1975) Lawshe	216
13	يلخص نتائج مؤشر صدق المحتوى (CVIs).....	219
14	يوضح قيم مؤشر صعوبة مفردات الاختبار وفقا للعينة الاستطلاعية ن=104.....	226
15	يوضح قيم مؤشرات الحساسية أو التمييز لمفردات الاختبار وفقا للعينة الاستطلاعية.....	227

16 يوضح نتائج تحليل بدائل مفردات اختبار مستويات التفكير الهندسي 28 مفردة وفقا للعينة الاستطلاعية ن= 104..... 229
17 يوضح قيم معاملات الارتباط (بيرسون) بين درجة كل مفردة والدرجة الكلية للمستوى الذي تنتمي إليه بعد حذف درجة المفردة من البعد، والمستويات الأخرى ن=104..... 231
18 يوضح قيم معاملات الارتباط (بيرسون) بين درجات مستويات الاختبار الأربعة (الفرعية) ودرجة الاختبار الكلي وفقا للعينة الاستطلاعية ن=104..... 234
19 يوضح قيم معاملات ثبات الاختبار باستخدام كيودر–ريتشاردسون (KR_{20}) ومعامل سبيرمان وبراون لتصحيح الطول، و معامل (Livingston,1972) وفقا للعينة الاستطلاعية 235
20 يوضح نتائج اختبار (ت) للفروق بين المتوسطات المستقلة تبعاً لمتغير الجنس..... 238
21 يوضح نتائج اختبار (ف) لتحليل التباين الأحادي تبعاً لمتغير المستوى الدراسي..... 239
22 يوضح قيم مؤشرات صعوبة مفردات اختبار مستويات التفكير الهندســـي وفقا للعينة الأساسية ن=577..... 240
23 يوضح قيم مؤشرات التمييز لمفردات اختبار مستويات التفكير الهندسي وفقا للعينة الإساسية ن=577..... 241
24 يوضح قيم معاملات الارتباط (بيرسون) بين درجة كل مفردة والدرجة الكلية وفقاً للعينة الأساسية ن=577 242
25 يوضح قيم معامل الثبات لاختبار مستويات التفكير الهندســي 28 مفردة وفقاً للعينة الاساسية 577 تلميذاً وتلميذة..... 243
26 يوضح قيم معاملات الارتباط بيرسون بين درجة كل مفردة والدرجة الكلية لكل مستوى من مستويات (فان هيل) التفكير الهندسي بعد حذف درجة المفردة من البعد ن=577..... 246
27 يوضح قيم معاملات الارتباط بيرسون بين الدرجات الكلية لمستويات اختبار مستويات التفكير الهندســي الاربعة، ودرجة الاختبار ككل ن=577..... 246
28 نتائج التحليل العاملي للمكونات الأساسية للبواقي(PCAR) باستخدام نموذج (راش) قبل التدريج..... 249
29 يظهر قيم مؤشرات التمييزلاختبار مستويات التفكير الهندســي..... 251
30 يوضح قيم المؤشر الإحصائي Q3 من خلال قيم ارتباط البواقــي المعياريــة..... 253
31 يوضح مؤشرات التخمين لمفردات اختبار مستويات التفكير الهندســي..... 254
32 ملخص نتائج التحليل لقدرات الأفراد..... 256

33	ملخص نتائج التحليل لمعالم الصعوبة لمفردات اختبار مستويات التفكير الهندســـي.....	257
34	يوضح المفردات غير المطابقة وقيم إحصائيات المطابقة الداخلية والخارجية، ومتوسطات المربعات الموزونة لهذه المفردات في التدريج الأولى لاختبار مستويات التفكير الهندسي.	258
35	ملخص نتائج التحليل للقيم المتحررة من قدرات الأفراد بعد استبعاد الأفراد والمفردات غير المطابقة.....	259
36	يوضح ملخص نتائج التحليل المتحرر من صعوبة المفردات بعد التدريج.....	260
37	يوضح صعوبة المفردات وأخطاؤها المعيارية مرتبة تنازلياً وفق صعوبة المفردات.....	261
38	يتضمن تقديرات قدرة الفرد المقابلة لكل درجة كلية محتملة على الاختبـــار، وإنحرافاتها المعيارية، وكذا كمية المعلومات التي يقدمها اختبار مستويات التفكير الهندســـي عند كل مستويات القدرة المختلفة، والعلاقة بين الدرجة الخام والقدرة بوحدة اللوجيت.....	266
39	تقديرات القدرة (باللوجيت) المقابلة لكل درجة محتملة على الاختبار المشتقة من كل من العينة الكلية – العينة المرتفعة – العينة المنخفضة – وكذا الإخطاء المعيارية.....	272
40	تقديرات الصعوبة (باللوجيت) المقابلة لكل مفردة من مفردات لاختبار المشتقة من كل من العينة الكلية والعينة المرتفعة والعينة المنخفضة، وكذا أخطائها المعيارية.....	274
41	يوضح تقسيم مفردات الاختبـــارين الصعـــب والسهـــل ومتوسط القدرة ومقدار الإزاحة للاختبار.....	277
42	يوضح عملية إجراء الموازنة بين صعوبة مفردات كل من الاختبـــارين الصعب والسهل والتدريج المرجعي بعد إجراء التعادل الرأسي.....	278
43	يوضح تقديرات قدرات عينة من الأفراد المشتقة من الاختبـــار الصعب والسهل والكلي وأخطائهما المعيارية.....	280
44	لتقدير صعوبـــة المفردات مقدرة باللوجيت والواط مرتبة تصاعدياً تبعاً للصعوبة 22 مفردة.....	282
45	يوضح معاييـــر القدرة لعينة التقنين 566 تلميذاً وتلميذة، باستخدام نموذج راش (وحدة اللوجيت – وحدة الواط)، و باستخدام النظرية الكلاسيكية في القياس (المعايير التائية – الرتب المئينية).....	283
46	يمثل توزيع متغير القدرة باللوجيت على مجموعة المعايرة 566 تلميذاً وتلميذة، وتوزيع	285

متغير درجة صعوبة مفردات اختبار مستويات التفكير الهندسي باللوجيت 22 مفردة.....
47 نتائج التحليل العاملـــي للمكونات الأساسية للبواقـــي (PCAR) تقديرات نموذج (راش) بعد التدريج..... 287
48 ملخص لنتائج معاملات الثبـــات والفصل وأخطاءهما المعيارية للأفراد والمفردات وفق نموذج راش بعد التدريج النهائي لاختبار مستويات التفكير الهندســـي في صورته النهائية 22 مفردة..... 290
49 يوضح نتائج التحليل العاملـــي المكونات الأساسية للبواقـــي(PCAR) وفق تقديرات راش بعد التدريج وفقاً للجنس (الذكور، والإناث)..... 291
50 يوضح نتائج الأداء التفاضلـــي لمفردات اختبار مستويات التفكير الهندسي في صورته النهائية حسب متغير الجنس (ذكور، إناث)، وباستخدام نموذج راش..... 292
51 يوضح النسب المئوية لتوزيع تلامذة المرحلة الثانوية (الشعب العلمية) على مستويات (فان هيل) بحسب المستوى الدراسي..... 296

قائمة الأشكـــال

الرقم	عنوان الشكـــــل	الصفحة
01	يوضح مستويات القياس الرئيسة.....	28
02	منحنيات خصائص المفردة لمفردتين وفقاً للنظرية الكلاسيكة في القياس.....	33
03	رسم بياني لثورستـــون (Thurstone,1925) يوضح نسبة الإجابـــة الصحيحة على مفردات اختبار "بينيه- سيمون" للأطفال تبعاً لمتغير العمر الزمني.....	64
04	المواضع المقدرة للمفردات على مقيـــاس للعمر ذي نقطة أصل اعتباطيـــة ووحدات قياس متساوية.....	64
05	يوضح المنحنيات المميزة لأربع مفردات.....	70
06	يوضح توازي المنحنيان المميزة للمفردات بحسب نموذج راش أحادي البارامتر.....	71
07	يوضح منحنى خصائص المفردة (ICC) والتنبأ باحتمـــال الاجابـــة الصحيحة في الاستجابـــة الثنائية، في نموذج الاستجابـــة للمفردة ثنائية المعلم. صعوبة المفردة عند القدرة صفر، وتمييز المفردة وهو ميل خط الظل عند القدرة صفر.....	72
08	يوضح أنواع نماذج الاستجابة للمفردة أحادية البعد.....	75
09	التمثيل البياني لإربعة مفردات في النموذج ثنائي البارامتر.....	76
10	التمثيل البياني للنموذج ثلاثى البارامتر.....	78
11	التمثيل البياني لصيغة نموذج راش لإربعة مفردات مختلفة الصعوبة.....	79
12	يوضح تعريف متغير القياس من خلال عملية تدريج أو معايرة مفردات الاختبار.....	88
13	العلاقة بين معلومـــات الاختبار والخطأ المعياري لتقدير الكفاءة (الموضع الذي يكون فيه أكبر معلومات يكون فيه الخطأ المعياري للقياس أقل).....	94
14	يوضح الأداء التفاضلـــي المنتظم.....	107
15	يوضح الأداء التفاضلـــي غير المنتظم.....	108
16	يوضح مستويات التفكير الهندســـي الخمسة حسب نظرية (فان هيل).....	132
17	يوضح منحنى المراحل الثلاثة في اكتساب التفكير الهندســـي وفق نظرية (فان هيل) لمستويات التفكير الهندســـي.....	144

18 تصور(فان هيل) لنموذج التعلم..... 146
19 يوضح نتائج تدريج اختبار (فان هيل) للتفكير الهندســي لأربعة مستويات..... 171
20 يوضح توزيع عينة البحث على الثانويات الستة والمستويات الدراسية الثلاثة..... 195
21 يوضح توزيع أفراد عينة البحث بحسب المستوى الدراسي والجنس..... 198
22 يوضح توزيع أفراد عينة البحث بحسب متغير السن وتكراره..... 199
23 يوضح توزيع عينة البحث بحسب الشعب العلميــة والمستويات الدراسية على الثانويات الستة..... 201
24 يوضح نتائج التحليل العاملــي المتوازي، و الجذر الكامن، والاحداثيات المثلى، والعامل المتسارع..... 248
25 يوضح خريطة اختبــار (Scree Plot) لتباين البواقــي المعياريــة بواسطة برنامج (winsteps)..... 250
26 منحنى خصائص مفردات اختبار مستويات التفكير الهندسي 28 مفردة..... 252
27 يوضح مؤشرات المطابقة للمفردات وحدود المطابقة ومواقع للمفردات من هذه الحدود وفقاً لإحصائي متوسط المربعات التباعــدي والتقاربــي (MNSQ)..... 262
28 منحنى خصائص مفردات اختبار مستويــات التفكير الهندســي بعد التدريج باستخدام نموذج راش..... 263
29 يوضح خريطــة رايت (Wright) لتدريج الأفراد ومفردات اختبار مستويــات التفكير الهندســي..... 264
30 العلاقة التقييسية بين الدرجة الكلية المحتملة على الاختبارعند كل مستوى من مستويات القدرة المختلفة للأفراد..... 267
31 يوضح علاقة دالــة معلومات الاختبار مع الخطأ المعياري للقدرة الكامنة..... 268
32 يوضح خريطة المسار(The Pathway Map) أو المخطط الفقاعي، للأخطاء المعيارية للقياس لكل من صعوبة مفردات اختبار مستويات التفكير الهندسي، وقدرات الأفراد..... 270
33 يوضح منحنى العلاقة التقيسية بين الدرجة الكلية المحتملة على الاختبار وتقديرات القدرة المشتقة من كل من العينة المرتفعة والعينة المنخفضة..... 273
34 يوضح تقديرات الصعوبة لكل مفردة المشتقة من مفردات الاختبار الكلي و أفراد 275

العينتين المرتفعة والمنخفضة المستوى.....
35 توزيع العتبات الفارقة للمفردات والأفراد على عينة التدريج 566 تلميذاً وتلميذة..... 286
36 يوضح اختبار (Scree Plot) لتباين البواقي المعياريـــة (أخطاء القياس)، مستخرج بواسطة برنامج (winsteps)..... 289
37 لمنحنى بياني يوضح الأداء التفاضلي لاختبار مستويات التفكير الهندسي في صورته النهائية، بالنسبة للاختبار ككل (DTF) للمجموعة المرجعية(الذكور)، والمجموعة المستهدفة (الإناث)..... 294
38 لمنحنى بياني يوضح الأداء التفاضلي لمفردات اختبار مستويات التفكير الهندسي في صورته النهائية، بحسب المفردات(DIF) للمجموعة المرجعية ذكور(1)، والمجموعة المستهدفة الإناث(2)..... 295
39 يوضح توزيع التلامذة على مستويات التفكير الهندسي لفان هيل وفقاً للمستوى الدراسي..... 297

قائمة الملاحق

الرقم	عنوان المـــلحق	الصفحة
01	استمارة تحكيم اختبار مستويات التفكير الهندسي لفان هيل في صورته الأولية.....	357
02	جدول يوضح قائمة بأسماء الأساتذة الخبراء المحكمين للاختبار.....	374
03	كراس اختبار مستويات التفكير الهندسي لفان هيل في صورته النهائية.....	375
04	ورقة الإجابة على الاختبار.....	385
05	مفتاح تصحيح الاختبار.....	386
06	طلب رخصة تربص ميداني للباحث من قسم علم النفس والأرطوفونيا بجامعة وهران2 إلى مديرية التربية لولاية الوادي.....	387
07	ترخيص مقدم للباحث بالتطبيق الميداني في الثانويات المعنية من طرف مديرية التربية لولاية الوادي.....	388
08	يوضح بعض مخرجات التحليل الإحصائـــي وفق نظرية الاستجابـــة للمفردة (نموذج راش) باستخدام برنامج وينستبس (Winsteps v4.0.0) لمايك لياكر (2017).....	389
09	تواصل الباحث بالبريد الالكتروني مع الدكتور يوسيسكين (Usiskin) مدير مشروع جامعة شيكاغو(1982)	394
10	يوضح الحوار بين الباحث و مايك لياكر(Mike Linacre) مطور برنامج (Winsteps) ومجموعة من الباحثين المهتمين في موقع التواصل الإجتماعي الفايس بوك.....	395

مقدمـــــة:

إن المتأمل في التراث العلمي للقياس والتقويـــم التربوي والنفسي منذ العشرينات وحتى الآن يستطيع أن يتلمس منحيين رئيسيين في مجال تصميم وبناء الاختبـــارات والمقاييس النفسية والتربوية، حيث يتمثل المنحى الأول في نظرية القياس الكلاسيكيـــة، والتي أحرزت انتشاراً واسعاً بين الباحثين، فامتدت سيادتها لتهيمن على مجال القيـــاس والتقويـــم التربوي والنفسي على مدى خمسة عقود منصرمة.

وبالرغم من اسهاماتها في مجال بناء الاختبـــارات والمقـــاييس و التحقق من جودتها، إلا أنها ظلت عاجزة على إيجاد حلول قاطعة للعديد من مشكلات القيـــاس والتقويـــم النفســـي والتربـــوي.

وباعتبار أن الاختبارات محكيـــة المرجع من أكثر أنواع الاختبارات التحصيلية ملاءمة لقياس وتقويم تحصيل التلامذة، ذلك لأنها تقوم على تحديد المهارات والكفايـــات المطلوب إتقانها، لكي يتمكن المعلم من ملاحظتها وقياسهـــا بشكل مباشر، ومن ثم تقدير مدى ما حققه التلميذ من تلك الأهداف، بناء على مستوى أداء محدد، مما يساعده على عملية التشخيص والتصنيف الدقيق للتلامذة. إلا أنه قد ظهرت مجموعة من المشكلات عند استخدام نظرية الاختبـــارات الكلاسيكية في بنائها، من بينها اتساع حيز الدقة والتغير في الدرجات. ولكن عند استخدام نماذج نظرية الاستجابة للمفردة في عملية تصميم وبناء الاختبارات مرجعية المحـــك، يمكن الحصول على معلومات دقيقة موجودة متساوية بين المدى الواسع للقدرات، وتقديرات لبارامترات المفردة بصورة مستقلة عن بارامتر الفرد.

فنظرية الاستجابـــة للمفردة نشأت منذ مطلع السبعينات، كمدخل جديد في القيـــاس، هدفه الأساســـي تلافي عيوب القياس الكلاسيكي، من اعتمادية نتائج القياس على خصائص الاختبار المستخدم، وخصائص العينة المستخدمة في تقنينه، حيث اتسمت بانجازاتها المتسارعة في ميدان القياس، وتحديداً في مجال تطوير الاختبـــارات والمقاييس النفسيـــة والتربوية، والتوجه نحو القياس الموضوعي للظواهر السلوكية، مما جعلها تقدم مجموعة من الحلول الناجعة للعديد من المشكلات المزمنة في مجال القياس والتي أفرزتها نظرية القيـــاس الكلاسيكية. وبما أن موضوعية وصدق نتائج الاختبـــارات تعتمد على دقـــة الأساليب المستخدمة في بنائها واختيار مفرداتها وتفسير نتائجها، ووصفا للقدرة التي يقيسها الاختبار لذلك فإن الأمر يتطلب ضرورة استخدام نظرية الاستجابـــة للمفردة التي أثبتت البحوث التجريبيـــة أنها تحقق الدقة والموضوعية المنشودة في العلوم النفسية والتربوية، حيث أن أداة القياس تتدرج بوحدة مطلقة ثابتة تتوافق

مع تدريج مستويات المتغير موضوع القيـــاس. أي يمكن من خلالها الحصول على مؤشرات إحصائية للمفردة لا تعتمد على خصائص المفحوصين، وعلامات لقدرة المفحوصين لا تعتمد على صعوبة مفردات الاختبـــار، وبالتالـــي شكلت قفزة على طريق القيـــاس السلوكي نحو تحقيق الموضوعيـــة وتكامل أهداف القيـــاس.

لقد بنيت هذه النظريـــة والنماذج الرياضيـــة المتعلقة بها، على افتراضـــات قوية وصارمة لتحقيق الموضوعية المنشودة للقياس، ومن أهم هذه الافتراضات: أحادية البعد؛ استقلالية القياس؛ منحنيات الخصائص المميزة للمفردات؛ السرعة؛ وعدم التخمين. ويعد نموذج (راش) أحادي البارامتر المختار في بحثنا الحالـــي، من أبسط نماذج الاستجابة للمفردة الاختبـــارية، وأكثره استخداماً في بناء الاختبارات التحصيلية وغيرها، وتحليل مفرداتها. كما يعد من أكثر النماذج التي أجريت حولها دراسات متعددة للتحقق امبريقيـــاً من خصائصه وقدرته في بناء الاختبـــارات، ومواجهة القصور في المدخل الكلاسيكـــي للقياس.(ميمي السيد، 2007، ص.05)

حيث توصلت نتائج البحوث والدراسات المقارنة بين نظرية القيـــاس الكلاسيكية ونموذج راش في اختيار مفردات الاختبـــار، لكل من: (كاظم، 1988ب؛ بوسالم، 2008؛ عوض الله، 2000؛ كريسليب وآخرون، 2001؛ بريتو وآخرون، 2003؛ الهاشمي، 2005؛ ميمي السيد، 2007؛ ماجد، 2012؛ أيمن القهوجي، 2014)، إلى أنها تميل لصالح نموذج راش لأنه أكثر جاذبية نظراً لما يتمتع به هذا النموذج من مميزات تحقيق الموضوعية في القيـــاس. بينما أثبتت دراســـات أخرى، ومن بينها دراسة ويلسون (Wilson, 1990)، (أحمد سليمان عودة، 1992؛ علي عبد الله زكري، 2009) إلى أن الخصائص السيكومترية للاختبارات المبنية وفق النظرية الكلاسيكيـــة لا تختلف كثيراً عن الاختبارات التي يتم تدريج مفرداتها باستخدام نموذج راش. إلا أن القريشي (El-Korashy, 1995) يرى كما ورد عن ميمي السيد بأن استخدام إجراءات النظرية الكلاسيكية في القياس يمكن أن يكون مفيداً في المرحلة الأولى لبنـــاء الاختبار وتنقية صياغة مفرداته، أما نمـــوذج راش فهو الأكثر فعاليـــة فى توفير معالم موضوعية وثابتة لتدريجات المفردات وقياسات الأفراد.(2007، ص.08)

هذا من جانب، ومن جانب آخر تمثل الهندسة أحد الفروع المهمة في علم الرياضيات وأحد مكوناتها الأساسية، لأنها تزود المتعلمين بالمهـــارات الأساسية الضرورية للحياة العملية، مثل: مهارات الحس

المكاني والاستكشاف والقدرة وحل المشكلات والتعليل الاستنتاجي والقدرة على التخمين، كما أنها تتضمن جوانب تعلم معرفيـــة لازمة لفهم وتفسير جوانب التعلـــم المعرفيـــة الأخرى المتضمنة لفروع الرياضيات المختلفة. لهذا أوصى المجلس القومي لمعلمـــي الرياضيـــات الأمريكية (National Council of Teachers of Mat hematics- NCTM) في مؤتمره المنعقد سنة 1989 بضرورة زيادة التركيز على الهندســـة في جميع المستويـــات، واعتبارها من أبرز معايير عقد التسعينات في القرن العشرين، ذلك لأن المعرفة الهندسيـــة وإدراك علاقتها أمران مرتبطان ببيئة الفرد وحياته اليوميـــة، علاوة على ارتباطها الوثيق بمواضيع رياضية وعلمية أخرى، مما يشير إلى اهتمام أكبر بالهندســـة وكيفية تدريسها. (مجدي عزيز، 2005، ص.350)

ومن بين أبرز النظريـــات والأساليب شبه المتفق عليها في تدريس الهندســـة والتي لقيت رواجاً وانتشاراً واسعاً، نظرية (فان هيل) لمستويات التفكير الهندســـي، التي تحتوي على إطار نظري يساعد على تصميم وإعادة بناء مناهج الهندسة في مراحل التعليم العام، حيث قدم (بيير فان هيل) وزوجته (دينا فان هيل) (van hiele, Pier and van hiele, Dina) عام 1957 نموذجاً في التفكير الهندســـي يعرض شكلاً تفصيلياً لمحتويات المناهج الدراسية في الهندسة بتسلسل متتابع في خمسة مستويات رئيسية، وهي: المستوى البصري (الإدراكي)، المستوى التحليلـــي، مستوى الاستدلال غير الشكلـــي، مستوى الاستدلال الشكلـــي، والمستوى الأخير الاستدلال المجرد الكامل. وهذه المستويات الخمسة متسلسلة ومتتابعة حيث لا يستطيع التلميذ أن يتقن مستوى دون أن يكون قد اتقن المستوى أو المستويات السابقة له أو المكونة له، ويوجد لكل مستوى لغته ومصطلحاته، والمفاهيم الهندسيـــة المناسبة له والخاصة به، والانتقال من مستوى أرقى منه لا يعتمد فقط على السن أو النمو البيولوجـــي، بل يعتمد في جزء كبير منه على مستويـــات التدريس ومستوى المادة الهندسيـــة ذاتها. (الشويخ، 2005، ص.09)

مما حدا بالعديد من الباحثين في بحث خصائص النظريـــة، ومنه إلى بناء وتصميم أدوات لقياس مستويات التفكير الهندســـي وفق نظرية (فان هيل)، على غرار دراسة كل من يوسيسكين (Usiskin, 1982)، و(ستولز، 2015؛ الشويخ، 2005؛ مخلوف، 1994، ماكبريد، 1995؛ عفانة، 2002؛ ماجد، 2014؛ نوري، 2015)، والتحقق من مستويـــات التفكير فيها من حيث كونها هرميـــة أم لا؟ وطبيعة هذه المستويـــات هل هي منفصلة عن بعضها أم متصلة؟ وكذلك دور اللغة فيها، لهذا أوصت هذه الدراسات

بضرورة الاهتمام بتطوير أدوات لقياس مستويات التفكير الهندســـي وفق تصور نظرية (فان هيل). ورغم أهمية هذه الأدوات إلا أنها لم تتوفر إلا بأعداد محدودة. لذا فإن الحاجة إلى اختبار لقياس مستويات التفكير الهندســـي بقيت قائمة نظرا للأهمية الكبيرة لمثل هذه الاختبارات خاصة في البيئة الجزائرية.

ومن هذا المنطلق، هدف بحثنا الحالي إلى مقارنة نتائج التحليل بين النظرية الكلاسيكية للاختبارات ونموذج راش اللوغاريتمي أحادي البارامتر في بناء اختبار تحصيلي محكي المرجع في ضوء نظرية (فان هيل) لمستويات التفكير الهندســـي، والذي يقيس تطور التفكير الهندســـي لدى تلامذة المرحلة الثانوية.

وقد أنجز هذا البحث بناء على ما تمليه البحوث السيكومتريـــة المقارنة، والذي جاء في تسعة فصول متسلسلة، كالتالي:

تضمن الفصل الأول مدخل لتقديـــم موضوع البحث، حيث تم تحديد إشكاليـــة البحث، و فرضياتـــه، ثم التحديد والضبط الاجرائـــي للمفاهيم الأساسية للبحث، ومنه إلى توضيح حدود البحث، ثم أسباب اختيار موضوع البحث، وأخيراً أهمية وأهداف البحث.

أما الفصل الثاني، تطرقنا فيه إلى القياس النفسي والتربوي، و كيفية بناء الاختبارات التحصيلية وفقاً للنظرية الكلاسيكية للقياس، وذلك بتسليط الضوء على التطور التاريخي للقياس النفسي والتربوي، و تحديد مفهومه، ثم التطرق إلى مستويات القياس النفسي والتربوي، ثم إلى نظريات القياس النفسي والتربوي، و أخيراً إلى كيفية بناء الاختبارات التحصيلية معيارية المرجع، والمحكية المرجع.

في حين تناولنا في الفصل الثالث كيفية توظيف نظرية الاستجابـــة للمفردة الاختبارية في بناء الاختبارات التحصيلية، وذلك بتوضيح نماذج نظرية الاستجابـــة للمفردة أحادية البعد وافتراضاتها، وتقديم تصنيفاً لنماذج نظرية الاستجابة للمفردة، ثم تحديد مفهوم نموذج راش اللوغاريتمي أحادي البارامتر، ومعنى الموضوعية والجانب السيكولوجي لنموذج راش ووحدات القياس المستخدمة فيه، وإلى كيفية تقدير بارامتراته من حيث صعوبة المفردات، وقدرات الأفراد، وصدق وثبات الاختبار في ضوء نموذج راش، وإلى كيفية التدريج وفق هذا النموذج، ومنه إلى كيفية إجراء اختبار حسن مطابقة وملاءمة البيانات لنموذج راش، وتعريف لمفهوم دالـــة المعلومات و معادلة أو تكافؤ الاختبارات، وأخيراً إلى كيفية الكشف عن الأداء التفاضلـــي للمفردة الاختبارية.

أما الفصل الرابع، فقد خصص للحديث عن التفكير الهندسي، وذلك من خلال تقسيم الفصل إلى ثلاثة أقسام، حيث عالجنا في القسم الأول موضوع التفكير وماهيته وتعاريفه وأهم نظرياته وتصنيفاته وأدواته، وكذلك دور المفـــاهيم في التفكير؛ أما القسم الثاني تطرقنا فيه إلى موضوع الهندســـة، وماهيتها وتعريفها وبنيتها وطبيعة المفاهيم الهندسية والنظرة الحديثة للهندسة، ثم الهندسة المدرسية وأهداف وأهمية تدريسها، و أخيراً عرض مختصر لأهم مهارت التفكير الهندسي؛ أما القسم الثالث استعرضنا فيه التفكير الهندســـي ونموذج (فان هيل)، من خلال التعريف بهذا النموذج وتطوره التاريخـــي و التعريف كذلك بمستوياته الخمسة وخصائصه والتعديلات التي أجريت لاحقاً عليه، وأخيراً المقارنة بين أفكار (جان بياجيه) وأفكار (فان هيل) حول تطور التفكير الهندســـي.

كما تطرقنا في الفصل الخامس، إلى الدراسات والبحوث السابقـــة في هذا الموضوع، والتي قسمت إلى أربعة محاور: المحور الأول تناولنا فيه دراسات المقارنة بين نموذج راش أحادي البارامتر والنظرية الكلاسيكية في القيـــاس والتعليق عليها، في حين تضمن المحور الثانـــي دراسات استخدمت النظرية الكلاسيكية في القياس في تطوير اختبارات لمستويات التفكير الهندسي وفق نموذج (فان هيل) والتعليق عليها، أما المحور الثالث فتناول دراسات استخدمت نظرية الاستجابة للمفردة الاختبارية ونموذج راش في بناء وتدريج اختبارات لمستويات التفكير الهندسي وفق نموذج (فان هيل) والتعليق عليها، في حين تناول المحور الرابع دراسات الأداء التفاضلـــي للمفردة، باستخدام نظرية الاستجابة للمفردة الاختبارية والتعليق عليها، وأخيراً تم التعليق العام على الدراسات السابقة، و تحديد موقع البحث الحالـــي منها.

و تناول الفصل السادس الإجراءات الميدانية للبحث، من خلال عرض المنهـــج المتبع، و مجتمع وعينة البحث و طريقة اختيـــارها وحجمها، و تقديم تعريف مختصر لأداة البحث، ثم توضيح خطوات تطبيق الاختبار على أفراد العينة، وتجهيز البيانات للتحليل الإحصائي وفق المقاربتين الكلاسيكية ونموذج راش في القياس، والتعريف بالأساليب والبرامج الإحصائية المستخدمة في تحليل بيانات البحث.

أما الفصل السابع فقد تناولنا فيه الخطوات الإجرائية عند بناء اختبار محكـــي المرجع لمستويات التفكير الهندسي وفق نظرية (فان هيل)، ثم التحليل الإحصائي لمفرداته بعد التجريب الاستطلاعي له وفقا للنظرية الكلاسيكية في القياس، و تقييم فاعليته من خلال التحقق من مؤشرات صدق درجاته وثباتها.

كما استعرضنا في الفصل الثامن فرضيات البحث وتحليل نتائجها. وأخيراً الفصل التاسع و الذي خصص لمناقشـــة وتفسير نتائج فرضيات البحث، على ضوء الإطار النظـــري والدراسات والبحوث السابقة.

و اختتم البحث بخلاصة للنتائج والاستنتاجات، و أهم مقترحات البحث لمعالجة المشكلة المطروحة، وأهم التوصيات في ما يخص الدراســـات والبحوث المستقبلية، وخاتمة للبحث، وارفق هذا البحث بقائمة للمراجع المعتمدة والملاحق.

ونأمل أن يكون هذا البحث وسيلة للمساهمة في معالجة القصور الواضح في أدوات جمع البيانات في قياس تطور التفكير الهندسي وفقاً لنظرية مستويات التفكير الهندسي لـــ(فان هيل)، والمطورة باستخدام نموذج راش والنظرية الكلاسيكية للقياس. وإن كنا لم ندخر جهداً للإلمام بهذا الموضوع وإعطائه حقهُ من البحث، إلا أنه أثناء ذلك واجهتنا في إنجازه العديد من الصعوبات التي نرى من الضروري الإشارة إليها، وأهمها قلة البحوث والدراسات التي تناولت بناء اختبارات مستويات التفكير الهندســـي وفق نموذج (فان هيل) وباستخدام نظرية الاستجابة للمفردة ونموذج راش خصوصاً، واستخدام الأداء التفاضلي للمفردة، وبنفس تناول البحث الحالـــي، كذلك الصعوبة الكبيرة التي وجدها الباحث في الحصول على البرامج الإحصائية التي تستخدم في التحليل وفقاً لنظرية الاستجابة للمفردة، و باستخدام نموذج راش والتدرب عليها، كذلك الصعوبات المتعلقة بالتطبيق الميداني للاختبار على تلامذة المرحلة الثانويـــة، خاصة وأن الاختبار يتطلب تطبيقه مدة زمنية طويلة نوعاً ما، ومع ذلك حاولنا التغلب على هذه الصعوبات على قدر المستطاع ليظهر البحث في شكله الحالـــي.

الفصـــل الأول

مدخل لتقديـــم موضوع البحث

1- إشكاليـــة البحث:

سيطـــرت النظرية الكلاسيكية للقياس، وما يرتبط بها من نماذج وأساليب إحصائية تتعلق بالمفردات الاختبارية على منهجيات القيـــاس والتقويـــم خلال الثلث الأخير من القرن الماضي، وهي التي قدمت إطاراً مرجعياً لبناء وتقييم الاختبـــارات والمقاييس النفسيـــة والتربويـــة، وتفسير درجاتها (علام، 2013، ص.48). وبالرغم من سيطرتها وانتشارها وما ارتبط بها من مقاييس إحصائية خاصة عند تحليل مفردات الاختبار، إلا أنه تبين في ما بعد قصورها في مواجهة كثير من المشكلات السيكومترية المعاصرة التي تقلل من دقـــة و موضوعية القيـــاس (محاسنة، 2013، ص.96)، والتي من أهمها اعتماد خصائصها السيكومترية، كمعاملات التمييز والصعوبة والثبات على خصائص عينة الأفراد التي يطبق عليهم، وعلى مدى صعوبـــة عينة المفردات التي يشتمل عليها الاختبار (علام، 2001، ص.205)، و عدم وجود وحدة قياس ثابتة، إذ لا تحدد مواضع القيـــاس متصل المتغير بصورة خطيـــة، فاعتماد درجـــات الأفراد على مفردات الاختبار قد يؤدي إلى اختلاف المسافة بين درجتين متتاليتين، ويؤدى هذا إلى اختلاف المعنى الكمي لأي فرد محدد عبر مدى درجات الاختبار، كذلك تأثر الدرجة الكلية للفرد في اختبار ما بمفرداته، إذ تكون درجة الفرد عندما يختبر بمفردات سهلة أعلى منها في حال المفردات الصعبة، فلا يمكن تقدير قدرته فيما تقيسه هذه المفردات قياساً دقيقاً، لذا تختلف نتيجة القياس باختلاف الاختبار المستخدم (محاسنة، 2013، ص.98)، كما تتأثر خصائص مفردات الاختبار بقدرة الأفراد، إذ تختلف معاملات الصعوبة أو السهولة والتمييز لمفردات الاختبار باختلاف قدرة أفراد العينة، فالمفردة التي يختبر بها أفراد ذوي قدرات عالية تبدو سهلة، بينما تبدو نفس المفردة صعبة لذوي القدرات المنخفضة. وإذا كانت العينة متجانسة نسبياً، فإن قيم معاملات التمييز تكون أقل من القيم التي نحصل عليها من عينة غير متجانسة. (Hambleton & Swaminathan, 1989, p.05)، إضافة إلى أنها تقتصر في الموازنة بين الأفراد في السمة أو القدرة التي يقيسها الاختبار على تطبيق نفس مفردات الاختبار أو مجموعة مفردات مكافئـــة أو موازية لها على كل فرد من الأفراد. وبالتالي لا نستطيع الموازنة بين مستويـــات القدرة إذا أجاب الأفراد على مفردات مختلفة ومتباينة في صعوبتها (عبد المسيح، 1991، ص.345)، كما أن ثبات الاختبار يتأثر بالموقف الاختباري، إذ يعتمد ثبات الاختبار في إطار هذه النظرية إما على تطبيق الصورة الاختبارية مرتين على أفراد العينة، أو على إعداد صور متكافئة من الاختبار ويعد هذا في الواقع أمراً صعباً،

وبالرغم من أهمية ذلك، إلا أنه غير كاف، إذ يمكن أن يختلف الموقف الاختباري وظروف التطبيق في هاتين المرتين، والذي اعتبره (هامبلتون وسوامنيثان) أمراً سيؤثر على دقة ثبات الاختبار. (Hambleton & Swaminathan, 1989, p.05)، أيضاً تساوي تباين أخطاء القياس لجميع أفراد العينة موضع الاختبار، بالرغم من أنه قد يكون أداء بعض الأفراد على الاختبار أكثر اتساقاً من غيرهم من الأفراد، وأن درجة هذا الاتساق تختلف باختلاف مستوى قدرة الأفراد أو بمستوى القدرة التي يقيسها الاختبار (Randall, 1998, p.06).

علاوة على ذلك، ظهرت مشكلات أخرى متعلقة باستخدام النظرية الكلاسيكية للقياس، في بناء الاختبارات مرجعية المحك، وهي اتساع حيز الدقة والتغير في الدرجات. ولكن مع ظهور الحاجة الماسة إلى هذا النوع من الاختبارات، قدم خبراء القياس النفسي والتربوي طرقاً و إجراءات جديدة لتطويرها، مما أدى إلى بروز نظرية الاستجابة للمفردة، كنظرية جديدة يمكن استخدامها في بناء وتصميم الاختبارات مرجعية المحك. حيث ترى ميمي السيد(2014)، أنه باعتماد هذه النظرية في تصميم وبناء الاختبارات مرجعية المحك، يمكننا الحصول على معلومات دقيقة موجودة متساوية على المدى الواسع للقدرات وتقديرات لمعالم المفردة بصورة مستقلة عن معالم الفرد. (ص.25)

لهذا شاع استخدام نظرية الاستجابة للمفردة في مجال القياس التربوي، لإعداد الاختبارات التحصيلية، وإنشاء بنوك للأسئلة، كما استخدمت أيضاً في مجال القياس النفسي في بناء وتطوير اختبارات القدرات العقلية. نظراً لعديد المزايا المتعلقة بتطبيق نماذجها في تقييم الاختبارات مرجعية المحك، ومن بينها أنها تجعل مصمم الاختبار لا يقتصر على مجرد تجميع أو صياغة عينة من المفردات التي تقيس سمة معينة، وإنما تجعله أكثر استبصاراً لطبيعة البيانات المتحدة في الاختبار. ويتضح التطبيق المباشر لنماذج الاستجابة للمفردة في حالة الاختبارات مرجعية المحك، إذا أدركنا أن أحد استخدامات هذه الاختبارات هو الحصول على تقدير مستوى اتفاق فرد معين لهدف تعليمي محدد، فنماذج الاستجابة للمفردة تمدنا بهذا التقدير، وخاصة أن فرض أحادية البعد يتخفف عادة في هذه النوعية من الاختبارات. وبالتالي فإن نظرية الاستجابة للمفردة ونماذجها المتعددة يمكن أن تفيد في بناء الاختبارات مرجعية المحك، بحيث تتفق هذه الاختبارات مع المستويات المختلفة لقدرات التلامذة. كما يتم انتقاء المفردات

الاختباريـــة التي تناسب مستوى معيناً من القدرة وتسهم إسهاماً كبيراً في دقة تقدير هذه القدرة. (ميمي السيد، 2014، ص.24).

ونظراً للدور الذي تلعبه الهندســـة في مناهج التعليم العام الذي كان ومازال أحد اهتمامات جميع المهتمين بتدريس الرياضيات خلال العقود الماضية. فعلى المستوى العالمي أوضحت دراسات المعايير الدولية (NCTM 1991; NRC,1989) الأساليب التي ينبغي تبنيها في تدريس الهندســـة في مراحل التعليم العـــام، ومن بين أبرز هذه الأساليب شبه المتفق عليها في تدريس الهندســـة هو نموذج (فان هيل) للتفكير الهندسي، والذي يحتوي على إطار نظري يساعد في تصميم وإعادة بناء مناهج الهندسة في مراحل التعليم العـــام. (clements and Battista,1992 ;Burger and Culpepper,1993 ;Geddes Freudenthal,1973 ; and Geddes and Fortunato,1993 ,1992 ;) (مذكور في: طلال، 2003، ص.30). وهو ما أكدت عليه كذلك وثيقة معاييـــر مناهـــج الرياضيات وتقييمها الصادرة عام 1989 من قبل المجلس الوطني لمعلمي الرياضيات (NCTM ,1989) بالولايات المتحدة بأنها تتفق ومستويات التفكير في الهندسة لـــ(فان هيل)، وتتفق بدورها مع تطور هذه المستويـــات من منطلق أن كل مستوى هو متطلب سابق للمستوى الذي يليه، فتطوير ونمو المعرفة الهندسيـــة لا تتم إلا بالاكتشاف والنقاش والوصف والتحليل.(مجدي عزيز، 2005، ص.350). ونظراً لهذه الأهمية استخدمها السوفييت في تطوير مناهجهم الهندسيـــة، حيث لقي نموذج (فان هيل) اهتماماً خاصاً من ورزيوب (Wirzup) الرائد في التحدث عن هذه النظرية عام 1974، وبعده (هوفر) في كتاباته عام 1981 (.Usiskin, 1982 ; Fuys et al, 1988) .

ركزت أبحاث الثنائـــي (بيير فـــان هيل) وزوجته (ديان فـــان هيل) على تعليم الهندســـة والتفكير، ومستويات التفكير فيها، ودور التعلم في تحسين تلك المستويات لدى المتعلمين في نهاية الخمسينيات. وقد طور هذا الثنائي نظرية مميزة تتعلق بمستويات التفكير الهندســـي تقوم على فكرة مفادها أن عملية التعلم ليست متصلـــة، بل توجد قفزات في منحنى التعلـــم، وخلال سبعينات القرن الماضي، اهتم كثير من التربويين في الولايات المتحدة بهذه النظريـــة ولاقت قبولاً لديهم.(سلامة، 1995، ص.216)، ومن خلال تكثيف البحث في النظرية، واختبار مدى فاعليتها في تعلم الهندســـة منذ بداية الثمانينات، أعلن عن ثلاثة مشاريع كبيرة كان لها بالغ الأثر على البحث في نظرية (فان هيل)، من أهمها مشروع جامعة شيكاغو(Usiskin, 1982) والذي كان موضوعه: "مستويـــات (فان هيل) والتحصيل في هندســـة المدارس

الثانوية" ومن بين أهم نتائجه تصميم اختبار موضوعي يقيس مستويات التفكير الهندسي وفق نظرية (فان هيل)، و مشروع جامعة أوريغون (Burger & Shaugnessy, 1986) الذي صممت فيه المقابلات، ومشروع كلية بروكلين (Fuys, Geddes & Tischler, 1988). وأهم ما تناولته هذه المشاريع هو كيفية تطور التفكير الهندســـي من خلال قياسه سواء بالاختبارات أو المقابلات، ولكن باستخدام النظرية الكلاسيكية للقياس. وباستخدام نظرية الاستجابـــة للمفردة والمتمثلة في نموذج (راش)، الذي يحدد كيفية ارتبـــاط أداء الأفراد على المفردات وعلاقـــة هذا الأداء بالقدرة الكامنـــة لديهم وهو موضوع مخرجات التعلم، قام كل من ويلسون (Wilson 1990) بدراسة بحث فيها " قياس تسلسل هندسة (فان هيل): إعادة التحليل"، وهي دراسة أعاد فيها تحليل بيانات مشروع جامعة شيكاغو(CDASSG) ولكن باستخدام نموذج (راش)، ودراسة ستولز (Stols, 2015) والتي كان موضوعها "استخدام نموذج (راش) لقياس مستويات التفكير الهندســـي".

ومن هذا المنطلق، ورغم المزايا السابقة الذكر لنظرية الاستجابة للمفردة، إلا أن ما توصلت إليه عديد البحوث و الدراسات السابقـــة التي اجريت للموازنة بين فعاليتها مقارنة بالنظرية الكلاسيكية للقياس، والتحقق من مدى قدرة كل نظرية على استخلاص خصائص سيكومترية ذات جودة لأدوات القياس النفسي و التربوي، على غرار دراسة كل من (كاظم، 1988ب؛ عوض الله، 2000؛ بريتو وآخرون، 2003؛ الهاشمي، 2005؛ ميمي السيد، 2007؛ بوسالم، 2008؛ علي زكري، 2009)، والتي كانت نتائجها متضاربة لم تحسم بعد أيّ المقاربتين أكثر كفاءة في استجلاء الكفاءة السيكومترية لأدوات القياس من أجل الوصول إلى تقديرات أكثر دقة لمعالم المفردات وقدرات الأفراد.

وعلى الرغم من أهمية التفكير الهندسي في المجال التربوي، إلا أن الباحث لاحظ بأن هناك قصوراً واضحاً في استخدام نظرية الاستجابة للمفردة في بناء وتطوير أدوات قياس لهذه القدرة، وإلى عدم وجود دراسات جزائرية – في حدود علم الباحث – تناولت إعداد اختبـــارات موضوعية تشخيصية تقيس القدرة على التفكير الهندســـي مبنية وفقاً لنظرية الاستجابة للمفردة باستعمال نموذج راش، والذي يسمح بتحديد ودراسة المسارات التطورية للنظريـــات المتأصلة في تطوير المفاهيم الرياضيـــة، وكذلك تنمية قدرات التلامذة (Callingham and Bond, 2006, p.01). التي يمكن من خلاله الكشف عن مستويات التفكير الهندســـي لدى تلامذة المرحلة الثانوية في ضوء نظرية (فان هيل)، ومقارنة نتائجه بنتائج النظرية الكلاسيكيـــة في القياس.

ومنه يمكن تلخيص مشكلة البحث وصياغتها في السؤال الرئيس التالـــي:

-هل يمكننا التوصل إلى بنـــاء اختبار موضوعي تشخيصي محكـــي المرجع يقيس مستويات التفكير الهندســـي وفقاً لنموذج (فان هيل)، وباستخدام مقاربتين مختلفتين في القيـــاس ومقارنة نتائجهما، يمكن فيه تصنيف تلامذة المرحلة الثانوية وفق هذا المتغير؟

و انطلاقاً من السؤال الرئيس يمكننا طرح التســـاؤلات الفرعية التالية:

1- ما الخصائص السيكومترية لاختبـــار مستويـــات التفكير الهندســـي باستخدام النظرية الكلاسيكية في القياس؟

2- ما درجة ملاءمة نموذج راش لبيانـــات اختبار مستويات التفكير الهندســـي المستمدة من استجابات العينة؟

3- ما تدريج صعوبة مفردات اختبار مستويات التفكير الهندســـي باستخدام نموذج راش؟

4- ما تقدير القدرة ودالـــة المعلومات المقابلة لكل درجة كلية خام محتملة على مفردات اختبار مستويات التفكير الهندســـي باستخدام نموذج راش؟

5- هل يحقق تدريج اختبار مستويات التفكير الهندســـي باستخدام نموذج راش استقلالية القياس بإختلاف عينة التدريج و الاختبار المستخدم؟

6- هل يمكن اشتقاق معايير كميـــة تفسر على أساسها تقديرات الأفراد على اختبـــار مستويات التفكير الهندســـي في صورته النهائية بعد التدريج على تلامذة المرحلة الثانوية؟

7- هل يتحقق صدق وثبات اختبار مستويات التفكير الهندســـي في صورته النهائية بعد التدريج باستخدام نموذج راش؟

8- هل يظهر تدريج صعوبة مفردات اختبار مستويات التفكير الهندسي في صورته النهائية أداءً تفاضلياً تبعاً لمتغير الجنس؟

9- ما هي مستويات (فان هيل) التي يبلغها تلامذة المرحلة الثانوية بعد تطبيق اختبـــار مستويات التفكير الهندســـي في صورته النهائية؟

2- فرضيات البحث:

بناءً على مشكلة البحث وهدفها الرئيس المتمثل في بناء اختبار موضوعي تشخيصي محكي المرجع لقياس مستويات التفكير الهندسي في ضوء نظرية (فان هيل)، وباستخدام النظرية الكلاسيكية للقياس، و نظرية الاستجابة للمفردة الاختبارية والمتمثلة في نموذج (راش) اللوغاريتمي أحادي البارامتر، والإطار النظري لموضوع البحث، وما تم الإطلاع عليه كذلك من دراسات وبحوث سابقة في هذا المجال، و على ضوء تساؤلات البحث، تم صياغة الفروض، التي سيحاول الباحث التحقق منها من خلال مجريات البحث الحالي، كالتالي:

1- يتوفر اختبار مستويات التفكير الهندسي على خصائص سيكومترية مقبولة باستخدام النظرية الكلاسيكية في القياس.

2- تتوفر لبيانات اختبار مستويات التفكير الهندسي المستمدة من استجابات العينة على ملاءمة مقبولة لنموذج راش.

3- تتباين تقديرات صعوبة مفردات اختبار مستويات التفكير الهندسي وتدرجها باستخدام نموذج راش.

4- تتباين تقديرات قدرات الأفراد ودالة المعلومات المقابلة لكل درجة كلية خام محتملة على مفردات اختبار مستويات التفكير الهندسي باستخدام نموذج راش.

5- يحقق تدريج اختبار مستويات التفكير الهندسي باستخدام نموذج راش استقلالية القياس باختلاف عينة التدريج و الاختبار المستخدم.

6- يمكن اشتقاق معايير كمية تفسر تقديرات الأفراد على اختبار مستويات التفكير الهندسي في صورته النهائية بعد التدريج على تلامذة المرحلة الثانوية.

7- يتحقق صدق وثبات اختبار مستويات التفكير الهندسي في صورته النهائية بعد التدريج باستخدام نموذج راش.

8- لا يظهر تدريج صعوبة مفردات اختبار مستويات التفكير الهندسي في صورته النهائية أداءً تفاضلياً تبعاً لمتغير الجنس.

9- يبلغ تلامذة المرحلة الثانوية مستويات متقدمة في مستويات (فان هيل) للتفكير الهندسي بعد تطبيق اختبار مستويات التفكير الهندسي في صورته النهائية.

3- التحديد والضبط الإجرائـــي للمفاهيم الأساسية للبحث:

3-1- النظرية الكلاسيكية في القيـــاس (CTT):

وهي واحدة من نظريات القيـــاس التي تستخدم بغرض تحديد العوامل التي تؤثر على الدرجة التي يحصل عليها الفرد في الاختبار، والتي ترتكز على مفهوم الدرجة الحقيقية والدرجة الخطأ، وتعتمد على موازنة أداء المفحوص بأداء أقرانه وفقاً للمجموع الكلي لدرجاتهم في الاختبار، سواء اجريت على أساس الدرجات الخام، أو الدرجات المعيارية بأنواعها المختلفة، وهذه الدرجات تخضع لخصائص عينة الأفراد التي تستخدم في اشتقاق معايير الاختبـــار، كما تخضع لخصائص عينة المفردات التي يشتمل عليها الاختبار، وهذا الاختلاف يؤثر في صدق إجراء هذه الموازنات وبالتالي صعوبة تعميم نتائجها أو الاستفادة العملية منها. (علام، 1995، ص.18).

3-2- نظرية الاستجابـــة للمفردة (IRT):

تعتمد نظرية الاستجابة للمفردة على فرضية أساسية مؤداها أن القيمة الاحتمالية لاستجابة فرد لمفردة اختبارية تكون دالة للقدرة أو السمة التي يفترض أن الاختبار يقيسها لدى الفرد، وكذلك خصائص المفردة التي يحاول الإجابة عنها، وعادة نحتاج إلى قيمة عددية واحدة تتعلق بالفرد وهو ما يطلق عليه (بارامتر) القدرة المقاسة لدى الفرد، وقيمة عددية أو أكثر تتعلق بالمفردة الاختبارية وهو ما يطلق عليه بارامتر أو بارامترات المفردة . ومن أهم نماذج القيـــاس التي تستخدم للتنبؤ بأداء المفحوص على المفردة نموذج (راش) اللوغاريتمي أحادي البارامتر.(أمين، 2009، ص.77)

3-3- نموذج راش اللوغاريتمـــي أحـــادي البارامتر Rasch Model:

اقترح هذا النموذج عالم الرياضيات الدانماركي جورج راش عام (1960)، والذي يفترض أن جميع مفردات الاختبار تميز بين الأفراد بنفس القدر، إلا أنها تختلف في صعوبتها، حيث لا تتقاطع المنحنيات المميزة للمفردات المختلفة. كما يهتم هذا النموذج بتحديد موقع المفردة الاختبارية على ميزان صعوبة جميع المفردات التي تشكل الاختبار، كما تهتم بتدريج مستويات قدرة الفرد باختبار معين على نفس ميزان تغيير المفردات. (أمين، 2009، ص.79)

3-4- **معلم القدرة Ability Parameter:**

هو مستوى القدرة لدى الأفراد من خلال استجاباتهم لمفردات الاختبار (Hambleton et al, 1991, p.59). و تقدر قدرة الفرد بوحدة اللوجيت، وهي اللوغاريتم الطبيعي لمميز نجاح الفرد على المفردات التي تعبر عن نقطة صفر التدريج عن صعوبتها. (كاظم، 1988ب، ص.27)

3-5- **معلم الصعوبة Difficulty Parameter:**

وهي موقع المفردة على محور مقياس القدرة عندما يساوي احتمال الاستجابة الصحيحة $\frac{C+1}{2}$، حيث (C) احتمال التخمين (Baker, 2001, p.29). وتقدر صعوبة المفردة بوحدة اللوجيت، وهي اللوغاريتم الطبيعي لمميز الفشل لدى الأفراد التي تعبر نقطة صفر التدريج عن قدرتهم. (كاظم، 1988ب، ص.27)

3-6- **تدريج أو معايرة المفردات Item Calibration:**

تعنى عملية التدريج بتحديد التناظر بين البيانات المشاهدة ومواقع الأشخاص على المتغير الكامن. فإذا ما حددنا مواقع الأفراد على المتغير الكامن، نستطيع عندئذ مقارنتهم ببعضهم بعضاً.(دي إيالا، 2017، ص.03).

و يعرف ويتزمان (weitzman) تدريج المفردات، بأنها: " تقدير بارامترات الصعوبة لكل مفردة " (weitzman, 1996, p.783)؛ في حين يرى القريشي، بأنها: " تحديد لمواقع المفردات على متصل فتري (حسب صعوبتها)، فتراته مقسمة بوحدات لوغاريتمية معيارية متوسطها صفر وانحرافها المعياري يساوي1، وعند نقطة الصفر فإن احتمالية الإجابة الصحيحة على المفردة تساوي 0.5" (EL-korashy, 1995, p.756)

ويعرف إجرائياً في بحثنا الحالي، بأنه تدريج لوغاريتمـــي لصعوبة مفردات اختبار مستويات التفكير الهندســـي، وقدرات التلامذة على نفس متصل القدرة وفقاً لنموذج راش اللوغاريتمي أحادي البارامتر، باستخدام البرنامج الإحصائي وينستبس (Winsteps v4.0.0).

3-7- **تحليــل المفــردات:**

وهي عملية فحص استجابات المفحوصين على مفردات الاختبار للحكم على مستوى نوعية كل مفردة، وأكثر المؤشرات التي يرتبط تحليل المفردات في بحثها، معامل صعوبة المفردة، ومعامل تمييزها، وكذلك معامل فعالية أو جاذبية مموهات تلك المفردة. (النبهان، 2013، ص.521)

ويتحدد إجرائياً في بحثنا الحالي، بفحص مؤشرات الصعوبـــة والتمييز و كذلك جاذبية المموهات (المشتتات) لمفردات الاختبار المحكي المرجع لمستويات التفكير الهندســـي المبني وفق نظرية (فان هيل) للتفكير الهندســـي، باستخدام النظرية الكلاسيكية في القياس.

3-8- الأداء التفاضلـــي للمفردة (Differential Item Functioning, DIF):

يعرف أنجوف (Angoff, 1993) الأداء التفاضلي للمفردة، على أنه إبداء المفردة الاختبارية خصائص إحصائية مختلفة في مجموعات مختلفة ظاهرة للعيان بعد معادلـــة المجموعات على مقياس القدرة.(دي إيالا، 2017، ص.396)

ويتحدد إجرائياً في بحثنا الحالي، بتقديرات معالم مفردات اختبار مستويات التفكير الهندسي والتي توصف باللاتغاير في المجموعات الظاهرة (ذكور، إناث)، أي مطابقة معالم المفردات في المجموعتين.

3-9- مستويات التفكير الهندســـي لفان هيل:

قدم فان هيل وزوجته دينا فان هيل(Dina van hiele) عام 1957 نموذجاً، يتعلق بتطور مستويات التفكير الهندســـي، ويفترض هذا النموذج أن جميع التلامذة يتقدمون بتسلسل هرمـــي من خلال خمسة مستويات للتفكير في الهندسة، وهي:

3-9-1- مستوى التعرف (Recognition):

ويطلق عليه أيضا مستوى التصور Visualization، وفي هذا المستوى يدرك أو يتعرف المتعلم على الأشكال الهندسية كوحدات كلية أكثر منها خصائص ومكونات.

3-9-2- مستوى التحليل (Analysis):

وفي هذا المستوى يتمكن المتعلم من تحليل الأشكال الهندسية على أساس خصائصها أو مكوناتها والعلاقات المتداخلة بين تلك المكونات، وذلك من خلال التجريب والملاحظة والقيام ببعض الأنشطة كالقياس.

3-9-3- مستوى الاستنتاج غير الشكلـــي (Informal Deduction):

ويطلق عليه أيضا المستوى شبه الاستدلالي، وفي هذا المستوى يتمكن المتعلم من صياغة واستخدام التعاريف وإكمال برهان استنتاجي لمشكلة معينة.

3-9-4- **مستوى الاستنتاج الشكلـــي (Formal Deduction):**

ويطلق عليه أيضا بمستوى الاستدلال المجرد، وفي هذا المستوى يتمكن المتعلم من إدراك دور الاستنتاج ويصبح ذا معنى بالنسبة له من خلال بناء البراهين الرياضية البسيطة، ويفهم دور المسلمات والتعاريف والنظريات ضمن خطوات البرهان.

3-9-5- **المستوى الاستدلالـــي المجرد الكامل (Rigor Dedction):**

وهو أرقى مستويات التفكير الهندســـي في نموذج (فان هيل)، وفي هذا المستوى يتمكن المتعلم من استخدام المنطق الشكلي في البرهان وفهم دور البرهان غير المباشر.

وطبقاً لهذا النموذج، فإن هذه المستويات تتحقق بمرورها خلال فترات تعلم مختلفة، والإنتقال من مستوى إلى آخر ليس عملية طبيعية، بل تأخذ مكانها تحت تأثير برنامج تعلـــم وتعليم مناسب. (مجدي عزيز، 2005، ص.347-349)

وتتحدد إجرائياً في بحثنا الحالي، من خلال بناء اختبار لمستويات التفكير الهندسي وفق نظرية (فان هيل)، والمكون من الأربعة مستويات الأولى فقط، وهذا بناءً على نتائج عدة دراسات سابقة في هذا الموضوع، ومن أهمها دراسة يوسيسكين (Usiskin,1982) في مشروع جامعة شيكاغو، حيث كان من أبرز نتائج المشروع أن المستوى الخامس يشكل معضلة بالنسبة للنظرية، وإلغائه يعطي نتائج أفضل، فهو إما غير موجود أو لا يمكن قياسه أو فحصه. (الشويخ، 2003، ص.19)

3-10- **الاختبارات التحصيلية محكية المرجع(CRT):**

قدم جليزر و نيتكو (Glaser and Nitko,1980) تعريفا للاختبارات مرجعية المحك بحيث يمكن أن يتضمن أنواعا مختلفة من هذه الاختبارات. فهما يقترحان أن الاختبار مرجعي المحك هو ذلك الاختبار الذي يصمم ليزودنا بقياسات يمكن تفسيرها مباشرة في ضوء مستويات أداء تتحدد عن طريق تعريف فئة أو نطاق من المهام التي ينبغي أن يؤديها الفرد، وتستمد هذه القياسات مباشرة من هذا النطاق وتنسب إليه. (ميمي السيد، 2014، ص.24)

و يتحدد الاختبار التحصيلي محكي المرجع في بحثنا الحالي، باختبار مكون من ثمنيه وعشرين 28 مفردة، من اختيار من متعدد من أربعة بدائل، مبني وفق نظرية (فان هيل)، يقيس مستويـــات التفكير الهندسي الأربعة الأولى فقط.

3-11- **تعريف درجة القطع:**

نظراً لأهمية درجة القطع في القياس محكـــي المرجع، الأمر الذي أدى إلى تعدد تعريفات علماء القياس التربوي والنفسي لمفهوم درجة القطع، ومن تلك التعريفات ما أشار إليه هامبلتون (Hambleton, 1978) حيث عرف درجة القطع، بأنها:" نقطة على متصل درجات الاختبار تستخدم لتصنيف التلامذة إلى فئتين تعكس مستويـــات الأداء المختلفة بالنسبة لهدف معين أو (الأهداف) المراد قياسها في الاختبار". كما عرفها بيرك (Berk)، بأنها :" تلك الدرجة التي تزيد من نسبة التصنيفات الصحيحة أو في المقابل تقلل من نسبة التصنيفات الخاطئة". كما عرفها بابام(Popham)، بأنها: "مقياس لمدى ملائمة التلميذ لهدف محدد". (سوسن شاكر، 2005، ص.186)

وتتحدد درجة القطع إجرائياً في بحثنا الحالي، بالدرجة التي يمكن الإعتماد عليها في إصدار الحكم على تحقيق التلميذ لكل مستوى من مستويات (فان هيل) الأربعة للتفكير الهندســـي، و هي بمثابة نقطة فاصلة بين المتقن وهو من كانت درجته تساوي أو أعلى من درجة القطع، وغير متقن إذا حصل على درجة دون درجة القطع المحددة بالإجابة الصحيحة على 7/5 بالنسبة للمحك المتفائل، و7/6 بالنسبة للمحك الصارم.

3-12- **المحك أو معيار الأداء:**

وهو الحدود المقبولة للأداء على متصل يبدأ قطبه الأدنى من لا كفاءة على الإطلاق و ينتهي قطبه الأعلى بأداء محكم تماما وفق تعريفات وحدات الاختبار مسبقة التحديد. (سوسن شاكر، 2014، ص.70)

ويتحدد المحك إجرائياً في البحث الحالي في نجاح التلميذ في الإجابة على خمس مفردات 5 من سبعة مفردات 7 المكونة لكل مستوى من مستويات (فان هيل) الأربعة للتفكير الهندســـي، أي ما نسبته 71%، و (احتمال الإجابة الصحيحة على 5 من 7 مفردات عن طريق التخمين العشوائي) تساوي = 0.03232 عند مستوى دلالة احصائية 0.01.

4- **حدود البحث:**

يتحدد البحث الحالي بموضوعه المتمثل، في البحث السيكومتري المقارن بين نظرية القياس الكلاسيكية ونظرية الاستجابة للمفردة باستخدام في نموذج راش أحادي البارامتر، في بناء وتحليل وتدريج اختبار لقياس تطور مستويات التفكير الهندسي وفقاً لنموذج (فان هيل).

كما يتحدد بالعينة التي شملها البحث والمكونة من 577 تلميذاً وتلميذة في الدراسة الأساسية، و104 تلميذاً وتلميذة في العينة الاستطلاعية، من تلامذة المرحلة الثانوية الذين يزاولون دراستهم في الشعب العلمية (تقني رياضي، ورياضيات، وعلوم تجريبية). انحصر التطبيق الميداني للاختبار على تلامذة ستة ثانويات بمنطقة وادي ريغ بولاية الوادي (ثانوية حساني لخضر بجامعة، ثانوية متقن جامعة، ثانوية حداج لخضر بالمرارة، ثانوية عمراني العابد بسيدي عمران، ثانوية عبد المجيد حبة بالمغير، ثانوية تندلة). خلال السنة الدراسية 2016-2017، والممتدة من 10 سبتمبر 2016 إلى 15 جوان 2017.

5- أسباب اختيار موضوع البحث:

5-1- سبب اختيار نموذج راش أحـــادي البارامتر:

فعلى الرغم من أن نموذج راش أحـــادي البارامتر يستند إلى عدة افتراضات تتعلق بعدم التخمين، وتساوي جميع المفردات في التمييز، وقياسها سمة أحـــادية البعد، وإتخاذ المنحنى المميزة للمفردة شكل منحنى الترجيح اللوغاريتمي، إلا أن مخالفة بعض هذه الافتراضات في البيانات الاختبارية ربما لا تؤثر كثيراً في دقة التقديرات الناتجة عن استخدام هذا النموذج في مواقف تطبيقية معينة.

ففي النموذج أحادي البارامتر فإن عينة من المختبرين تشتمل على 200 فرد تكون كافية، لذلك فإنه من الناحية العملية يمكن تبرير اختيار نموذج بسيط مثل نموذج راش، إذ رأينا أن تطبيقه في موقف معين يتحمل مخالفة الافتراضـــات فنموذج راش يتميز بخصائص تجعله أكثر تطبيقاً واستخداماً في المواقف العملية، ومن بين هذه الخصائص، مايلي: (علام ، 2013، ص.86)

1. يعد هذا النموذج أقل نماذج الاستجابة للمفردة في عدد الفروض اللازم توافرها في البيانات لكي يعطي تقديرات دقيقة ومتسقة.

2. تمكن علماء القياس من إيجاد حلول مناسبة لمشكلة تقدير بارامتر الصعوبة، وبارامتر قدرة الفرد لهذا النموذج، في حين أنهم يواجهون حتى الآن مشكلات سيكومتريـــة وإحصائية في تقدير بارامترات النموذجين الآخرين، وبخاصة النموذج ثلاثي البارامتر.

3. القيم التقديرية لقدرة فرد حصل على نفس الدرجة في الاختبـــار الذي بني على أساس نموذج راش تكون متساوية بغض النظر عن عينة المفردات التي اختبر بها، في حين أنه ليس من الضروري أن يحدث ذلك في حالة النموذجين الآخرين، لذلك يطلق عليه "نموذج البارامتر الحر في تحليل المفردات ".

5-2- سبب اختيار مفهوم التفكير الهندســي:

1. باعتبار أن الهندســة إحدى فروع الرياضيــات، والتي تعتمد بالدرجة الأولى على الأساليب المتقدمة في التفكير، فهي من أفضل المجالات التي يمكن استثمارها في تنمية التفكير، وهذا ما دفع الباحث إلى اختيار هذا الفرع بالذات من فروع الرياضيات، فالهندســة بفروعها المختلفة تعتبر مجالاً خصبا للتدريب على كيفية استخدام أنماط التفكير في الوصول إلى الحلول المطلوبة، وبالتالي فإن للمضامين الهندسية مميزات خاصة في تنمية الملاحظة والتجريب والقياس والاستنتاج المنطقي وكتابة البراهين وإثباتها، وذلك من خلال إدراك المتعلم للعلاقات الهندسية القائمة على المسلمات والنظريات، ومحاولة تطبيق تلك المسلمات والنظريات في ضوء ما هو معطى لإثبات المطلوب(عفانة، 2003، ص.01)

2. احتلال الهندســة الجزء الأكبر من الرياضيــات الواقعية (المحسوسة) حيث يشاهدها الجميع ويستطيع التلميذ الإحساس بها على العكس من بعض المواضيع الرياضية الأخرى والتي تعد تجريدية بالكامل وليس من السهل على التلميذ التعامل معها وخاصة الجبرية منها، لذا فمعظم المفاهيم الهندسيــة مفاهيم فيزيائية يسهل التعامل معها وتعلميها بيسر وسهولة إذا أحسن المعلم استخدام الوسائل التعليمية اللازمة لفهمها وإتقانها (أبو ملوح، 2002، ص.15).

3. العجز وافتقار البيئة الجزائرية -على حد علم الباحث- لأداة موضوعية للقيــاس والتقويم التربوي، مصممة وفق نظرية القياس الكلاسيكية، و نموذج راش اللوغاريتمي أحادي البارامتر، والمتمثلة في اختبار محكي المرجع لقياس وتشخيص نقاط القوة والضعف بالنسبة للقدرة على التفكير الهندســي لدى تلامذة المرحلة الثانوية على ضوء نظرية (فان هيل) لمستويات التفكير الهندســي، بغض النظر عن خصائصهم وباختلاف الثانويات التي ينتمون إليها.

6- أهمية البحث:

تأتي أهمية البحث الحالي في ما يضيفه من معلومات جديدة على المستوى المحلي والعربي والدولي باعتباره يستخدم نموذج راش والنظرية الكلاسيكية في القياس في بناء اختبار موضوعي تشخيصي محكي المرجع لقياس مستويات التفكير الهندســي تم بناؤه وفق نموذج (فان هيل)، كما أنه من البحوث القليلة في البيئة الجزائرية – على حد علم الباحث- التي استخدم فيها نموذج راش اللوغاريتمي أحادي البارامتر

الذي يتمتع بخاصية استقلالية القياس التي تتيح الفرصة لتعادل نتائج القياس والتعبير عنها بوحدة قياس واحدة ومعرفة، ومدرجة على ميزان تدريج واحد.

كما تكمن أهمية البحث الحالي أيضا في الأهمية الكبيرة لكل من القدرة على التفكير الهندسي بوصفها جانباً من جوانب القدرة الرياضية، والتي لها تأثير كبير على التحصيل الدراسي للتلامذة في بعض المواد الدراسية مثل: الرياضيات، والهندسة، والفيزياء، كما أثبتته العديد من الدراسات الأجنبية والعربية على غرار دراسة يوسيسكين (1982)، وكذلك أهمية المرحلة الثانوية باعتبارها مرحلة تتطلب مستوى عال من الذكاء والقدرات العقلية، و متوجة للمراحل التعليمية السابقة، وما لهذه المرحلة من تأثير على الأداء في امتحان نهاية المرحلة الثانوية أو بما يعرف بشهادة البكالوريا.

7- أهداف البحث:

يكمن الهدف الرئيس لهذا البحث في التوصل إلى بناء اختبار موضوعي تشخيصي محكي المرجع يقيس مستويات التفكير الهندسي وفقاً لنموذج (فان هيل)، وباستخدام مقاربتين مختلفتين في القياس النفسي والتربوي، وهما النظرية الكلاسيكية في القياس ونموذج (راش)، يمكن فيه تصنيف تلامذة المرحلة الثانوية وفق هذا المتغير، كما ينبثق عن هذا الهدف الرئيس مجموعة من الأهداف الفرعية، وهي:

1- محاولة التحقق من الخصائص السيكومترية لاختبار مستويات التفكير الهندسي المبني وفق نموذج (فان هيل) باستخدام النظرية الكلاسيكية في القياس.

2- محاولة التحقق من مدى مطابقة البيانات المستمدة من مفردات اختبار مستويات التفكير الهندسي المبني وفق نموذج (فان هيل) لافتراضات نموذج (راش) أحادي البارامتر.

3- محاولة تدريج صعوبات مفردات اختبار مستويات التفكير الهندسي المبني وفق نموذج (فان هيل) باستخدام نموذج راش.

4- تحديد العلاقة بين الدرجة الكلية ومستويات القدرة على الاختبار المحكي المرجع المبني وفق نموذج (فان هيل) لقياس مستويات التفكير الهندسي باستخدام نموذج راش.

5- محاولة التأكد من أن تدريج اختبار مستويات التفكير الهندسي باستخدام نموذج راش يحقق استقلالية القياس باختلاف عينة التدريج و الاختبار المستخدم.

6- محاولة إيجاد معايير لتفسير قدرات التلامذة على اختبار مستويات التفكير الهندســـي في صورته النهائية بعد تدريجه.

7- محاولة التحقق من مؤشرات صدق وثبات اختبار مستويات التفكير الهندســـي في صورته النهائية باستخدام نموذج راش.

8- محاولة التأكد من أن تدريج صعوبة مفردات اختبار مستويات التفكير الهندسي في صورته النهائية لا يظهر أداءً تفاضلياً تبعاً لمتغير الجنس.

9- محاولة معرفة المستويات التي يبلغها تلامذة المرحلة الثانوية بعد تطبيق اختبار مستويات التفكير الهندســـي المبني في ضوء نموذج (فان هيل).

الفصـــل الثانـــي

القيـــاس النفســـي والتربوي و بنـــاء الاختبارات وفق النظرية الكلاسيكية للقياس

تمهيد:

يعد القياس أمراً على جانب كبير من الأهمية في أي علم من العلوم، فجميع العلوم تسعى لتطوير أساليب موضوعية دقيقة لقياس الظواهر المتعلقة بها، من أجل فهم هذه الظواهر وتفسيرها، والتنبؤ بالعلاقات القائمة بين متغيراتها، ومحاولة ضبطها والتحكم فيها (صلاح علام، 2000، ص.13). فتقدم أي علم من العلوم مرهون بدرجة الدقة التي يصل إليها في قياس السمات التي يتناولها، وعلى حد قول كارل بيرسون (Karel Pearson) الذي يرى بأن الأساس الذي تقوم عليه العلوم ليس طبيعة مادتها أو موضوعها، ولكنها الأسلوب العلمي أو الطريقة العلمية المتبعة في دراسة السمات (بوسالم، 2014، ص.05)، هكذا يسعى علماء النفس والتربية إلى دراسة السمـــات النفسية والتربوية من منظور علمي يقوم على الملاحظة المضبوطة واستعمال الوسائل المناسبة لدراسة الظواهر، إنطلاقاً من مسلمة أساسية مفادها أنه على الرغم من التشابه بين الأفراد إلا أنه هناك تباين وفروق كثيرة بينهم، بعضها قد يظهر للملاحظة العرضية المباشرة كالاختلاف في الطول أو الوزن أو لون البشرة مثلاً، ولكن كثير من الصفات لا تتضح بسهولة بل تحتاج إلى التدقيق والدراسة والاستعانة بوسائل مساعدة لإدراكها وتمييزها (بوسالم، 2008، ص.38). فإدراك الفروق في أي صفة يتطلب التعرف عليها وفهمها أولاً، ثم اكتشاف واستنباط وسائل لقياس هذه الفروق ثانياً، قصد الوصول إلى تفسير علمي لها، هذه العملية عرفت تطوراً تاريخياً يمكن رصد أهم معالمه.

1- التطور التاريخي للقياس النفســـي والتربوي:

منذ اكتشاف الفروق الفرديـــة والجهود قائمة لتقييمها باستخدام مقاييس خاصة لقياسها والدلالة عليها كمياً. وهي التي بدأت بوادرها على أيدي عالم فلكـــي اسمه ماسكلين (Maskelyne) سنة 1776 عندما طرد من مرصد غرينيتش مساعده كينبروك (Kinne brook) ، لأنه تأخر في رصد أحد النجوم فترة تقرب من الثانية. وحدث أن قرأ العالم بيسال (Bessel) هذه القصة فبدأ يهتم بما أصبح يسمى فيما بعد بالمعادلة الشخصية "Personal Equation" (عبد الحفيظ، 2011، ص.15)، ومضمونها أن الأفراد يختلفون من حيث سرعة "زمن الرجع" بين حدوث المثير والاستجابة. وقد أدى هذا الحدث التاريخي إلى إهتمام الباحثين في النصف الأول من القرن التاسع عشر بقياس الفروق الفرديـــة ولم يكن هدف علماء النفس التجريبيون الأول قياس الفروق الفرديـــة، إذ كان الظن السائـــد أنها أخطاء ولذلك اهتموا بدراستها للتخلص منها والوصول إلى

صياغة أوصاف معممة أو قانون عام يصف السلوك الانسانـــي، وهذا هو الاتجاه عند فونت (Wundt) الذي أنشأ أول معمل لعلم النفس في ليبزغ سنة 1879 (أحمد، 1960، ص.06)، والذي انعكس تأثيره على من تعلموا على يديه بعلم وظائف الأعضاء والطبيعيات خلال بحوثهم، فاهتموا بالحساسية للمثيرات الحسية مثل البصرية والسمعية، وبدراسة زمن الرجع "Reaction time". وكان لهذا الاتجاه صداه في الاختبارات النفسيـــة الأولى التي ركزت على الظواهر الحسية، وكان لذلك الاتجاه تأثيراً آخر، إذ نبه إلى أهمية الضبط الدقيق للظروف التي تسجل فيهما الملاحظات، ومن هنا أصبح تقنيـــن الإجراء "Standardization" طابعا يميز الاختبارات النفسية.

وجاء بعد فونت تلميذه سير فرانسيس غالتون (Francis Galton) البيولوجي البريطاني الشهير (1822–1911) حيث تعتبر إسهاماته في القيـــاس أهم معلم في تطور حركة القياس في بريطانيا. فقد كان غالتون مهتم بدراسة الوراثة لدى الانسان وخاصة دراسة الصفات المتشابهة والمختلفة بين الأقارب وغير الأقارب. ومن أهم إسهاماته تطويره لعمليات إحصائية بالتعاون مع تلميذه كارل بيرسون (Karl Pearson)، ومن أهم المقاييس الإحصائية التي تم التوصل إليها نذكر معامل الارتباط بيرسون، والوسيط، والسلم الترتيبي وغير ذلك، حيث كان له الفضل في تأسيس منهج جديد في البحث النفســـي وهو المنهج الإحصائـــي . (بوسنه، 2007، ص.23)

كما كان للسيكولوجي الأمريكي جيمس ماكين كاتل (james Makeen cattell) جهود كبيرة في تطور حركة القيـــاس النفســـي وعلم النفس التجريبـــي. فقد أكمل دراسة الفروق الفردية في زمن الرجع، وتعاون مع جالتون وإن لم يؤيدهما فونت. وأنشأ وأدار معمل علم النفس التجريبي، ونشر حركة القياس واستخدم إصطلاح "الاختبار العقلي mental test" سنة 1890 لأول مرة إذ تعرض لبطاريات اختبارات لقيـــاس الكثير من السمات الحسيـــة الحركيـــة والنفسيـــة. (أحمد، 1960، ص.07).

وفي عام 1895 أشارت جمعية علم النفس الأمريكية إلى أهمية دراسة الفروق الفرديـــة وشكلت لجنة لهذا الغرض كان كاتل أحد أفرادها. وكان هدف اللجنة تنمية دراسة الفروق الفردية بالتعاون مع مختبرات علم النفس الموجودة في ذلك الوقت. واستنتج من دراسات عديدة حول الفروق الفردية في الذكاء والقدرات العقلية، أن أكثر ضعاف العقول ذكاء لا يمكن تمييزهم بسهولة عن أقل الأسوياء ذكاء. وأدت هذه الملاحظة إلى التعرف على حقيقة هامة تعتبر جوهرية في مجال القياس النفســـي، وهي أن الفروق بين الأفراد في

الخصائص العقلية، وغيرها من الخصائص السيكولوجية، مهما كانت هذه الخصائص، هي فروق في الدرجة (الكم) وليس في النوع (الكيف). وبدأت معرفة السيكولوجيين الأمريكيين باختبارات الذكـــاء على نمط اختبارات بينيه بعد ترجمة هنري جودارد 1866–1957 (H.H.Goddard) لاختبارات بينيه سنة 1910. (بشير معمرية، 2007، ص.56)

كما اهتم لويس ترمان 1877– 1956 (Lewis Terman) بالاختبارات العقلية وأعد بحثه للدكتوراه فيها سنة 1905 وبدأ في إنجاز بحث ضخم على اختبار بينيه على عينات تقنين أمريكية تكونت من 2300 طفل من أطفال المدارس والراشدين. وتضمنت دراسته تغييراً وتبديلا في صياغة ومضامين وأماكن العديد من البنود في المستويات العمريـــة المختلفة، وإلغـــاء البعض الآخر وإضافة بنود جديدة، حتى كاد الاختبار أن يكون جديداً يختلف جذرياً عن اختبار بينيه الأصلي. ونشر بحثه لأول مرة عام 1916، وكانت الطبعة الأولى للاختبار، وصار يعرف منذ ذلك التاريخ بإسم "اختبار ستانفورد – بينيه للذكاء". ولم يتوقف قياس الذكـــاء في أمريكا على اختبار "ستانفورد– بينيه للذكـــاء"، إذ بمجرد ظهوره وتطبيقه، وجهت له ملاحظات على أنه لفظي ويقيس الذكـــاء عند فئة الأطفال فقط. لذا اتجهت الجهود نحو إنجاز اختبار يتجنب الملاحظتين السابقتين، فظهر عام 1939 اختبار وكسلر– بلفيو 1896– 1981 "Wechsler Bellevue Test" الذي وضعه ديفيد وكسلر الإخصائي النفسي في مستشفى بلفيو في محاولة لقياس الذكاء، بعد أن ظهرت عيوب مختلفة في اختبار "ستانفورد– بينيه للذكاء". ويتميز اختبار وكسلر الذي لقي استحساناً كبيراً في أوساط السيكولوجيين لوفائه بحاجتهم إلى قياس ذكـــاء الراشدين، بأنه يفيد في الاستخدامات العياديـــة، ويتضمن نوعين من الاختبارات، لفظية وأدائية، وله القدرة على التمييز بين القدرات التي يقيسها.

وفي مجال استخدام التحليل العاملـــي في قياس الذكاء والقدرات العقلية، بدأ (ثرستون) أبحاثه في مجال تحليل القدرات العقليـــة، فتوج جهوده بإصدار كتاب بعنوان: (التحليل العاملـــي المتعدد Multiple Factor Analysais) عام 1931. (أحمد، 1960، ص.57–58).

وفي أعقاب الحرب العالمية الثانية، تطور القياس النفســـي وتقدم بشكل ملحوظ، وحدث ذلك كنتيجة لفعالية وأهمية الاختبـــارات والمقاييس النفسيـــة ونجاحها في المهام التي أوكلت للسيكولوجيين خلال الحرب، ولم تقتصر الاختبارات على تقديم الخدمة النفسيـــة العملية، بل أسهمت بقدر واضح في إرساء أنساق نظرية

أساسية في مجال التنظير في أعمق معاينة في علم النفس، سواء في مجال القدرات العقليـــة أو في مجال السمات الشخصيـــة أو غيرها من مجالات علم النفس المتعددة. (أحمد، 1960، ص.67)

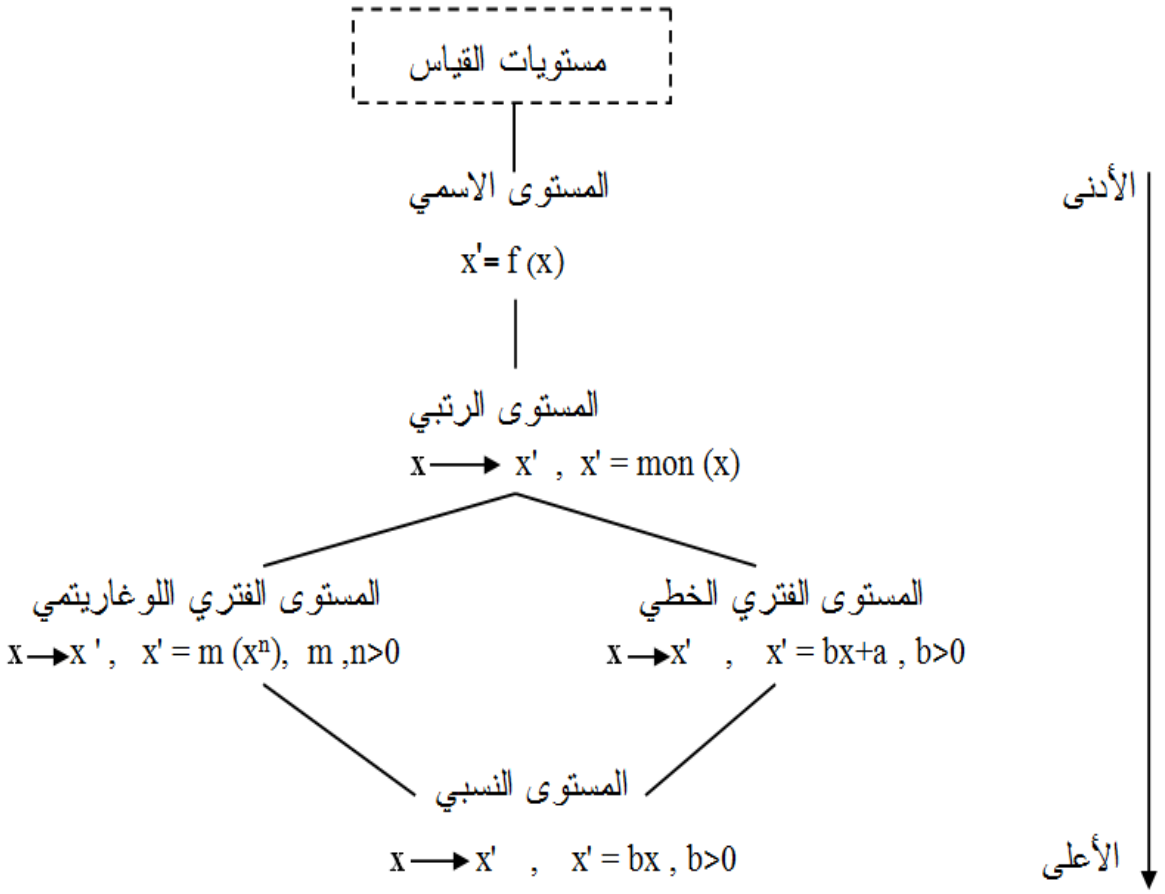
يتضح للمتأمل في التطورات التاريخية التي مر بها القيـــاس والتقويـــم النفسي والتربوي، بأنه حظي بمكانة خاصة لدى علماء النفس والتربية، نظراً لأهميته البالغة في عملية بناء وتصميم وتطوير الاختبارات والمقاييس النفسية والتربوية، لهذا انتشرت عديد المؤسسات العالمية، والتي أوكلت لها مهمة تطوير وتقنين هذه الأدوات، لاستخدامها في الأغراض البحثية المختلفة في حقول علم النفس والتربية.

2- مفهوم القياس النفســـي والتربوي:

إذا كان التقويـــم يعرف بتوفير معلومات موضوعية لأجل إصدار حكم، فإن القياس هو التعبير الكمي لهذه المعلومات الموضوعية بلغة رقمية أو حسابية بمقدار معياري (بوحدة قياس ما) تعبر عن الظاهرة المقاسة، أو بتعبير أخر القياس هو عملية (تكميم) أو تعبير بلغة كمية أو حسابية عن صفات أو عوامل أو ظواهر لموضوعات نوعية أو معنوية أو سلوكية تتطلب إصدار حكم أو تقويم عنها. فالقيـــاس التربـــوي أو النفســـي لا يكون مباشراً كما في العلوم الطبيعية، إنما هو استدلالي شأن أغلب الظواهر السلوكية. وهذا يتطلب أساليب متطورة ودقيقة للقياس تتوافر فيها صفات الصدق والثبات والقدرة على التمييز في المستويات الافتراضية. وتقاس سمات السلوك كالذكاء، والقدرات الفعلية المختلفة أو الميول والاتجاهات أو أي سمة من سمات الشخصية قياسا استدلالياً غير مباشر بافتراضات لوحدات سلوكية يعبر عنها في الغالب بمفردات الاختبارات النفسية لأي من مظاهر السلوك، كلما ازداد تعقد عملية القياس، ارتفعت احتمالات الخطأ، وهذا يعني أن للقياس مهما كان نوعه عوامل تساعد على الخطأ أو تكون مصدراً له، ولكن نسبة الخطأ في العلوم التربويـــة والنفسيـــة هي أكثر مما في العلوم الطبيعية.(سوسن، 2014، ص.25)

3- مستويات القياس النفســـي والتربوي:

إذا كان القياس هو تعيين أعداد للسمات أو الخصائص تبعاً لقواعد معينة معروفة، فالصياغة العامة لهذه القواعد وما يقابلها من مستويات القياس، جاء بها ستيفنز (1946-1951) Stevens وفق مجموعة من القواعد يشار إليها بمستويات القياس (بوسالم، 2014، ص.23). وعلى الرغم من تعدد هذه المستويات، إلا أن هناك مستويات رئيسة اتفق عليها علماء القياس النفســـي والتربوي، يوضحها الشكل رقم(01):



شكل (01) يوضح مستويات القياس الرئيسة.

يوضح الشكل رقم(01) امكانية تصنيف مستويات القياس في خمسة أقسام رئيسة تراكمية، حيث أن أي عملية إمبريقية تميز أحد الأقسام الأعلى ينبغي أن تكون متضمنة في جميع العمليات التي تميز الأقسام الأدنى، وكل من هذه الأقسام يمثل مستوى تكميم مختلف للمتغير موضع القياس، ويسمح بعمليات حسابية وأساليب إحصائية مختلفة. وفي ما يلي توضيح لهذه المستويات بشئ من التفصيل: (علام، 2013 ، ص.17)

3-1- **المستوى الإسمي (Nominal Level):**

وهو أدنى مستويات القيـــاس من حيث مدى ملاءمة الاجراءات الحسابية المعروفة ومدى تطبيقها فيه، لأن الأعداد (الأرقام) تستخدم فيه لتعيين فئات ينتمي إليها الأشخاص أو الأشياء فقط، ولذلك فإن الأعداد لا تعد دالـــة على كميات من خصائص.

3-2- **المستوى الرتبي (Ordinal Level):**

يمكن تنظيم الوحدات في هذا المستوى في سلسلة تمتد بين الأدنى والأعلى في الخاصية التي نقيسها، لكننا لا نستطيع أن نحدد بدقة الفرق بين أي رتبتين، ويتم تحديد مرتبة الشيء أو مكانته في مقياس يقدم وضعا كيفياً مثل قليل أو كثير، كبير أو صغير، الأول والثاني و الأخير.(سوسن، 2014، ص.26)

3-3-1- **المستوى الفتري الخطي (Linear Interval Level):**

يعد هذا المستوى أكثر المستويـــات استخداماً في القيـــاس في العلوم السلوكية. إذ أنه يفي بخصائص المستويين الإسمي و الرتبي، إضافة إلى توافر وحدات أو مسافات متساوية بين النقط على المتصل الخطـــي للسمة المقاسة. وهذا يجعل من الممكن المقارنة بين المسافـــات المتتالية على هذا المتصل، وإجراء عمليتي الجمع والطرح على هذه المسافـــات. غير أنه لا يوجد صفر حقيقي أو مطلق على هذا المتصل، حيث أن نقطة الصفر تكون اعتباريـــة و يتم تعيينها بين الباحثين، لذلك لا يسمح المستوى الفتري الخطـــي بإجراء عمليتي القسمة والضرب على الدرجات الناتجة. ولعل درجات الحرارة تعد مثالاً شائعاً لاستخدام هذا المستوى في القياس الفيزيائي.

3-3-2- **المستوى الفتري اللوغاريتمـــي (Logarithmie Interval Level):**

نظراً لأن القيـــاس يختص بتعيين أعداد أو رموز رقمية وفقا لقاعدة معينة، فإن العلماء يمكنهم ابتكار قواعد جديدة تختلف عن القواعد التي استندت إليها مستويات القيـــاس الثلاثة السابقة. إذ يمكن أن يستند أحد هذه المستويـــات على العمليات الإمبريقية الثلاث، وهي: تحديد التماثل أو التكافؤ، وتحديد الأكبر من أو الأصغر من، وتحديد النسب المتساوية. أما العملية الإمبريقية المتعلقة بتحديد تساوي الفترات فإنه يفترض عدم توافرها، وفي هذه الحالة يمكننا تحديد قيم معينة مناظرة لإحدى الخصائص أو السمات، ثم نقوم بترتيب هذه القيم والربط بينها بعلاقة التناسب رقم(01):

$$\frac{a}{b} = \frac{b}{c} = \frac{c}{d} \quad \text{.....(01)}$$

حيث تشير الرموز a، b، c، d، ... إلى نقط متتالية على متصل أو ميزان السمة المقاسة، ويمكن تحويل هذه العلاقة إلى الصيغة اللوغاريتمية رقم(02):

$$\log a - \log b = \log b - \log c = \log c - \log d = \ldots\ldots\ldots\ldots (02)$$

وهذه الصيغة تمثل ميزان القيـــاس الفتري اللوغاريتمـــي، فإذا لم نستطع تحديد فترات متساوية، فإننا نقتصر على نسب متساوية اعتبارية، مما يؤدي إلى الإقتصار على المستوى أو الميزان الفتري اللوغاريتمي. والأساليب الإحصائية التي يمكن استخدامها في القياسات التي تستند إلى الميزان الفتري اللوغاريتمي يمكن أن تشمل تلك التي تناسب الميزان الفتري الخطـــي. غير أن هذا يتطلب التعامل مع لوغاريتمات قيم الميزان بدلا من قيم الميزان في حد ذاتها، فعند إيجاد المتوسط مثلا ينبغي أن نوجد متوسط لوغريتمات قيم الميزان. فالكثير من المقاييس في علم النفس الطبيعي يعتمد على المستوى الفتري اللوغاريتمي، مثل التقدير أو الحكم الذاتي على شدة الإضـــاءة، ودرجة إرتفاع الصوت، وثقل الوزن، وغير ذلك. فقد تبين أن أحكام الفرد على الفترات المتساوية أو المختلفة تخضع لتحيزات منتظمة لا تتواجد في الأحكام على النسب المتساوية (علام، 2013، ص.20). ولعل نموذج راش اللوغاريتمي أحادي البارامتر أحد نماذج الاستجابة للمفردة الاختبارية التي يتناوله البحث الحالي يعد خطوة أساسية في هذا الشأن.

3-4- **المستوى النسبي (Ratio LeveI):**

يتميز هذا المستوى للقيـــاس بأن الفروق بين المسافات أو الفترات على متصل السمة المقاسة، والنسب بين الأعداد التى تمثل نقط هذا المتصل يكون لها معنى، وذلك لأن الميزان النسبـــي له صفر مطلق. ومن أمثلة ذلك في القيـــاس الفيزيائي الطول، والوزن، ويتميز هذا المستوى من مستويات القياس بجميع خصائص المستويات الأدنى السابقة، بالإضافة إلى توافر صفر مطلق، أي يناظر إنعدام السمة بالفعل، لذلك يعد أعلى مستويات القياس.

عموماً، ينظر تعريف ستيفنز إلى أن القيـــاس شبيهاً باستخدام مسطرة في قياس طول جسم، من حيث إن المسطرة تعني " قواعد " والقيم العددية المقترنة بها يعبر عنها في علامات الترقيم على المسطرة، غير أن تعريف ستيفنز يستدعي سوء تفسير، فبالرغم من أن المرء يمكن أن يستنتج من التعريف أنه يصف فعلياً أو عملياً، إلا أن هذا لا يبرر بشكل واضح في التعريف.

4- نظريات القياس النفســـي والتربوي:

تخضع نظريات القياس سواء في العلوم النفسيـــة أو العلوم التربويـــة لكثير من المفاهيم الرياضية والإحصائية في أساسها وفي إجراءاتها وفي طرق إستنتاج النتائج للتأكد من مطابقتها للنماذج النظرية الشائعة، وتأتي أهمية نظريات القياس في النقاط التالية:

(أ) تساعد في تحديد مصادر خطأ القياس ومستوى التأثير التي تحدثه هذه المصادر؛

(ب) تمثل الإطار الذي يتم من خلاله البحث عن العوامل التي تؤثر على الدرجة التي يحصل عليها المفحوص، ويكون لهذا الإطار مجموعة من الفروض وقد تتحقق هذه الفروض فتكون النتائج المبنية عليها واقعية وممثلة للحقيقة وقد تكون غير واقعية ومن ثم ستكون الاستنتاجات غير واقعية وخاطئة، ونظراً لأن جميع عمليات القياس يشوبها بعض الخطأ فإن العلماء يبحثون باستمرار عن طرق لزيادة دقة القياس ولتحقيق ذلك فإنهم كثيراً ما يحصلون على متوسط عدد مرات القياس التي تم تعريفها بواسطة عدة معايير. ويفيد هذا المتوسط كتقدير للقياس الحقيقي أو المثالي الذي يمكن أن نحصل عليه نظرياً من جميع حالات القياس لصفة من الصفات؛

(ج) نظريات القيـــاس تساعدنا في بناء أدوات القياس وفي تفسير الدرجات الناتجة عن التطبيق، فنجد أن علماء القيـــاس يهتمون بمدخلين رئيسيين في بناء أدوات القياس، أولهما المدخل التقليدي الذي يعتمد على النظريات الكلاسيكية في بناء الاختبارات، والثاني هو المدخل الحديث في بناء الاختبارات والذي يعتمد على نظرية الاستجابة للمفردة. وفي المجال النفســـي والتربوي يوجد العديد من النظريات التي استخدمت في تطوير أدوات القياس، إلا أننا سنقتصر على مناقشة النظرية الكلاسيكية للاختبارات في هذا الفصل، ونترك مناقشة نظرية الاستجابة للمفردة الاختبارية أو بما يسمى سابقاً نظرية السمات الكامنة إلى الفصل القادم من هذا البحث.

4-1- النظريـــة الكلاسيكيـــة للقيـــاس:

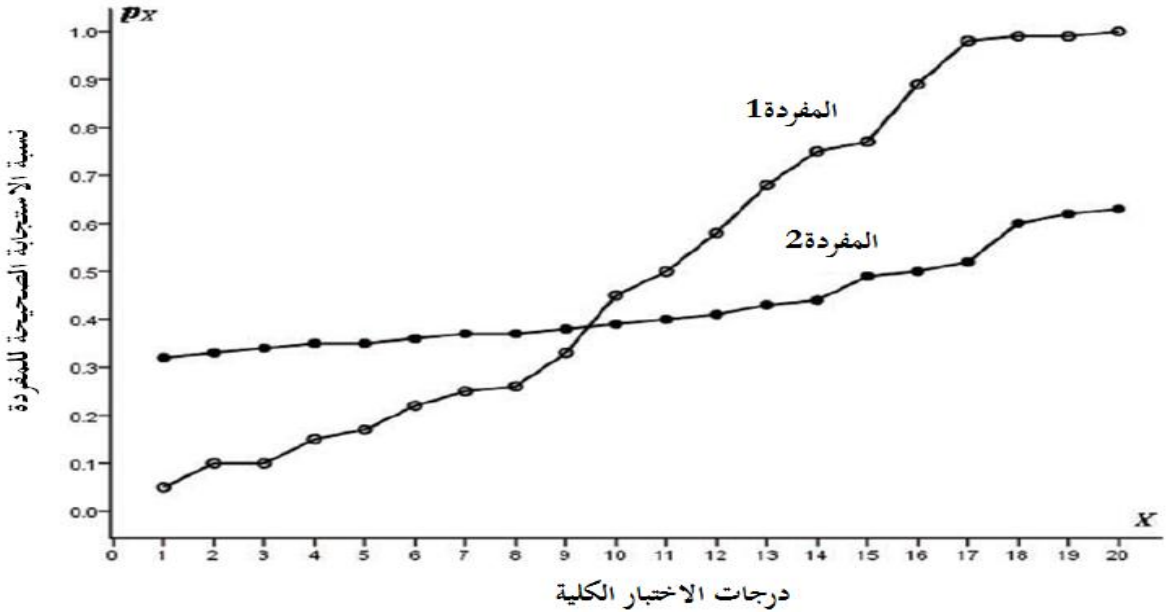
تعتبر نظرية القيـــاس بشكل عام نموذج يستخدمه الباحثون في مجال القياس بغرض تفسير العوامل الداخلة في التأثير على الدرجة التي يحصل عليها المفحوص في اختبار من الاختبارات، باعتبار أن الدرجة المحصل عليها ليس من الضروري أن تمثل قدرته أو المستوى الحقيقي للفرد في الصفة أو الخاصية المقاسة (الطريري، 1997، ص.34). وترجع الأسس الأولى للنظرية الكلاسيكية في القيـــاس إلى العالم جاليكسنا

(Gulliksen) عام 1950 الذي طورها واستخدمها على نطاق واسع في القياس النفسي والتربوي. حيث أطلق على عليها مسميات عديدة منها: نظرية الدرجة الحقيقية (The true Score Theory)، أو نظرية الخطأ في القياس (Theory Of Measurment Error)، أو نظرية الدرجة العددية الصحيحة، أو نظرية الاستجابة للاختبار (Test Respons). والتي تعد من أهم وأقدم الطرق المستخدمة في القيـــاس النفســـي والتربوي، وقد انتشر استخدامها في بناء وتطوير مختلف أنواع الاختبارات سواء النفسية، مثل: (اختبارات الذكاء أو الاستعدادات أو القدرات الطائفية أو الخاصة أو الميول أو الاتجاهات ومقاييس الشخصية وغيرها من الاختبارات النفسية)، أو الاختبارات التربوية التحصيلية، وتعتمد في تفسيرها للثبات على الصور المتكافئة للمقاييس المستخدمة (أمين، 2009، ص.68). ولغرض تفسير الدرجة من حيث صحتها من عدمه تضع هذه النظرية مجموعة من الافتراضات.

4-1-1- المسلمات و الافتراضات الأساسية للنظرية الكلاسيكية للقياس:

(1). الدرجة الملاحظة (س) في الاختبار هي عبارة عن مجموع جزيئين هما الدرجة الحقيقية (ح)، ودرجة خطأ القياس (خ)، ويمكن التعبير عليها بالمعدلة التالية: الدرجة الملاحظة = الدرجة الحقيقية + درجة خطأ القياس؛ (2). لا يمكن قياس الدرجة الحقيقية بل يمكن الاستدلال عليها أو تقديرها من خلال الدرجة الملاحظة، وذلك بحساب متوسط الدرجة الملاحظة أي أن الدرجة الحقيقية للفرد تساوي متوسط الدرجة الملاحظة التي يحصل عليها الفرد عندما يطبق عليه عدد لا نهائي من نسخ الاختبارات المتكافئة، وعندما يكون متوسط خطأ القياس يساوي صفراً؛ (3). التوزيع الاعتدالي لدرجات الأفراد في الاختبارات؛ (4). لا يوجد ارتباط بين الدرجات الحقيقية (ح) وأخطاء القياس للمفحوصين (خ) على الاختبار؛ (5). لا يوجد ارتباط بين أخطاء القياس لاختبارين مختلفين على نفس المشارك؛ (6). درجات الخطأ على الاختبار الأول (خ1) لا ترتبط بالدرجة الحقيقية على الاختبار الآخر (خ2)؛ (7). يمكن بناء اختبارين متكافئين (متوازيين) يقيسان نفس السمة بحيث تتساوى درجاتهما في المتوسط، والانحراف المعياري، والتباين، وصعوبة المفردات الاختبارية وتمييزها؛ (8). خطأ القيـــاس هو خطأ عشوائي يحدد دقة القياس أو ما يسمى بثبات الاختبار، بينما الخطأ المنتظم هو خطأ يتعلق بصدق القيـــاس أو صدق الاختبار؛ (9). تفترض هذه النظرية تساوي تباين أخطأ القيـــاس لجميع المشاركين الذين يطبق عليهم الاختبار أو أداة القيـــاس. (محاسنة، 2013، ص.95) (أمين، 2009، ص.69)

و ترى النظرية الكلاسيكية للقياس كذلك، بأن درجة الفرد في الاختبـــار أو المقيـــاس، هي دالـــة خطية مطردة، بمعنى كلما زادت درجة الفرد على الاختبـــار زاد مقدار وجود السمة لديه (Cronbach & Gleser, 1970, p.116). كما يوضحه الشكل رقم(02):



شكل (02) منحنيات خصائص المفردة لمفردتين وفقاً للنظرية الكلاسيكة في القياس.

يوضح الشكل رقم(02) تمثيل لدرجات الاختبـــار الكلية للممتحنين على محور السينات(X)، و نسبة الاجابة الصحيحة للممتحنين لمفردتين من مفردات الاختبـــار على محور الصادات (Px)، صمم الاختبار لتمييز الأداء العالي مع تحديد الدرجة كلية تزيد عن17 على هذا الاختبار (نلاحظ أن المفردة 2 تفرق بشكل جيد نسبياً بين الممتحنين عند الدرجة X >17). ومن خلال ملاحظة إنحدار منحنى خصائص المفردتين يمكننا تحديد تمييز وصعوبة المفردتين، أي أن الدرجة الكلية على الاختبـــار تقابل نسبة الإجابة الصحيحة 0.50، ما معناه (احتمال ما نسبته 50% للنجاح في الإجابة الصحيحة على المفردة).

اضافة لذلك يظهر الشكل رقم (02)، بأن الدرجة الكلية على الاختبـــار التي تقابل 0.5 للمفردة 1 هي الدرجة X=11، وبالتالي مؤشر الصعوبة القائم على منحنى خصائص المفردة هو $b1 = 11$. ومن ناحية أخرى فإن الدرجة الكلية على الاختبـــار التي تقابل 0.5 للمفردة 2 هي X=16، مما يشير إلى أنها مؤشر لصعوبة المفردة 2 أي $b2 = 16$. ونستطيع القول بأنها في المتوسط، ومنه فإن المفردة 2 أكثر صعوبة من المفردة 1 ($b2 > b1$). ومع ذلك، لا يمكننا القول أن المفردة 2 هي دائما أكثر صعوبة من المفردة 1. حيث يوضح الشكل رقم (02) أن منحنى خصائص المفردات للمفردتين يتقاطع تقريباً عند الدرجة X=9 ، أي أن

الممتحنين الذين درجاتهم الكلية أقل من الدرجة 9، فإن المفردة 1 تكون بالنسبة لهم أكثر صعوبة، أما الممتحنين الذين لديهم درجة كلية أكبر من 9، فإن المفردة 2 تكون بالنسبة إليهم هي الأكثر صعوبة. (Dimitrov, 2012, p.179)

يتضح للباحث أن النظرية الكلاسيكية تعتمد في الحكم على مؤشر صعوبة المفردات من خلال الدرجة الكلية للاختبار وليس كل مفردة على حدة، لهذا عند مقارنة تلميذين حصلا على نفس الدرجة الكلية في الاختبار بمفرداته السهلة والصعبة، هذا لا يعني أنهما يمتلكان نفس الكفاءة التي يقيسها الاختبار، لأنه ربما يكون أحد التلامذة أجاب على مفردات سهلة والآخر أجاب على مفردات صعبة، وهذا الإشكال المطروح تجيب عليه النظرية الحديثة في القياس أو بما يسمى نظرية الاستجابة للمفردة الاختبارية.

4-1-2- الخطـــأ المعيـــاري للقيـــاس:

يعد الخطأ المعياري للقياس من الخصائص السيكومترية الأساسية للاختبارات، حيث أنه يستند إليه في تفسير كل درجة من درجاتها. فالخطأ المعياري للقياس يعد عاملاً أساسياً في تقرير نتائج الاختبارات والمقاييس وتفسيرها، وهو مرتبط ارتباطاً وثيقاً بمفهوم الثبـــات ولكنه أيسر فهماً لمستخدمي الاختبارات والمختبرين، ولا يتأثر كثيراً بتشتت الدرجات على عكس الثبـــات الذي يعتمد على التشتت (التباين). وتعتبر النظرية السيكومترية للقياس أن الخطأ المعياري ثابت بالنسبة لمجتمع معين من الأفراد المختبرين. (علام، 2000، ص.176). ويمكن التعبير عن ثبات الاختبار من حيث الخطأ المعياري للقياس (SEM)، ويسمى أيضا الخطأ المعياري للدرجة. وهذا المقياس يناسب بدرجة جيدة تفسير الدرجات منفردة ولذلك فإنه بالنسبة لكثير من الأغراض الاختبارية يكون أكثر فائدة من معامل الثبات. لذلك إذا أردنا المقارنة بين ثبات "اختبارات مختلفة"، فإن معامل الثبات يكون هو المقياس الأفضل، ولتفسير "الدرجات منفردة"، فإن الخطأ المعياري للقياس يكون أكثر ملاءمة. فالدرجة الحقيقية ثابتة، والدرجة الملاحظة تتذبب حولها، فإذا افترضنا أن قيمة معامل ثبات درجات الاختبار تساوي 0.80، وقيمة الانحراف المعياري للدرجات الملاحظة 6.3 فإن الخطأ المعياري للقياس الذي يستخرج وفق المعادلة التالية: $SEM= SD\sqrt{1-R}$؛ يساوي 2.82. فإن حصل فرد في الاختبار على الدرجة 12 مثلا، فإنه يمكننا أن نستدل على أن درجته الحقيقية سوف تقع بين 12− 2.82؛ و 12+2.82، أي بين 9.18؛ و 14.82 وثقتنا في هذا الاستدلال 68% تقريباً. كذلك يمكننا أن نستدل بدرجة ثقة 95% تقريباً من خلال ضرب قيمة الخطأ المعياري للقياس في قيمة مجال مستوى الثقة

1.96 على أن درجته الحقيقية سوف تقع بين ±5.53، أي بين 6.46؛ و 17.53. كما يمكننا أن نستدل بدرجة ثقة 99% تقريباً كذلك، من خلال ضرب قيمة الخطأ المعياري للقياس في قيمة مجال مستوى الثقة 2.58 على أن درجته الحقيقية سوف تقع بين ± 7.27، أي بين 4.73 و 19.27. (علام، 2000، ص.176)

غير أنه لا يمكن افتراض أن معاملات الثبات أو أخطأ القياس تظل ثابتة عندما تباين مستوى القدرة تبايناً كبيراً، والفروق في معاملات الثبات تظل ثابتة عندما يتم حساب أخطاء القياس عند مستويات مختلفة لنفس الاختبار. وأحد الحلول الشاملة لهذه المشكلة تقدمها أساليب نظرية الاستجابة للمفردة الاختبارية في تحليل المفردات. وبمراعاة هذه الأساليب لمدى متسع من القدرة فإنها تقدم وسيلة للتعبير عن دقة قياس الاختبار كدالة لمستوى القدرة، وينتج هذا الإجراء "منحنى معلومات الاختبار" الذي يعتمد فقط على المفردات المتضمنة في الاختبار، ويسمح بتقدير خطأ القياس عند كل مستوى قدرة.(أناستازي، 2015، ص.142)

4-1-3- مميزات النظرية الكلاسيكية للقياس في بناء الاختبارات:

(أ). من السهل على غير المتخصصين في مجال القيـــاس النفسي والتربوي من تحليل بيانات الاختبارات، أي المعالجة الإحصائية لمفرداتها من خلال حساب السهولة والصعوبة، و معامل التمييز لمفرداتها، وإستيعاب المفاهيم المتعلقة بها؛ (ب). بساطة الافتراضات التي تقوم عليها النظرية الكلاسيكية للقياس والتي يمكن أن تلائم بيانات مختلف أنواع الاختبـــارات النفسية والتربوية، إضافة إلى أنها لا تحتاج لنماذج رياضية معقدة؛ (ج). يمكن تحليل بيانات الاختبـــارات وفق هذه النظرية دون الحاجة لبرامج أو أشخاص متخصصين.

4-1-4- أوجه قصور وعيوب النظرية الكلاسيكـــة للقيـــاس:

(أ). جميع الخصائص السيكومترية للاختبارات أي معامل الصدق و الثبات، وكذلك الخصائص السيكومترية لمفردات الاختبار مثل معامل الصعوبة و التمييز تعتمد في خصائصها على عينة الأفراد التي يطبق عليها الاختبار؛ (ب). تفترض النظرية أن درجات الاختبار والتي تمثل السمة أو القدرة المقاسة هي دالـــة خطية مطردة، بمعنى أنه كلما زادت درجة المشارك دل ذلك على زيادة السمة أو القدرة المقاسة، ولكن الواقع يشير إلى أن بعض الأفراد ذوي القدرة العقلية العالية يحصلون أحياناً على درجات منخفضة في الاختبارات والعكس صحيح؛ (ج). تفترض هذه النظرية أن الاختبارات تقيس متغير أحادي البعد، هذا الفرض صحيح في العلوم الفيزيقية لأنها تقيس شيئاً واحداً فقط مثل الطول، الوزن، الحجم، ..الخ. إلا أن اختبارات الذكاء قد تقيس متغيرات أخرى بجانب الذكاء مثل: ثراء البيئة، أو الإمكانيات المادية، والاجتماعية، وغيرها.

وبالتالي يصعب تقدير أحادية البعد في كثير من الاختبارات النفسية؛ (د) اعتماد الدرجة الكلية للاختبـــار على مفردات (أسئلة) هذا الاختبـــار؛ (هـــ) تغير تكوين الاختبار، ومعنى مفردات الاختبار بمرور الزمن؛ (و) يعتمد تعريف ثبـــات الاختبـــارات على مفهوم الاختبارات المتكافئة رغم أن مفهوم التكافؤ يصعب تحقيقه عملياً.

5- بناء الاختبارات التحصيلية وفقاً للنظرية الكلاسيكية للقياس:

اهتم علماء القيـــاس والتقويـــم التربوي بالتوصل إلى أساليب متعددة لقياس مدى تحقق الأهداف التعليمية وفعالية البرامج التربوية المقدمة، وكذلك لقياس مستوى التحصيل الدراسي للتلميذ والحكم عليه في ضوء أسس معينة، ولذلك ظهرت أنماط مختلفة في بناء الاختبارات التحصيلية وفق الغرض المطلوب منها:

5-1- الاختبارات مرجعية المعيار (المرجعية الجماعة) NRT أو الاختبارات السيكومترية:

والهدف من هذه الاختبارات هو تحديد مقدار المعلومات والمهارات التي سبق وأن حصل عليها التلميذ في موضوع معين. وتفسر درجة التلامذة في ضوء متوسط أداء الجماعة التي ينتمي إليها للحصول على تقدير لتحصيله الدراسي. وقد وجه إلى هذا النوع من الاختبـــارات الانتقاد بأنها لا تقيس الأهداف التعليمية التي يسعى المعلم إلى تحقيقها في حجرة الدراسة، كما أنها لا تفيد في تشخيص جوانب القوة والضعف في التحصيل الدراسي للتلميذ. (مراد وسليمان، 2005، ص.141)

5-2- الاختبارات مرجعية المحك (المرجعة للهدف) CRT أو الاختبارات الإيديومترية:

والهدف من هذه الاختبارات هو التعرف على مستوى التمكن أو الإتقان (Mastery) للأداء، وبالتالي فإن درجة التلميذ تفسر في ضوء محك (ميزان) سابق التحديد وليس متوسط أداء الجماعة التي ينتمي إليها، وبالتالي فهو يحدد مستويات الأداء المطلوب تحقيقها.

وبالعودة للاختبارات مرجعية الجماعة أو الاختبارات السيكومترية، والتي سوف نستعرض خطوات تصميم أو بناء اختبار تحصيلي موضوعي وفقها.

6- الاختبارات مرجعية المعيار (المرجعية الجماعة) NRT أو الاختبارات السيكومترية:

6-1- خطوات إعداد اختبـــار تحصيلي موضوعـــي مرجعـــي المعيـــار(Norm-Refrenced Test (NRT)): (مراد وسليمان، 2005، ص.143) (أمين، 2009، ص.237)

6-1-1- **تحديد الأهداف التعليمية/ التدريسية وتعيين الوزن النسبي لكل هدف:**

بعد الإنتهاء من دراسة مقرر معين نحدد الأهداف التي حظيت باهتمام وتدريب سابق في المجالات المعرفية، و الانفعالية، والنفسحركية، ثم نحدد الأهمية النسبية لكل هدف على حدة، ويمكن الإعتماد على آراء الخبراء في الميدان لتحديد الأهمية النسبية، والخبراء هم (معلم متميز له خبرة أكثر من 5 سنوات في التدريس أو معلم أول أو موجه للمادة أو أحد خبراء مراكز البحوث التربوية من ذوى العلاقة).

- **تحليل محتوى المقرر الدراسي السابق دراسته وتعيين الوزن النسبي لعناصر المحتوى:**

لا شك أن تحليل مكونات المقرر الدراسي أو تحليل مكونات العمل (المهارة) السابق إكتسابها أو تحليل المهمة وتحديد الموضوعات الرئيسية والفرعية التي أهتم بها المعلم أثناء عملية التدريس يسهل من مهمة إعداد الاختبار الصادق، وتحديد الأهمية النسبية لعناصر المحتوى الرئيسية والفرعية أو مكونات المهارة أو المهمة السابق دراستها يتم في ضوء واحد أو أكثر من المحكات التالية: عدد الصفحات التي يشغلها كل موضوع على حدة بالنسبة للعدد الكلي للكتاب (وهو محك غير جيد)، أو الزمن المستغرق أثناء عملية التدريس لكل موضوع من الموضوعات على حدة أو آراء الخبراء في مادة التخصص بشأن كل موضوع على حدة وهو أفضل المحكات. (أمين، 2009، ص.238)

6-1-2- **إعداد جدول مواصفات الاختبار التحصيلي:**

وهو عبارة عن مخطط تفصيلي ثنائي البعد أحد أبعاده قائمة الأهداف (نواتج التعلم المرغوب تحقيقها)، والبعد الثاني هو عناصر المحتوى التي يشملها الاختبار، ويمكن تلخيص فوائد إعداد جدول مواصفات الاختبار فيما يلي: (أ). يجبر واضع الأسئلة على توزيع أسئلة الاختبـــار على مختلف أجزاء المقرر الدراسي بحيث يعطي كل جزء من أجزاء المحتوى وزنه النسبي أي الأهمية النسبية عند صياغة الأسئلة؛ (ب). عدم تركيز الأسئلة لقياس مستوى محدد داخل المجال سواء كان (معرفي أو نفسحركي) بل يأخذ جميع المستويات المطلوبة تحقيقها بعين الاعتبار؛ (ج). شعور التلميذ الذي بذل جهداً في الاستذكار بأنه لن يضار من الاختبار؛ (د). شمول المحتوى في إعداد أسئلة الاختبار والذي يؤدي إلى صدق المحتوى.(مراد وسليمان، 2005، ص.146)

6-1-3- صياغة مفردات (أسئلة) الاختبار التحريري الموضوعي:

لا شك أن صياغة الأسئلة الجيدة شأنها شأن صياغة الأهداف (التعليمية/ التدريسية، والأهداف التقويمية، وأهداف القياس) هي عملية طويلة ومعقدة. ويجب على المعلمين وواضعي الأسئلة بمراكز البحوث التربوية ألا يقللوا من مهمة صياغة الأسئلة الجيدة سواء كانت من النوع الموضوعي أو غير الموضوعي، بل يجب عليهم أن يحسنوا من صياغة الأسئلة سواء من حيث الشكل أو من حيث المستوى. و ذلك من خلال التمكن من مادة التخصص، والمراجعة المستمرة للأسئلة التي أعدوها من قبل ذلك، وتقبل النقد من وزملاء التخصص، وإجابات التلامذة عن الأسئلة، كما يمكن الإستعانة بكتب القيـــاس للتعرف على الأشكال المختلفة من الأسئلة وشروط صياغتها، أضف إلى التدريب على بعض الأساليب الإحصائية اللازمة لتحليل أسئلة (مفردات) الاختبار للتعرف على عيوب تلك الأسئلة ومحاولة تعديلها. **وفي ما يلي أنواع الأسئلة التحريرية الموضوعية وأسس تصنيفها:**

1. **التصنيف حسب درجة تعقدها:** (أ). أسئلة الاختيـــار من بديلين (الصورة البسيطة، و التصحيحية، والعنقودية)؛ (ب). أسئلة المزاوجة (الصورة البسيطة، والتصنيفية)؛ (ج). أسئلة الاجابات القصيرة (أسئلة التكملة، التعيين، المماثلة)؛ (د). إعادة الترتيب؛ (ه). التفسيرية/حل المشكلات؛ (و). الاختيار من متعدد (صورها: اختيار الإجابة الصواب، الإجابة الخاطئة، أفضل إجابة، الترتيب الصحيح، الترتيب الخاطئ).

ب. التصنيف حسب العمليات العقلية المتضمنة أثناء الإجابة: (أ). أسئلة انتقاء الاستجابـــة مثل: (أسئلة الاختيار من بديلين، المزاوجة، إعادة الترتيب، الاختيار من متعدد)؛ (ب). أسئلة إنتاج الاستجابة مثل: أسئلة الإجابات القصيرة (التكملة، التعيين، المماثلة). (أمين، 2009، ص.246)

6-1-4- إخراج الورقة الامتحانية:

يشتمل إخراج الورقة الامتحانية على الخطوات التالية:

أ. شكل الورقة الامتحانية (أي نظام عرض الأسئلة):

يفضل أن ترتب الأسئلة داخل الورقة حسب درجة تعقد النشاط العقلي اللازم للإجابة عليها، بحيث يتم تجميع الأسئلة المتشابهة الصياغة معاً، وتوضع لكل منها تعليمـــات خاصة بها، مما يساعد التلميذ على الاحتفاظ بحالة التهيؤ العقلـــي أثناء الإجابة.

ب. صياغة الأسئلة:

يتضمن إخراج الورقة الامتحانية حسن الصياغة والتحديد الجيد للمفردة أو السؤال ووضوح الخط وخلوه من الأخطاء الإملائية أو الأخطاء العلمية.

جـــ. تعليمات الاختبار:

تهدف تعليمات الاختبار إلى شرح طريقة الإجابـــة على المفردة أي طريقة تسجيل الإجابة والزمن المسموح به للإجابة، وأن نحذر التلامذة من التخمين إذا كان هناك تصحيح للدرجة من أثر التخمين. (مراد وسليمان، 2005، ص.205)

6-1-5- **تحليـــل مفردات الاختبار(التجريب والمعالجة الإحصائيـــة):**

بعد الإنتهاء من إعداد الورقة الامتحانية (يمكن إعداد صور متكافئة)، يتم تجربتها من خلال التطبيق على عينة من الأفراد ممثلة للمجتمع الذي سوف يطبق عليه الاختبار فيما بعد لتحديد صدق الاختبـــار (التأكد من قياس ما وضع لقياسه)، وثبات الاختبار (يعطي نفس النتائج إذا ما أعيد تطبيقه على نفس العينة تحت نفس الظروف) والتعرف على خصائص كل سؤال من أسئلة الاختبار. وتتلخص المعالجة الإحصائية لمفردات الاختبار في تحديد مستويات السهولة والصعوبة لكل مفردة؛ و تحديد معامل التمييز لكل مفردة؛ والتحقق من فعالية المشتتات(في حالة الاختبارات التحصيلية الموضوعية). (أمين، 2009، ص.312)

6-1-6- **الاختبار في صورته النهائيـــة:**

بعد إجراء المعالجات الإحصائية لمفردات الاختبار واستبعاد المفردات السهلة والصعبة والحفاظ على المفردات التي تتراوح معامل سهولتها بين 3 و 7 أي ما بين 30% و 70%، وكذلك استبعاد المفردات غير المميزة بين التلامذة والتي يقل مؤشر تمييزها عن 2، والاحتفاظ بالمفردات التي معامل تمييزها 3 فأكثر، وأخيراً الاحتفاظ بالمفردات ذات المشتتات الفعالة التي تجذب أفراد المجموعة الدنيا أكبر من المجموعة العليا. و إذا كان لدينا مجموعة كبيرة من المفردات فيمكن عمل اختبارين متكافئين، ويصبح كل اختبار جاهز للتطبيق النهائي. (مراد وسليمان، 2005، ص.211-221)

6-2- **أنواع مفردات(أسئلة) الاختبارات التحصيلية:**

تنقسم مفردات الاختبارات التحصيلية إلى نوعين، هما:

6-2-1- مفردات الاختبارات المقالية:

وهي الأكثر شيوعا في ظروفنا التربوية والأطول تاريخاً وقد تكون أهم وسيلة تستعمل لتحديد مستوى تحصيل التلميذ، ولهذه الطريقة إيجابيات وسلبيات كما لها شروط خاصة لتحسين إستعمالها في قياس تحصيل التلامذة.

6-2-1-1- قواعد صياغة الأسئلة المقالية:

لصياغة الأسئلة المقالية يجب إتباع مجموعة من القواعد، أهمها: (أ). تدرج الأسئلة في الصعوبة إذ ينبغي أن تكون الاختبارات مقياسا متدرجاً من السؤال السهل إلى الصعب إلى الأصعب؛ (ب). تصمم الأسئلة بشمولية لتحقيق العدالة في توزيع الأسئلة على أجزاء المادة المقصود اختبار التلامذة فيها؛ (ج). يرتبط السؤال بأهداف المقرر؛ (د). يجب تحديد المشكلة؛ (هـ). يجب ذكر تعليمات السؤال بوضوح؛ (و). بدأ جملة سؤال المقال بكلمات أو عبارات مثل (قارن- انقد- ميز- بين- حلل)؛ (ز). تجنب استخدام كلمات مثل (من- متى- ماذا)، لأنها تتطلب إستدعاء المعلومات فقط وبخاصة إذا كان المتعلم قد سبق له دراسة الموضوع؛ (ح). تأكد من أن السؤال عن أداء هو نفس الأداء الذي تريد بالفعل من التلميذ أن يظهره؛ (ط). عدل طول الإجابة المطلوبة، ودرجة التعقيد فيها بما يتلاءم مع مستوى النضج عند التلميذ؛ (ي).عدم وضع أسئلة اختيارية إلا إذا كان الناتج التعليمي يتطلب ذلك؛ (ك). وضوح الصياغة وعدم استخدام كلمات غامضة. (ميمي، 2014، ص.30) (سوسن، 2014، ص.245)

6-2-1-2- الخصائص الإيجابية للاختبارات التحريرية (المقالية):

من أهم ما يميز الاختبارات المقالية أنها: (أ). تسمح للتلميذ بأن يتأمل و ينظم أفكاره دون إحراج، أو الاثارة التي قد يسببها موقف الإمتحان الشفهي، كما تسمح له بأن يغير ويصحح ويضيف بوقت أكبر مما هو في موقف الامتحان الشفهي؛ (ب). تنمي القدرة الفكرية للتلميذ على التعبير وتنظم أفكاره وفهمه للموضوعات؛ (ج). أقل كلفة في الوقت وتتطلب وقتا أقل في الإعداد؛ (د). احتمالات الغش فيها أقل من احتمالاته في الاختبارات الموضوعية (هـ). تكشف عن قدرة التلامذة على النقد والتقويم. (سوسن، 2014، ص.243)

6-2-1-3- **سلبيات وعيوب الاختبارات التحريرية (المقالية):**

كما أن لهذا النوع من الاختبارات عيوب يمكن تلخيصها في، ما يلي: (أ). تفتقر إلى صدق المحتوى وثبات الفكرة وقد تكون مركزة في جانب من جوانب المنهج دون آخر؛ (ب). يتطلب تصحيحها وقتا طويلا مع خبرة موحدة في تقدير الدرجات الأمر الذي قد يؤثر على مصداقية التقويم؛ (ج). إبراز عامل الحظ والتوقع في إجابات التلامذة؛ (د). تؤدي إلى عدم فهم المطلوب من السؤال وتوقع إجابات متعددة؛ (هـ). تتفاوت إجابات التلامذة لعامل المهارة اللغوية وحسن التعبير؛ (و). تسبب الإرهاق والضجر للتلميذ مما يؤثر على إجابته؛ (ز). ضعف ثبات وموضوعية التصحيح، حيث تتأثر بعوامل عديدة أثناء التصحيح منها عوامل ذاتية للحالة النفسية للمصحح، أو بعوامل تخص أوراق الإجابة كحسن الخط والتنظيم؛ (ح). ضعف صدقها إذ يحتمل أن تقيس المفردات معلومات تعتمد على الذاكرة فقط بينما يكون الهدف عمليات عقلية أخرى. (سوسن، 2014، ص.243) (ميمي السيد، 2014، ص.30)

6-2-2- **المفردات الموضوعية:**

تتألف الاختبارات الموضوعية من عدد من المفردات (الأسئلة) تصمم بحيث يكون لكل منها إجابة واحدة صحيحة، وعلى المفحوص أن يتذكرها أو يتعرف عليها من بين عدة إجابات محتملة. فنوع المفردة المعتمدة على نظرية الاستجابة للمفردة هو نوع من أنواع مفردات الاختبارات التحصيلية الموضوعية التي تحقق مطالب الموضوعية في القياس وتتوافق درجات التلامذة مع خصائص مفردات الاختبار، وذلك من خلال الاعتماد في بنائها على نظرية الاستجابة للمفردة، حيث تتمكن مثل هذه الاختبارات من قياس القدرات الفعلية للتلامذة.

والهدف الأساسي من التعرف على الأنواع المختلفة للمفردات هو أن يتاح للتلميذ والمعلم، وواضعي المفردات فرصة الألفة بتلك الأنواع وأن يكون واضع المفردات لديه مهارة صياغة المفردات والتدريب على إعداد نماذج أخرى على شاكلتها، والتعرف على المواصفات الفنية المطلوبة في كل نوع على حدة، حتى يتمكن من عمل بنـــك أسئلة جيد الصياغة يستخدمه فيما بعد في إعداد الورقة الامتحانية حسب الهدف من الامتحان الذي قد يحدده المعلم أو المؤسسات التعليميـــة أو صانع القرار التربوي.(ميمي السيد، 2014، ص.30)

6-2-2-1- أنواع أسئلة (مفردات) الاختبـــارات التحصيلية الموضوعية:

1. أسئلة الاختيـــار من بديلين:

يطلق على أسئلة الاختيار من بديلين أسماء عدة من بينها أسئلة نعم/لا، أسئلة صح خطأ، أسئلة موافق/ غير موافق. يقيس هذا النوع من المفردات مستوى المعرفة (التذكر) سواء كان تذكر الخصوصيات مثل المصطلحات، والحقائق، والمفاهيم، والقوانين والنظريات والتواريخ والأماكن، أو تذكر طرق ووسائل تناول الخصوصيات المحكات التي تستخدم في الحكم على صحة الآراء، أو تكرار التعميمات، كذلك تقيس مستوى الفهم / الاستيعاب سواء من خلال الترجمة من صورة إلى أخرى أو من مستوى إلى مستوى آخر أو التفسير أو التنبؤ وإدراك العلاقات، وتكون في ثلاثة أشكال، وهي:

(أ). الصورة البسيطة؛ (ب). الصورة التصحيحية؛ (جـــ). الصورة العنقودية – التجميعية.

2. أسئلة (مفردات) المزاوجة / المقابلة – المطابقة:

يطلق على أسئلة المزاوجة أسماء عدة بينها أسئلة المقابلة، وأسئلة المطابقة. تقيس هذه الأسئلة (المفردات) مستوى المعرفة (التذكر)، سواء كان تذكر الخصوصيات (المصطلح، الأحداث، التواريخ، الأماكن)، أو تذكر طرق ووسائل تناول الخصوصيات مثل (تذكر المحكات التي تستخدم للحكم على صحة الآراء)، أو تذكر تعميمات مثل تذكر أسس التعلم، مبادئ إكتساب اللغة. وتكون على الصور التالية: (أ). الصورة البسيطة؛ (ب). الصورة التصنيفية؛ (ج). أسئلة إعادة الترتيب؛ (د). أسئلة الإجابات القصيرة.

3. الأسئلة التفسيرية (أسئلة حل المشكلات):

يتمثل التفسير في قدرة الفرد على الشرح أو التلخيص أو الترجمة والتفسير هو إحدى العمليات العقلية الضرورية لحدوث التفكير بجميع صوره (الابتكاري– الناقد، التقاربي، الاستدلالي)، يشيع استخدامها في مجال الرياضيات والعلوم.

4. أسئلة (مفردات) الاختيار من متعدد:

تعد أسئلة الاختيار من متعدد من أهم وأفضل أنواع الأسئلة الموضوعية، لأنها تشمل على معظم أنواع الأسئلة الأخرى، بمعنى يمكن صياغة الأنواع الأخرى من الأسئلة الموضوعية (الاختيار من بديلين– المزاوجة– إعادة الترتيب، الإجابات القصيرة، التفسيرية) في صورة الاختيار من متعدد، كما يمكن قياس جميع المستويات العقلية إذا أحسن صياغتها. (أمين، 2009، ص.245–279)

أ. مميزات وعيوب أسئلة (مفردات) الاختيار من متعدد:

المميزات	العيوب
1. يمكن أن تقيس جميع المستويات العقلية المعرفية التي اقترحها بلوم إذا أحسن إعدادها. 2. سهلة في تصحيحها ويمكن التحكم في مستوى سهولة وصعوبة الأسئلة عن طريق زيادة التشابه بين البدائل. 3. يقل فيها التخمين إلى أقل حد ممكن. 4. تغطي جزء كبير من المحتوى المقرر الدراسي. 5. توفر للمعلم وسيلة جيدة للتشخيص في مجال التحصيل الدراسي، أي تحديد نقاط القوة والضعف لدى التلامذة. 6. درجة الصدق والثبات مرتفعة بالمقارنة مع باقي أنواع الأسئلة الموضوعية. 7. تساعد في تنمية قدرة التلميذ على حل المشكلات من خلال تدريبه على التمييز بين الحلول المقترحة، ثم إتخاذ القرار المناسب .	1.لا تصلح لقياس مخرجات التعلم التي تتعلق بالتعبير الكتابي والمتمثلة في: التأليف، والتنظيم وتوليد الأفكار(الابتكار) وهذه المخرجات يصلح لقياسها أسئلة المقال بنوعيها. 2. صعوبة إعدادها وذلك لصعوبة الحصول على مشتتات على درجة عالية من الجودة. 3. تحتاج إلى وقت في إعدادها وكذلك وقت الإجابة عليها وخاصة إذا كانت الأسئلة تتطلب تمييزا دقيقا بين البدائل. 4. مكلفة ماديا نتيجة استهلاك كمية كبيرة من أورق والطباعة وكذلك أجور الطباعة. 5. الغش فيها أسهل بالقياس إلى أسئلة المقال. 6. لا تقيس قدرات التفكير الابتكاري.

ب.شروط صياغة أسئلة (مفردات) الاختيـــار من متعدد:

1. الشروط التي تتعلق برأس السؤال:

ومن بين أهم الشروط التي تتعلق برأس السؤال، نذكر مايلي: (أ). يجب أن تتعلق المشكلة التي يطرحها السؤال باحدى المخرجات الهامة وليست بمشكلة هامشية؛ (ب). يجب أن يحتوي أصل السؤال على مشكلة واضحة ومحددة تماماً بحيث يستطيع التلميذ أن يفهمها دون اللجوء إلى قراءة البدائل؛ (ج). إذا كان أصل السؤال على هيئة عبارة ناقصة يجب أن يضاف إليها كل الكلمات المتكررة في البدائل بمعنى تجنب تكرار الكلمات في كل بديل وذلك بوضع الكلمات المكررة في متن (رأس) السؤال حتى يسهل على التلميذ اختيار البديل المناسب؛ (د). تجنب الارتباطات اللفظية بين رأس السؤال (المتن) و الإجابات الصحيحة، أي عدم وجود تلميحات (إشارات) مقصودة بين المتن والبدائل لأن هذه التلميحات تساعد التلميذ على اختيار الإجابة

الصحيحة دون أن يكون ملما بالسؤال، (هـ). تجنب صياغة النفي كلما أمكن ذلك، وإذا استخدم النفي يجب وضع خط تحت الكلمة الدالة على النفي لينبه التلميذ على ذلك، ويستحسن أن يجئ النفي في نهاية المفردة المكتوبة بصيغة موجبة تميل إلى قياس النواتج التعليمية بشكل أفضل من تلك المكتوبة بصيغة سالبة؛ (و). إذا كان السؤال متعلقا بتعريف مصطلح معين فمن الأفضل وضع المصطلح في متن السؤال وإعطاء التعريف ضمن البدائل. (مراد وسليمان، 2005، ص.196) (أمين، 2009، ص.304)

2. الشروط التي تتعلق بالبدائل:

من أهم هذه الشروط: (أ). يجب أن يكون هناك إجابة واحدة صحيحة أو أفضل الإجابات؛ (ب). يجب أن تكون البدائل متجانسة في محتواها ومرتبطة بمجال المشكلة بحيث لا يتم استبعاد أي بديل، مع ملاحظة أن صعوبة السؤال تزداد بزيادة التجانس بين البدائل؛ (ج). يجب أن تتوزيع الإجابات الصحيحة توزيعاً عشوائياً داخل البدائل حتى لا يكتشف التلميذ نظام الترتيب ومنه يتعرف على موقع الإجابة الصحيحة؛ (د). يجب أن تكون كل البدائل في نفس الطول قدر الإمكان حتى يصعب تمييز الإجابات الصحيحة عن الخاطئة؛ (هـ). يجب أن يتراوح عدد البدائل بين 3-5 حسب طبيعة المرحلة العمرية وبنية المادة حتى تقلل من أثر التخمين. (كلما زاد عدد البدائل قل التخمين والعكس صحيح)؛ (و). تجنب استخدام المحددات الشائعة مثل: جميع الإجابات صحيحة/ خاطئة، لا شيء مما سبق، دائما، أبدا، ولا نلجأ إليها إلا إذا كانت هناك صعوبة في وضع عدد كاف من البدائل أو عندما نريد تحديد صحة الإجابة أو خطئها بدقة. (مراد وسليمان، 2005، ص.197-203)

7- الاختبارات مرجعية المحك (المرجعة للهدف) CRT أو الاختبارات الإيديومترية:

تقوم الاختبارات ذات المرجعية المحكيـــة Criterion-referenced tests، على عكس الاختبارات ذات المرجعية المعياريـــة Norm-referenced ، على أساس مقارنة الإجابات مع مجموعة من الأهداف التعليمية المراد تحقيقها، وليس على أساس مقارنة الإجابات بعضها مع بعض.(لويس، 2007، ص.91)

7-1- الاختبارات محكيـــة المرجع من منظور تاريخي:

بظهور مفهوم التعلم من أجل الإتقـــان لم يعد هدف القيـــاس التركيز أساساً على الفروق بين الأفراد والتمييز بينهم. ظهرت المناداة بالإبتعاد عن شكل الناقوس الذي يميز التوزيع الاعتدالي الذي اعتمدت عليه المقاييس الجماعية – المرجع في التمييز بين الأفراد على الأداء. وكانت الحجة في ذلك أن النشاط التربوي

نشاط مقصود، يبذل بهدف أن يتقن التلامذة ما تعلموه، ولا ينبغي أن يخضع توزيع الأداء لما تخضع له المتغيرات الطبيعية كالطول والوزن (كاظم، 1988أ، ص.18). فالانتقادات التي وجهت للقياس المعياري المرجع دفعت بعض الباحثين إلى التفكير في أساليب جديدة لبناء المقاييس النفسيـــة والتربويـــة، ومن أبرزها مقال روبرت جليزر (R.Glaser) عام (1963) الذي نشر في الدورية الأمريكيـــة (American the psychologis) بعنوان: تكنولوجية التعليم وقياس نواتج التعلم، والذي ميز فيه بين نوعين من القياس أحدهما القياس معياري المرجع، يهتم بترتيب الفرد بالنسبة لأقرانه في القدرة التي يقيسها الاختبار. والآخر محكي المرجع يهدف إلى موازنة أداء الفرد بمستوى أداء معين وفق محك يحدد مسبقا أو أهداف واضحة يبنى الاختبار من أجل قياسها، بغض النظر عن موقع الفرد ضمن الجماعة.

فالفضل يعود إذن إلى جليزر في إثارة الإهتمام بالقياس محكي المرجع، حيث أثار مقاله مناقشات بين علماء القياس عامـــة والمتخصصين في تطبيقات تكنولوجية التعليم خاصة، ولكن لم يحدث نشاط علمي ملحوظ نحو تحقيق ما نادى به جليزر إلا بداية من سنة (1969) حسب علام (2005)، حيث بدأ جيمس بابام (J.Popham) عالم النفس الأمريكي بجامعة كاليفورنيا بتزعم حركة القياس المحكي بهدف تجسيده في واقع تعليمي فعلي، فدعى إلى مؤتمر متخصص في ولاية مينابوليس الأمريكية لمناقشة القضايا والمشكلات السيكومترية المتعلقة بهذا المفهوم الجديد في ميدان القياس النفســـي. (بوسالم، 2014، ص.160)

7-2- الاختبار التحصيلي مرجعـــي المحـــك:

الاختبارات التحصيلية مرجعية المحك، هي تلك الاختبارات التي تقيس أداء الأفراد في ضوء مستوى التمكن المطلوب من الأداء أي أنها تشير إلى درجة تمكن الفرد من مجموعة من الأهداف السلوكية المرتبطة بعينة من السلوك، و بذلك فهي تقيس مستوى أداء الأفراد في ضوء مستوى أداء محدد سلفاً. ويصمم الاختبار مرجعي المحك بناء على مجموعة من نواتج التعلم تعريفاً جيداً ضمن محتوى تعليمي معين. ونشأت فكرة الاختبارات المرجعية إلى محك نتيجة رغبة الأخصائيين في الاختبارات في معرفة وسائل جديدة تمكنهم من الوصول إلى معلومات أكثر عن مستوى أداء الأفراد موضوع القياس. وتذكر مارتن (Martin, 1975) أن الاختبارات التحصيلية مرجعية المحك هي اختبارات تعتمد على موضوعية الأداء ويتم تصميمها لكي تسمح بتحديد ما إذا كان المتعلم أتقن الهدف أم لا.(ميمي السيد، 2014، ص.09). وتستخدم اختبارات الإتقان أو المحكية في برامج التعليم الإفرادي أو نموذج لإتقان التعلم (Blomm, 1968). إذ تتألف هذه البرامج من

وحدات، أو وحدات قياسية ينظر إليها على الأغلب بشكل هرمي، ويعتمد كل منها على هدف سلوكي واحد أو أكثر، ويطلب من التلميذ العمل في الوحدة (التعليمية) حتى يحقق مستوى تخصص أدنى من التحصيل. ثم يعتبر ذلك الفرد متقناً ومسيطراً على تلك الوحدة التعليمية. وفي مثل هذه البرامج لا تعتمد القرارات التعليمية حول التلميذ على مقارنة أدائه مع أداء غيره، أما إذا أدى بشكل دقيق تجاه الأهداف، يقتضي القرار إنتقاله إلى الوحدة الدراسية التالية، أما اذا لم يفعل، فعليه إذن أن يعيد دراسة المادة التي غطاه الامتحان حتى يؤدي أداء دقيقا، وهذا يعني إتقانه للمادة. و يتطلب اختبار الإتقان من القائم على الاختبار أن يضع درجة فاصلة، أي يجب أن يكون هناك أساس منطقي دقيق ومنهج لاختيار النقطة. (النبهان، 2013، ص.459)

7-3- أوجه الاختلاف بين الاختبـــارات التحصيلية مرجعية المحـــك و مرجعية المعيار:

يوجد عديد الاختلافات بين الاختبارات مرجعية المعيـــار والاختبارات مرجعية المحك والتي يمكن أن تقدم على النحو التالي:

أ. من حيث الهدف الرئيس للاختبار:

يهدف الاختبار مرجعي المعيار إلى مقارنة أداء التلميذ بأداء مجموعته المعيارية من خلال تحديد عدد الأسئلة التي أجاب عليها التلميذ بشكل صحيح، أما الاختبار مرجعي المحك يهدف إلى مقارنة أداء التلميذ بمحك محدد مسبقاً وذلك لمعرفة الأهداف التي حققها والتي لم يحققها التلميذ.

ب. من حيث استخدام الاختبار:

يستخدم الاختبار مرجعي المعيار لإظهار الفروق الفردية بين التلامذة، أما الاختبار مرجعي المحك يستخدم للتأكد من تحقيق التلميذ للأهداف السلوكية المحددة.

ج. من حيث خصائص أسئلة الاختبار:

تنتشر الأسئلة في الاختبار مرجعي المعيار بشكل واسع حول نطاق الأهداف، بينما تتجمع الأسئلة في الاختبار مرجعي المحك حول عدد محدد من الأهداف.

د. من حيث تفسير الأداء:

تفسر درجة التلميذ في الاختبار مرجعي المعيـــار بناءً على درجات معياريـــة، بينما تفسر الدرجة في الاختبار مرجعي المحك بناءً على درجة قطع معينة.

هـ. من حيث بناء المفردات:

يعتمد بناء المفردات في الاختبار مرجعي المعيـــار على تباين الدرجات ومعامل التمييز، بينما يعتمد بناء المفردات في الاختبار مرجعي المحـــك على مدى تحقيق المفردة للهدف السلوكي.

7-4- **مميزات الاختبار التحصيلي مرجعي المحك:**

لقد حدد لوموس 1980(Loumos) ست مميزات للاختبار التحصيلي مرجعي المحك، كالتالي:

(أ). تعطي تلك الاختبارات معلومات صالحة لسجل توافق التلميذ؛ (ب). تتطلب الاختبارات مرجعية المحك من التلامذة أداء ما هو مطلوب في هدف الأداء؛ (ج). تصميم هذه الاختبارات لاختبار توافق التلامذة ولا يوجد اختبارات في مثلها فهي إما أن تكون قصيرة للغاية أو سهلة للغاية أو صعبة أو محددة للغاية؛ (د). نتائج الاختبارات مرجعية المحك تمكن التلامذة من الصعود للمرحلة التالية أو البرنامج الدراسي التالي؛ (هـــ). يمكن لتلك الاختبارات كتابتها في شكل اختبارات أداء أو موضوعية؛ (و). تعطى هذه الاختبارات كل المعلومات الضرورية لمنفذي الاختبار وكذلك التلامذة.

7-5- **خطوات بناء الاختبارات محكيـــة المرجع:**

وحتى يكون لنا تبرير سليم لاعتبار اختبار ما اختباراً مرجعا إلى المحك، فإننا لابد أن نتبع الخطوات التالية في بنائه وترجيعه بحسب ما أشار إلى ذلك فؤاد أبو حطب وآخرون (2008، ص.495)، بما يلي:

1. إعداد تخطيط للمحتوى يذكر فيه المهارات والمعرفة التي يعتبر الاختبار محاولة لقياسها.
2. يميز الأداءات، أى المرامى القابلة للقياس التي سوف يكون المفحوص قادراً عليها، مع التسليم بأن هذا المفحوص قد اكتسب كفاءة في المهارات والمعرفة المقاسة بالاختبار.
3. يميز المجال أو النطاق الذي يحدده كل مرمى ويكتب العناصر أو المفردات في الاختبار وفق التفاصيل المعينة في ذلك المجال أو النطاق، ثم يختار عشوائياً عنصرين لكل مرمى ليبنى منها الاختبار.
4. يتأكد من صحة حقيقة أن المهارات والمعرفة المقاسة بالاختبار هي في الواقع لازمة للمرامي الأدائية أو المرامي السلوكية المميزة المطلوبة في الخطوة 2، وقد يكون هذا هو أكثر جوانب هذه العملية تغيراً، ذلك لأن عملية إثبات صدق الاختبار تبدأ بحكم الشخص الذاتي، أي بإجراء صدق ظاهري، ثم تمتد لتتضمن حكم مجموعة من الخبراء أو تتضمن بيانات أو معطيات حقيقية حصلنا عليها من إعطاء الاختبار لمجموعة أثبتت أداءها الكفء لترى ما إذا كانت لدى أفرادها المهارات والمعرفة الخاصة بالاختبار.

5. يحدد درجة قطع (Cutoff score) أو درجة محك (Criterion score) تعبر عن الأداء على الاختبار الذي يجب أن يحصل عليه المختبر ليظهر أنه قد أحرز كفاءة كافية في المهارات والمعرفة ليكون قادراً على أداء ألوان السلوك المحكي.

7-6- **تحليل مفردات الاختبارات محكيـــة المرجع:**

تركز الاختبارات محكية المرجع على معرفة مدى تحقيق المتعلم لمستويات الإتقان التي قررها المعلم بالنسبة لدراسة أو عمل أو مهارة معينة، وبالتالي فهي تسعى للحصول على معلومات محددة عن كفاءة المتعلم وما حققه من الأهداف الموضوعية لهذه الدراسة أو المهارة، وهذا يعني أن أداء المتعلم على الاختبار يحدد بالرجوع إلى محتوى الاختبار ذاته ومدى تحقيق المتعلم لمستوى معين في المهارة أو الدراسة التي يقوم عليها الاختبار، وبشكل أكثر تحديداً، يمكن القول أن الاختبارات محكيـــة المرجع تحدد أداء التلميذ أو موقعه في ضوء محكات محددة (مستوى الإتقان المطلوب) أو في ضوء أهداف سلوكية تصف الأداء المتوقع من التلميذ بعد الإنتهاء من وحدة تدريسية معينة. وهذا بخلاف الاختبارات المعيارية المرجع التي تحدد أداء التلميذ وموقعه من خلال مقارنة الدرجـــة التي يحصل عليها في الاختبـــار بدرجات المجموعة التي ينتمي إليها، وهم تلامذة صفه أو زملائه.

وينبغي أن يراعي المعلم أو مصمم الاختبار عند تحليله لمفردات الاختبار المحكي المرجع أنها تستخدم عادة في اختبارات الإتقان، لذا فينبغي للمعلم أو لمصمم الاختبار وضع مستوى الأداء المقبول لتحديد الإتقان المطلوب من المتعلمين بلوغه بالإشارة إلى كل هدف أو إلى كل مهمة محددة. وقد يكون هذا المستوى المنشود بالإشارة إلى سرعة الإنجاز (مثال: يحل التلميذ عشر مسائل حسابية بسيطة في دقيقتين)، أو بالإشارة إلى دقة الإنجاز (مثال: يقيس الزاوية المنفرجة إلى أقرب درجة صحيحة)، أو بالإشارة إلى النسبة المئوية للمفردات الاختبارية التي أجاب عنها المتعلم بشكل صحيح (مثال: يعرف 80% من المصطلحات الأساسية التي درسها). (أبو ناهية، 1994، ص.319)

7-6-1- **معامل صعوبة مفردات الاختبارات محكيـــة المرجع:**

تتحدد صعوبة المفردة في اختبارات التمكن مرجعية المحـــك، وفقا للطبيعة الخاصة للإنجاز المستهدف من عملية القياس، فإذا كانت المهمات المطلوب إنجازها سهلة، كانت مفردات الاختبار هي الأخرى سهلة، وإذا كانت المهمات (الواجبات) متوسطة الصعوبة، كان من الضروري أن تكون مفردات الاختبار هي

الأخرى متوسطة الصعوبة، ولا يجوز بأي حال من الأحوال محاولة إجراء أي تعديل في مستوى صعوبة المفردة في الاختبارات مرجعية المحـــك، أو حذف المفردات السهلة من الاختبـــار لتوسيع مدى درجات الاختبار. ولا يتأسس مستوى صعوبة المفردة في الاختبارات مرجعية المحـــك على قدرة المفردة على التمييز بين المستوى المرتفع والمستوى المنخفض للمفحوصين– كما في حالة الاختبارات معيارية المرجع– وإنما تتحدد صعوبة المفردة في الاختبار في ضوء طبيعة المهمة أو الواجب المطلوب القيام به.(رضوان، 2011، ص.322)، إلا أن هذا لا يعني التقليل من أهمية أن تتصف المفردات في هذا النوع من الاختبارات بدرجة معقولـــة من الصعوبة، ذلك أنه كلما زادت درجة الصعوبة في الاختبار المحكـــي المرجع كان ذلك مؤشراً على أن التدريس فعال .(أبو ناهية، 1994، ص.320). ففي الطرق التقليدية تكون المفردات غير المميزة هي: السهلة جداً أو الصعبة جداً، أما بالنسبة للاختبارات مرجعية المحك فإنه يجب أن تعدل أسس التمييز. فلا تستبعد المفردة غير المميزة، إذا كانت تعكس صفة هامة للمحك. والمفردة موجبة التمييز مرغوب فيها في الاختبارات مرجعية المحك، والمفردة سالبة التمييز في الاختبارات مرجعية المحك يجب أن يتشكك فيها وقد يكتشف بعد فحص وتحليل دقيق عيب في هذه المفردة. (ميمي السيد، 2014، ص.12)

ويتم حساب مؤشر معامل الصعوبة لمفردة الاختبار ذات المرجعية المحكية بنفس الطريقة المستخدمة مع الأنواع الأخرى من الاختبارات، هو النسبة المئوية للمتعلمين الذين أجابوا عنه إجابة صحيحة مقابل العدد الكلي للمجيبين، ويمكن استخدام الصورة الرياضية رقم(03) في حساب هذا المعامل:

$$\text{معامل صعوبة المفردة} = \frac{\text{عدد المتعلمين الذين أجابوا عن المفردة اجابة صحيحة}}{\text{العدد الكلي للمتعلمين}} \times 100 \quad \ldots\ldots\ldots\ldots\ldots\ldots (03)$$

قيمة معامل الصعوبة يساوي متوسط الدرجات التي حصل عليها المتعلمين في المجموعة على المفردة مضروباً في 100، وتتراوح قيم هذا المعامل بين الصفر و100، وكلما اقتربت القيمة من 100 دل ذلك على سهولة المفردة، فقيمة معامل الصعوبة تساوي 100 إذا أجاب جميع أفراد المجموعة إجابة صحيحة عن المفردة وقيمته تساوي الصفر إذا أجاب جميع الأفراد عن المفردة إجابة خطأ. وقد يحدث أن يتم إيجاد معامل صعوبة المفردات لمجموعة المتعلمين قبل التعليم وبعده أو لمجموعة تلقت التعليم ومجموعة لم تتلقى التعليم كمؤشر لمصداقية المفردة، ولكي نستطيع الحكم على فاعلية التعليم باستخدام الاختبار محكـــي المرجع يجب أن تختلف قيم هذا المعامل بعد التعليم عن قيمته قبل التعليم بشكل دال إحصائياً (بوسالم، 2014، ص.181).

ولكن من المهم أن يدرك المعلم أو مصمم الاختبار أن معامل صعوبة المفردة في الاختبار المحكـــي المرجع مرتبط بأمرين هامين، وهما : (أ). المستوى التعليمي الذي يوجد عليه المفحوصين؛ (ب). المحـــك أو نقطة القطع التي يجب اجتيازها من قبل المفحوصين. فإذا حدد المعلم المحـــك أو مستوى الإتقان المطلوب من المفحوصين، بالإشارة إلى النسبة المئوية للمفردات التي يجب أن يجيب عليها المفحوص بشكل صحيح ولتكن المحك على سبيل المثال أن يجيب المفحوص على 80% من المفردات المتعلقة بهدف ما إجابة صحيحة، فإن هذا المحك لا يمكن للمعظم أن يصلوا إليه أو يجتازوه إذا كان متوسط معامل صعوبة المفردات لذلك الهدف 70% وبالتالي فيجب على المعلم أن يضع في اعتباره مستوى المفحوصين وقدراتهم عند تحديد المحك أو مستوى الإتقان المطلوب أو نقطة القطع التي يجب اجتيازها.(أبو ناهية، 1994، ص.323)

7-6-2- حساب معامل الحساسية أو تمييز مفردات الاختبارات محكيـــة المرجع:

إن قدرة مفردات الاختبار على التمييز بين الأفراد في المجموعة العليا والأفراد في المجموعة الدنيا ليست حيوية لتقويم فعالية مفردات الاختبار محكـــي المرجع. فقد تكون المفردة جيدة في اختبار الإتقان محكي المرجع، رغم أنها ذات قدرة تميزية منخفضة أو قريبة من الصفر بالنسبة للاختبار المعياري المرجع. فإذا أجاب كل المفحوصين إجابة صحيحة على المفردة (فيكون معامل التمييز يساوي صفر) بعد الإنتهاء من تدريس وحدة تعليمية معينة، فإن هذا يدل على أن كل من التدريس والمفردة فعالان. وعلى الرغم من أن مثل هذه المفردات تحذف في الاختبارات معيارية المرجع بسبب عدم قدرتها على التمييز، إلا أنها في اختبارات الإتقان محكيـــة المرجع تمد المعلم بمعلومات مفيدة عن مدى إتقان تلامذته للمهام التعليمية. وبالتالي فمعامل تتميز مفردات الاختبارات مرجعية المحـــك تتميز بإنخفاض معامل تمييزها ذلك لأن الهدف من هذه الاختبارات ليس قياس الفروق الفردية بين الأفراد، بل قياس وتحديد ما إذا كان المتعلم قد وصل إلى مستوى الإتقان في المعرفة أو المهارة.

وقد تم تطوير عدة طرق لحساب معامل تمييز المفردة في الاختبار المحكي المرجع، إلا أننا سنكتفي بطريقتين في هذا البحث واللتين تعتبران من أكثر الطرق ملائمة لتحليل مفردات الاختبارات المحكيـــة المرجع، وهما: (أبو ناهية، 1994، ص.323)

7-6-2-1- التمييز باستخدام الاختبار القبلـــي و الاختبار البعــدي:

يمكن أن يحسب تمييز أي مفردة في الاختبار المحكي المرجع باستخدام معامل التمييز القبلي – البعدي (Pre- and Postdiscrmination Index) والذي يتم حسابه بعد تطبيق الاختبار على المفحوصين مرتين مرة قبل التدريس (Pretest) ومرة أخرى بعده (Posttest)، ويتم ذلك باستخدام المعادلة رقم(04):

م ت = م ص (ب) – م ص (ق) .....(04)

حيث أن: م ت = معامل التمييز؛ م ص (ب) = معامل الصعوبة قبل التدريس؛ م ص (ق) = معامل الصعوبة بعد التدريس. وتتراوح قيم معاملات التمييز بين+1 و-1 . وعندما يكون معامل التمييز قريبا من الصفر، فهذا يشير إلى أن القدرة التمييزيـــة للمفردة ضعيفة. كما أن المفــردات التي يجيب عليها جميع المفحوصين إجابة صحيحة في الاختبـــار القبلي (قبل التدريس)، وفي الاختبار البعدي (بعد التدريس) هي مفردات غير فعالة حيث يكون معامل تمييز مثل هذه المفردات عادة صفراً.

7-6-2-2- التمييز باستخدام مجموعتين محكيتين أو طريقة (ب)(The B index Approach):

طور برينان (Brennan,1972) طريقة لحساب معامل الحساسية أو التمييز، والتي تسمى بمؤشر(**ب**)، اعتمادا على مستويات الإتقان المطلوبة من المفحوصين، أو نقطة القطع التي يجب اجتيازها. بعد تطبيق الاختبار المحكي المرجع تستخدم نتائجه لتقسيم المفحوصين الذين أخذوا الاختبار إلى أربعة مجموعات فرعية في ضوء بعدين أو متغيرين أساسيين، وهما: الأداء على الاختبار (متقن- غير متقن)، والأداء على المفردة (إجابة صحيحة – إجابة خاطئة) . ويتم حساب معامل الحساسية أو تمييز مفردات الاختبارات محكية المرجع باستخدام الطريقة التي اقترحها برينان (Brennan,1972)، بالمعادلة رقم(05):

$$B = \frac{U(n_2 - L) - L(n_1 - U)}{n_1 n_2} \quad \text{.....(05)}$$

$$= \frac{Un_2 - Ln_1}{n_1 n_2} \quad \text{.....(05.1)}$$

$$= \frac{U}{n_1} - \frac{L}{n_2} \quad \text{.....(05.2)}$$

حيث أن : (B) مؤشر معامل التمييز؛ (U) عدد المفحوصين الذين أجابوا على المفردة إجابة صحيحة وحققوا مستوى الإتقان المطلوب؛ (L) عدد المفحوصين الذين أجابوا على المفردة إجابة صحيحة ولم يحققوا مستوى الإتقان المطلوب؛(n_1) عدد المفحوصين الذين حققوا مستوى الإتقان المطلوب؛ (n_2) عدد المفحوصين الذين لم يحققوا مستوى الإتقان المطلوب.(Brennan, 1972, p.292)

حيث يتم تسجيل البيانات الخاصة بمفردة الاختبار في جدول هو أقرب للجداول التي تستخدم في توضيح التصميم العاملي 2x2 . كما يبينها الجدول رقم (01) من خلال تخطيط لمكونات المعادلة يوضح طريقة تسجيل البيانات الخاصة، بتحليل إحدى مفردات الاختبار محكي المرجع وحساب معامل تمييزها:

جدول (01)

يوضح طريقة تسجيل البيانات الخاصة، لحساب معامل التمييز بطريقة معادلة برينان.

المجموع	إجابة المفردة		
	إجابة خاطئة (0)	إجابة صحيحة (1)	
n1		U	المتمكنون
n2		L	غير المتمكنيين

ويتراوح معامل التمييز لبرينان ما بين −1 و+1 . وتشير القيمة الموجبة لهذا المعامل إلى أن المفردة تميز بين المفحوصين الذين حققوا مستوى الإتقان المطلوب أو المحك، وأولئك الذين لم يحققوا مستوى الإتقان أو المحك، أما القيمة السالبة لمعامل التمييز فهي تدل على وجود مشكلة أو خلل حول المفردة، فقد تكون المفردة نفسها ضعيفة في محتواها أو في صياغتها، وقد تكون طريقة التدريس المستخدمة من قبل المعلم خاطئة، وقد يكون المفهوم الذي تدور حوله المفردة صعب، وواجب المعلم أن يفحص المفردة جيداً ويحدد السبب أو الأسباب التي أدت لضعفها، فإذا كان الأمر يتطلب منه إعادة صياغة المفردة وتحسينها فيجب أن يقوم بذلك، وإذا تبين له أن المفردة غير صالحة فيجب حذفها وإعادة تقدير درجات المفحوصين من جديد بدونها، أما إذا تبين له أن طريقة التدريس غير فعالة فيجب أن يخطط لتدريس جديد وفعال . (أبو ناهية، 1994، ص.326) (النبهان، 2013، ص.251)

7-6-3- صـــدق الاختبارات محكيـــة المرجع:

تعتمد معظم الطرق للحصول على صدق الاختبارات مرجعية المحك، على الارتباطات، وكذلك التباين. لذلك فإن النتائج بالنسبة للصدق، كما هي بالنسبة للثبات تكون مفيدة إذا كانت موجبة ولكنها لا تهدم بالضرورة

إذا كانت سالبة. ولكن يحدد صدق الاختبارات مرجعية المحـــك على أساس كفاءتها في تمثيل المحك. ويتم تقدير صدق الاختبارات مرجعية المحـــك باستخدام ثلاث طرق رئيسة تماثل إلى حد ما الطرق المستخدمة في التحقق من صدق الاختبارات مرجعية الجماعة أو المعيار، وهي: الصدق الوصفي، الصدق الوظيفي، صدق إنتقاء النطاق السلوكي.(ميمي السيد، 1994، ص.20)

7-6-3-1- **الصدق بدلالة المحتوى أو الصدق الوصفـــي:**

يتم حساب صدق الاختبارات محكيـــة المرجع باستخدام طريقة الصدق بدلالة المحتوى أو صدق المحتوى، والطريقة المستخدمة في حساب صدق المحتوى بالنسبة للاختبارات محكية المرجع هي نفس الطريقة المستخدمة مع الاختبارات معيارية المرجع، إلا أن التحقق من صدق المحتوى في الاختبارات محكية المرجع أسهل لحد ما منه في الاختبارات معياريـــة المرجع، والسبب في ذلك هو وجود مجال ذو مواصفات محددة مما يسهل تحديد مدى مطابقة مفردات الاختبار لذلك المجال. (أبوناهية، 1994، ص.350)

يعتبر الاختبار المرجعي المحـــك صادقاً بدرجة دقته في قياس فئة عناصر سلوك المتعلم التي وردت في تعريف المجال وتحديده وهذا يتوقف على القدرة الوصفية للاختبار. والطريقة التي توصلنا إلى تحديد الصدق الوصفي المناظرة في أساليب صدق الاختبارات المرجعية المعيـــار هي صدق المحتوى، بمعنى الاعتماد على أحكام مراجعي الاختبـــار. وفي القياس المرجعي المحـــك يؤدي وصف المجال وصفا مفصلاً وواضحاً إلى إتمام هذا العمل بدقة أكبر وعلى المحكم أن يحدد ويميز نسبة العناصر الاختبارية التي تتسق وتتطابق مع وصف المحك. (السيد، 1994، ص.20)

7-6-3-2- **الصدق الوظيفـــي:**

هو وضع اختبار صادق يقيس وظيفة معينة قياساً ملائماً. ويعتبر الصدق المرتبط بمحك هو نوع الصدق المناظر في الاختبارات المرجعية المعيار، حيث نحاول أن نتنبأ بأداء الفرد في موقف لاحق من أداء الفرد في اختبار يثبت صدقه، ووظيفة الاختبار هو المساعدة على التوصل إلى تنبؤات دقيقة.

7-6-3-3- **صدق النطاق السلوكـــي:**

ويقوم هذا النوع على دقة اختيار المجال، وهذا يتطلب أن نحدد مجالاً معينا اختاره واضع الاختبار يخدم على نحو دقيق وصحيح ويصلح كمؤشر لبعد أكثر عمومية صمم أو وضع ليعكسه. وهناك طرق عدة للتثبت من صدق إنتقـــاء النطاق السلوكـــي مثل: تدريس مجموعات صغيرة من المتعلمين حتى يتقنوا كل

مجالات المهارة والكفاءة ثم نختبرهم لنتبين قابلية هذه المهارة للتعميم وقدرتها على ذلك بالنسبة للمجالات الأخرى.

7-6-4- **ثبـــات الاختبارات محكيـــة المرجع:**

يشار إلى مصطلح الثبات في المقاييس المرجعية المحك كما هو في الاختبارات المرجعية المعيار، بعدم وجود خطأ في القياس. ولما كانت المفردات على الاختبار محكي المرجع تمثل عينة من الأعمال، فإن الدرجة الكلية للاختبارات محكية المرجع لا تكفي لتميز أداء التلامذة، فيمكن أن يحصل تلميذان على نفس الدرجة الكلية، لكن نمط استجاباتهما و المهارات المتعلمـــة، ربما تختلف إلى حد كبير.

فبالنسبة لثبـــات الاختبار مرجعي المعيار، المفردات التي يفشل في الإجابة عليها 50% منهم تكون ثابتة، كذلك فإن التباين في الدرجات يكون عاملاً مهما في تحديد الثبات. أما في الاختبار محكي المرجع قد يكون تباين الدرجات صغيراً جداً. إذ يمكن أن يفشل معظم التلاميذ في أداء الاختبار قبل شرح التعليمات، ويمكن أن ينجح معظمهم بعد شرحها. والمفردات الهامة ليست دائماً متوسطة الصعوبـــة، قد تكون سهلة أو صعبة ويمكن أن تكون اختباراً نقياً في مجال خاص، أي يجب أن يكون الاختبار متسقاً داخلياً. و يشير الثبات إلى تماسك واتساق درجات الاختبـــار التي يتم الحصول عليها من مقياس إلى مقياس آخر، ويحذر (هامبلتون) و (نوفك) (1973) من تفسير تقديرات الثبات التي تم حسابها بالطرق التقليدية للاختبارات مرجعية المحـــك، لان الثبات الذي يتم تقديره بالطرق التقليدية يقوم على أساس دوال مربع الأخطاء المفقودة الذي يقدم معلومات متعلقة بثبات المقاييس أو الدرجات.(السيد، 1994، ص.14)

ولما كانت الاختبارات محكيـــة المرجع تستخدم عادة في تصنيف المتعلمين إلى مجموعتين أو أكثر بحسب درجة إتقانهم لأهداف محتوى دراسي معين، لذلك اعتبر علماء القيـــاس النفسي والتربوي المعاصر أن مفهوم الثبات في هذا النوع من الاختبارات يشير إلى مدى إتساق التصنيف بناءا على الدرجات المستمدة من تطبيق الاختبار نفسه. وتوصل عدد من هؤلاء العلماء إلى طرق مختلفة لتقدير الثبات بهذا المفهوم الجديد، وتحققوا من مصداقية هذه الطرق سواء باستخدام أساليب المحاكاة عن طريق الحاسب الآلي، من مثل أسلوب مونت كارلو (Monte Carlo Method)، وهو أسلوب يعتمد على إصطناع بيانات تحقق فروض توزيعات البيانات الفعلية، أو أسلوب التحليل المتعـــدد الذي يعتمد على نتائج عدد من التطبيقات.(بوسالم، 2014، ص.182)

7-6-4-1- تقدير ثبـــات درجات الطلاب في النطاق السلوكي:

طرق هذه المجموعة تهتم بتقليل خطأ التباين الناتج عن الفروق بين درجات التلامذة في النطاق السلوكي بالإضافة إلى درجات النطاق، ومن تلك الطرق مايلي:

1- معامل أو مؤشر ليفينجستون (Livingston Index (1972)):

والذي يعتمد على مفاهيم النظرية الكلاسيكية في تقدير ثبات الاختبارات معيارية المرجع والتي تركز على معرفة انحراف درجة الفرد عن متوسط العينة. إلا أنها تستخدم الدرجة الفاصلة بدلاً من المتوسط، فبدلاً من إيجاد انحراف درجات الأفراد عن المتوسط تحسب انحرافات درجات الأفراد عن الدرجة الفاصلة المحددة.(أبو علام، 1995، ص.264)، والذي يعطى بالمعادلة رقم(06):

$$k^2(XT) = \frac{\mathrm{KR20}(S^2) + (M - \lambda)^2}{(S^2) + (M - \lambda)^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots(06)$$

حيث أن: $k^2(XT)$ = معامل (ليفنجستون) لتقدير الاعتمادية للاختبارات المحكية؛ S^2= تباين درجات الاختبار (أو مربع الإنحراف المعياري)؛ M = متوسط درجات الاختبار؛ λ = درجة القطع للاختبار؛ KR_{20}= معامل الثبات كودر – ريتشارسون. كما يمكن أن يعطي كذلك بالمعادلة رقم(07):

$$k^2 = \frac{\frac{I}{I-1}\left[V_x - \sum V_i\right] + (\bar{X} - C)^2}{V_x + (\bar{X} - C)^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots(07)$$

حيث أن : I = عدد مفردات الاختبار؛ $\bar{X}$ = المتوسط الحسابي لدرجات الاختبار؛ C= درجة القطع؛ V_x= التباين الكلي لدرجات الاختبار؛ V_i= تباين كل مفردة. (Brown & Hudson, 2002, p.194)

2- تقدير ثبات تصنيف التلامذة في مجموعات على حسب درجة تمكنهم من النطاق السلوكي:

تهتم طرق هذه المجموعة بتحديد الأخطاء الناتجة عن عدم اتساق التصنيف عند تطبيق اختبارين متوازيين أو في حالة إعـــادة تطبيق الاختبار على أفراد العينة، ومن تلك الطرق مايلي:

طرق تتطلب تطبيق الاختبار مرة واحدة:

أ- طريقة هاريس (Harris Method):

اشار هاريس بأن هذه الطريقة لا تعتمد في تحديدها على متغير طول الاختبار، وإنما تقوم على الارتباط بين متغير يمثل مجموع درجة التلميذ في الاختبار، ومتغير ثنائي آخر يمثل تصنيف التلميذ إلى متقن وغير متقن على ضوء درجة القطع المستخدمة.

ب- معامل كابا (هاينا) (Huynh Kappa Coafficient):

يمكن تقدير معامل (هاينا) في حال تطبيق اختبار واحد أو اختبارين متوازيين، ونظراً لصعوبة تقدير هذا المعامل في حال تطبيق اختبارين متوازيين، ظهرت عدة طرق تعالج هذا الاشكال منها:

أ- معامل كارفر (Carver Method):

يعد معامل كارفر من الطرق الأولى المقترحة لتقدير ثبات الاختبارات مرجعية المحك التي تتعلق باتساق قرار التصنيف. وتقوم هذه الطريقة على تطبيق اختبارين متوازيين على مجموعة واحدة من الأفراد، يليه مقارنة نسبة عدد التلامذة الذين تم تصنيفهم ضمن المتقنين في كلا الاختبارين، فإذا كانت النسبتان متماثلتان أو قريبتان من بعضهما فبذلك يعتبر الاختباران ثابتين.

ب- معامل (كابا) (Swaminathan-Hambleton-Algina):

تعتبر هذه الطريقة امتدادا لطريقة كارفر، حيث تعتمد على فرق النسبة بين التلامذة المتمكنين وغير المتمكنين في صورتي الاختبار المتوازيتين أو في مرتي تطبيق الاختبار، وتمتد قيمة معامل الثبات بين -1 و +1.(السيد، 2014، ص.16)

7-6-5- العوامل المؤثرة في حساب معامل ثبات الاختبار محكـــي المرجع:

تختلف العوامل المؤثرة في حساب معامل ثبات الاختبارات محكية المرجع، عن العوامل المؤثرة في حساب معامل ثبات الاختبارات معيارية المرجع، نظراً لإختلافهما في الهدف، والتصميم كما تم بحثه سابقاً، حيث أن العوامل المؤثرة في حساب معاملات ثبات الاختبارات محكية المرجع، والتي وردت في تراث القياس التربوي وتناولتها الدراسات السابقة في هذا المجال كطول الاختبار، درجة القطع، حجم العينة، قدرة التلميذ، خواص محتوى الاختبار، شكل توزيع الدرجات، طريقة اختيار المفردة، طريقة تقدير الدرجات، نوع المفردة (الاختيار من متعدد، الصواب والخطأ، المقال)، ترتيب المفردات. وسيتم التركيز في هذا البحث على أكثر العوامل أهمية في حساب معامل ثبات الاختبارات محكية المرجع، وهي: (أ). طول الاختبار؛ (ب). درجة القطع؛ (ج). حجم العينة.

7-6-5-1- **طول الاختبـــار:**

بما أن الاختبارات محكية المرجع تستخدم لتصنيف المختبرين في حالات إتقان محدد فإن إرتفاع قيمة معامل الثبات يعني زيادة إتساق قرارات التصنيف، وبالتالي زيادة الثقة في نتائج الاختبار. أي أن مشكلة تحديد طول الاختبار ترتبط بعدد أخطاء التصنيف المسموح التجاوز عنها، واحدى طرق تقليل احتمالات خطأ سوء التصنيف هي جعل الاختبار طويلاً جداً وربما يكون هذا مناسباً دائماً (Hambelton et al , 1978, p.24). ويشير جرونلند (Gronlund, 1981) في هذا السياق إلى أنه عند استخدام الاختبارات محكية المرجع في التعليم الصفي فإننا نستطيع زيادة احتماليـــة صدق وثبات النتائج باستخدام عينة كبيرة بشكل كاف من مفردات الاختبار لكل هدف تعليمي أو مجال مخصص من مهام التعلم الواجب قياسها. فإذا كانت المخرجات المقصودة كثيرة التخصيص وعالية التركيب فعدد قليل نسبياً من المفردات وليكن خمس مفردات ربما يكون كافياً لإعتماد قرار فيما يتعلق بالإتقان. (p.112)

7-6-5-1- **درجة القطـــع:**

تعرف درجة القطـــع على أنها الدرجة التي تفصل بين التلامذة الناجحين والراسبين في اختبار ما، وتحدد نسبة التلامذة القادرين على تحقيق الحد الأدنى المقبول كشرط للإتقان أو النجاح، وهي " تعبر عن الأداء على الاختبـــار الذي يجب أن يحصل عليه المختبر ليظهر أنه قد أحرز كفاءة كافية في المهارات والمعرفة وليكون قادراً على أداء ألوان السلوكـــات المحكية (فؤاد أبو حطب وآخرون، 2008، ص.495). وقد تعددت المسميات التي أطلقت على درجة القطع (Cutoff score)، فأحياناً يطلق عليها مستويات الإتقان (Mastery level)، أو درجة النجاح (Passing score)، أو الحد الأدنى للكفاية (Minimum competency level)، أو مستويات المحك (Criterion levels)، وأحياناً يطلق عليها درجات القطع (Cutoff score). (علام، 2001، ص.253)

فجميع معاملات ثبات الاختبارات محكية المرجع حساسة لموضع درجة القطع، وبالتالي فإن تفسير أي مؤشر يجب أن يتضمن تعيين درجة القطع، ولا أهمية لحسابه بدون تبرير اختيارها، لأن القيمة المرتفعة لمؤشر ثبات مبني على درجة قطـــع ليس لها تبرير أو لها تبرير ضعيف سيشير إلى أن الاختبار يستطيع تصنيف التلامذة باتساق في مجموعات خاطئة، فاتساق صنع القرار بدون دقته له قيمة مشكوك فيها في التقويم محكـــي المرجع (علام، 2000، ص.288).

والسؤال الذي يطرح نفسه، كيف سنختار الطريقة المناسبة من بين الطرق الكثيرة المتوفرة لتحديد درجة القطع؟

لقد وضع هامبلتون (Hambelton, 1980) عدة عوامل يجب أخذها بالحسبان عند اختيار طريقة تحديد المحكات ومن بين هذه العوامل:(أ). أهمية القرارات المتخذة؛ (ب). مقدار الوقت المتاح لوضع المحك؛ (ج). المصادر المتاحة (أشخاص ومال) للحصول على العمل المطلوب؛ (د). إمكانية الأحكام (فبعض الطرق تتطلب معرفة المحتــوى والمختبرين الواجب اختبارهم أكثر من غيرها)؛ (هــ). ملاءمة الطريقة لنوع الاختبار تحت الدراسة. ويضاف إلى ذلك التكاليف النسبيــة لأخطاء سوء التصنيف والتي تقسم إلى نوعين: (p.103)

1. الخطأ من النوع الأول، ويتمثل في نجاح تلميذ لا يستحق اجتياز الاختبار.
2. الخطأ من النوع الثاني، ويتمثل في رسوب تلميذ كان من الواجب أن ينجح (Geisinger , 1991, p.17).

7-6-5-1- حجم العينة:

يؤثر حجم العينة في قيمة معامل ثبات الاختبارات محكية المرجع، فكلما زاد حجم العينة ارتفعت قيمة معامل الثبات، وإن كان العكس غير صحيح، فربما يرجع ارتفاع قيمة معامل الثبات لعوامل أخرى غير زيادة حجم العينة. وقد اشارت دراسة داتشك (Dutschke ,1988)، التي هدفت إلى تعيين مميزات وخصائص مؤشرات ثبات الاختبارات محكيــة المرجع، إلى اقتراب خصائص التوزيع لمؤشرات الثبات من المميزات الطبيعية مع زيادة حجم العينة.

7-6-6- كيفية تحديد درجات القطع في الاختبارات محكيــة المرجع:

عادة ما يستخدم الاختبار محكي المرجع في تقدير أداء الفرد بالنسبة إلى نطاق سلوكي معين، وهو ما قد يمثل مجموعة من المعارف والمهارات المعرفة بطريقة إجرائية (سلوكية)، بغض النظر عن علاقة أدائه بأداء غيره من الأفراد الذين يطبق عليهم نفس الاختبار، كما يستخدم هذا التقدير في تصنيف المتعلمين إلى مجموعتين إحداهما متقنة والأخرى غير متقنة لهذا النطاق السلوكي، وإتخاذ مثل هذا القرار لابد أن يعتمد على تحديد مسبق لدرجة قطــع معينة في الاختبار تفصل بين أفراد المجموعتين، وبالطبع يمكن أن تحدد درجات قطع مختلفة بحسب مستويات الإتقان المطلوبة وفق الغرض من الاختبار.

وقد كثر الجدل والنقاش بين المربين وعلماء القياس التربوي المعاصر حول كيفية تحديد درجات القطع في مختلف الاختبارات، والدليل على ذلك كثرة ما ينشر من دراسات تتناول هذه المشكلة في دوريات القياس وبخاصة في أعدادها الحديثة، فقد أفردت مجلة القياس التربوي (Journal Of Educational Measurement) التي تصدر في الولايات المتحدة الأمريكية عددا كاملاً صدر عام 1998 لمناقشة استخدامات درجات القطع، والمشكلات المتعلقة بكيفية تحديدها في الاختبارات النفسية والتربوية، وخلاصة المقالات التي نشرت هي، تعدد طرق تحديد درجات القطـــع باختلاف مواضيع القيـــاس، والغرض من الاختبار، وخصائص العينات المطبق عليهم.(بوسالم، 2014، ص.190)

وتنقسم الطرق والدراسات المقترحة لتحديد درجات القطـــع في الاختبارات المحكية، إلى ثلاثة أقسام رئيسية هي:

أ- طرق تعتمد تقدير التحكيمية:

تستند هذه الطرق على آراء المتخصصين في مجال القيـــاس أو في محتوى الاختبار المراد قياسه، وذلك لتحديد درجة القطع المناسبة، ومن هذه الطرق ما يلي:

1- **طريقة إيبل (Ebel's Method):** تعتمد هذه الطريقة على تعلقيات المحكميـــن على مفردات الاختبار، وذلك من خلال اعتبار أن مواصفات المفردات لها بعدين هما: بعد الصلة، وبعد الصعوبة.

2- **طريقة أنجوف (Angoff Method):** يطلب من كل محكم في هذه الطريقة فحص كل مفردة من مفردات الاختبار، ثم يطلب منهم تصور مجموعة من الأفراد الذين حققوا الحد الأدنى للكفاية التي يقيسها الاختبار، ثم يقدر نسبة عدد الأفراد الذين يحتمل أن يجيبوا إجابة صحيحة عن كل مفردة من مفرداته، ويمثل متوسط هذه النسب الحد الأدنى لمستوى الاجتياز في الاختبار.

3- **طريقة ندلسكي (Nedlesky's Method):** هذه الطريقة تصلح لتحديد درجة القطع للاختبارات التي تشتمل على أسئلة ذات اختيار من متعدد وفي هذه الطريقة يطلب من كل محكم فحص كل مفردة من مفردات الاختبار، ثم يحدد من بين الخيارات الخاصة لكل مفردة تلك التي يمكن أن يتجنب التلامذة ذوو الحد الأدنى للكفاية اختيارها، لأنها لا تمثل في نظرهم الإجابة الصحيحة للمفردة، بعد ذلك يضع المحكم درجة السؤال بناء على ما تبقى من الخيارات وذلك بأخذ مقلوب الخيارات المتبقية.

ب- الطرق التي تعتمد جزئياً على التحكيم وتسترشد ببيانات تجريبية:

نتيجة وجود قصور ملحوظ في الطرق التحكيميـــة، ومن ذلك القصور اعتمادهم على تحكيم مفردات الاختبـــار دون النظر إلى الأداء الفعلـــى للمختبرين، مما يؤدي ذلك إلى الحصول على معايير غير متميزة بدرجة عالية من الصدق، ومن أهمها طريقة التحكيم المعزز بالمعلومات، وطريقة أنجوف المعدلة، والطريقة التوفيقية.

جـــ- طرق تعتمد على البيانات التجريبية وتسترشد بالتحكيم:

ومن أهمها طريقة المجموعات المحكية، وطريقة المجموعات الحدية، وطريق المجموعات المتناقضة.

د- طرق تعتمد المفاهيم الإحصائية:

يتطلب استخدام هذه الطرق توفر معلومات معينة عن كلفة وعائد ما يترتب على التصنيف الخاطئ للمتعلمين على أساس درجات القطع المختلفة، فمن المعلوم أن هناك نوعين من الخطأ مرتبطين بتصنيف الأفراد إلى متقنين وغير متقنين لمجموعة معينة من الأهداف التعليمية، أحدهما هو الخطأ السالب والآخر هو الخطأ الموجب، فالخطأ السالب يحدث عندما تكون الدرجة التي يحصل عليها تلميذ أتقن الأهداف التي يقيسها الاختبار، أقل من درجة القطع، والخطأ الموجب يحدث عندما تكون الدرجة التي يحصل عليها تلميذ لم يتقن الأهداف أعلى من درجة القطع، وبالطبع لكل من الخطأين أثره على عملية صنع القرارات التعليمية، ولكن يختلف هذا الأثر في خطورته تبعا لنوع وأهمية القرار التصنيفي، فتوجيه المتعلمين الى دراسة مجموعة من الأهداف دون التأكد من اتقانهم للأهداف السابقة عليها، يؤدي إلى الوقوع في خطأ موجب، وهذا الخطأ يكون بالطبع له ضرر أكبر على تحصيل المتعلمين من الخطأ السالب الناتج عن تدريس محتوى الأهداف مرة أخرى للمتعلمين برغم اتقانهم لها.(بوسالم، 2014، ص.193) (ميمي السيد، 2014، ص.18)

ملخص الفصل:

تعتمد عمليتي القيـــاس والتقويـــم التربوي على أدوات ومقاييس تقدم بيانات كمية تتيح للتربويين فهم الظاهرة التربوية بشكل دقيق ومحدد، وتعتبر الاختبارات والمقاييس جزءاً هاماً من أجزاء العملية التربوية فمن خلالها يستطيع المعلم الحكم على مدى تحقق أهداف البرامج والمناهج التعليمية التي يقوم بتدريسها، كما أنها تفيد التلميذ كتغذية راجعة في التعرف على مستوى تحصيله الدراســـي من خلال أدائه في الاختبار. والمتتبع لحركة بناء الاختبارات التحصيلية وفق النظرية الكلاسيكيـــة في القيـــاس، يلاحظ أن هناك إتجاهين

رئيسيين، يختلف كل إتجاه عن الآخر في الافتراضـــات التي يقوم عليها، وفي الإطار المرجعي الذي يعتمد عليه في تفسير نتائج التلامذة في الاختبـــار، وفي إعداد أدوات القيـــاس وهي القياس المعيـــاري المرجعي والقياس محكي المرجع، حيث يعتمد هذا الأخير على تحديد النطاق الأهداف السلوكية تحديداً دقيقاً، ومن ثم قياس مدى اتقان التلميذ لهذه الأهداف من خلال مقارنة أداء التلميذ بمستوى معين من الأداء (درجة القطع).

وعليه حاولنا من خلال هذا الفصل التطرق إلى القيـــاس النفسي والتربوي، ومنه إلى بنـــاء الاختبارات التحصيلية وفقاً للنظرية الكلاسيكية للقيـــاس، وذلك بتسليط الضوء على التطور التاريخي للقيـــاس النفسي والتربوي، و تحديد مفهومه، ثم التطرق إلى مستويات القيـــاس النفسي والتربوي، ثم إلى نظريات القيـــاس النفسي والتربوي، و أخيراً إلى كيفية بناء الاختبارات التحصيلية مرجعية المعيار، ومحكيـــة المرجع.

الفصـــل الثالث

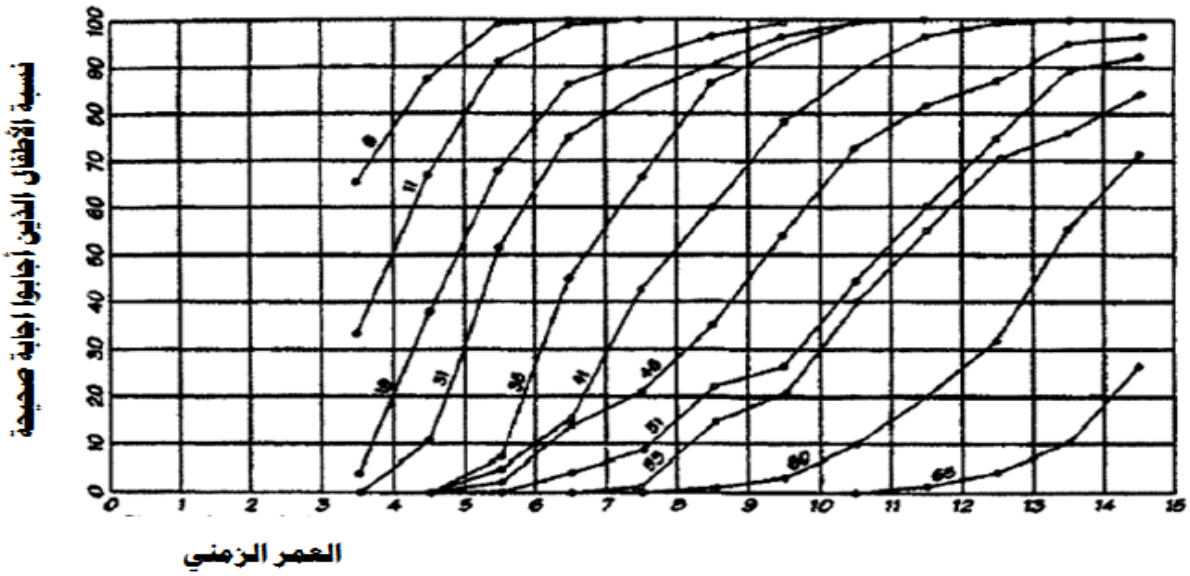
توظيف نظرية الاستجابـــة للمفردة في بناء الاختبـــارات التحصيلية

تمهيـــد:

تعتمد نظرية الاستجابـــة للمفردة على مفاهيم ومبادئ تختلف اختلافاً جوهريا عن تلك التي تعتمد إليها النظرية الكلاسيكية في القياس النفسي والتربوي. فمنهجية القيـــاس التي تميز نظرية الاستجابة للمفردة تستند إلى نماذج رياضية احتمالية تعمل على ضبط العوامل المربكة المختلفة التي تؤثر في المقارنات بين الدرجات، وذلك باجراء نوع من النمذجة الرياضية لبارامترات العملية الاختبارية. ويرجع الفضل في تقديم هذه المنهجية إلى كل من لورد ونوفيك (Lord and Novick,1968)، الذين عرضوها تفصيلياً في كتابهما الأساســـي "النظريات الإحصائية لدرجات الاختبارات العقلية Statistical Theories of Mental Test Scores". واقتراح نماذج رياضية تمثل ما يحدث عندما يستجيب فرد معين لمفردة اختبارية، بحيث يمكن التنبؤ بأدائه أو تفسير هذا الأداء عن طريق خصائص معينة يطلق عليها السمـــات. وتستخدم هذه النماذج في تقدير بارامترات للمفردات، وكذلك للأفراد في هذه السمات، حيث يمكن باستخدام القيم التقديرية لهذه البارامترات تفسير أداء كل فرد في الاختبار. ونظراً لأن هذه السمات يصعب ملاحظتها وقياسها قياساً مباشراً، وإنما يستدل عليها باستخدام هذه القيم التقديرية، والتي يطلق عليها عادة السمات الكامنة "Latent Traits" أو القدرات "Abilities". (علام، 2013، ص.47)

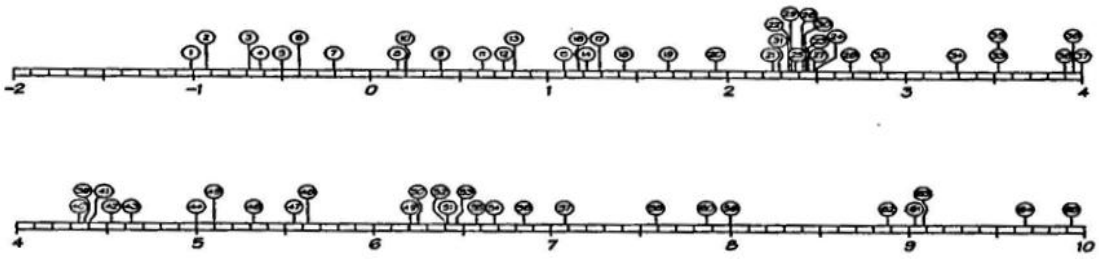
1- التطور التاريخي لنظرية الاستجابـــة للمفردة (IRT) ونماذجها:

ربما تعود الإشارات الأولى لنظرية الاستجابة للمفردة إلى الورقة البحثية التي قدمها (لويس ثرستون) عام (1925)، بعنوان "منهجية للمقايسة الاختبارية النفسية والتربوية" (A Method of Scaling Psychlogical and Educational Tests) والتي قدم خلالها حل لمشكلة كانت تشغل الباحثين النفسيين في حينه، حول أفضل طريقة لمقايسة القدرة العقليـــة من خلال مفردات مقياس "بينيه- سيمون" للأطفال على تدريج العمر الزمني، حيث قام برسم العلاقة لعينة كبيرة بين نسبة الأطفال الذين نجحوا في القيام بمهام اختبـــار "بينيه – سيمون" على المحور الرأسي، والعمر الزمني على المحور الأفقي، فنحصل على منحنيات لها شكل الحرف (S) توضح زيادة نسبة الاستجابة الصحيحة على مقياس "بينيه – سيمون" باعتبارها دالة لزيادة العمر الزمني، هذه المنحنيات التي يوضحها الشكل رقم (03) تشبه منحنيات الترجيح اللوغاريتمي لنموذج (راش) التي تسمى "منحنيات القدرة".



شكل (03) **رسم بياني لثورستون (Thurstone,1925) يوضح نسبة الإجابة الصحيحة على مفردات اختبار "بينيه-سيمون" للأطفال تبعاً لمتغير العمر الزمني.**(Wilson, 2005, p.89)

قام ثيرستون بتحديد موضع المفردات الـــ (65) في اختبار "بينيه – سيمون" على مقياس العمر، نقطة أصله الاعتباطية هي صفر، ووحدته المسافية تساوي (1) انحراف معياري، في حين أن الشكل رقم (04) يوضح التدريج العمري لمفردات الاختبار العقلي الناتجة من عملية المقايسة التي قام بها، وموضع النقاط على الشكل يعكس مستوى صعوبتها وفق العمر الزمني، وبذلك قدم لمطبقي اختبار " بينيه – سيمون" أساساً لتقديم مفرداته، وفق تزايد صعوبتها، وتحديد العمر العقلـــي من عدد الاستجابات الصحيحة وفق ما حدده بينيه. (ديانا فهمي، 2010، ص.36)



شكل (04) **المواضع المقدرة للمفردات على مقياس للعمر ذي نقطة أصل اعتباطية ووحدات قياس متساوية.** (Louis, 1925, p.449)

يمثل الشكل رقم (04) مقياس مطلق لمفردات اختبار بينيه، المسافة الفترية بين الأرقام تعبر عن الانحراف المعياري لذكاء أطفال في الثالثة والنصف من العمر على اختبار بينيه. نقطة الأصل تعبر عن

متوسط ذكاء أطفال في الثالثة والنصف من العمر على اختبار بينيـــه. وبذلك يعتبر (ثيرستون) من الإشارات المبكرة لنظرية الاستجابة للمفردة قبل تسميتها باسمها الحالي لأنهما يشتركان في النقاط التالية:

1. كلاهما يضع نموذج استجابة يتضمن احتمالية الإجابة الصحيحة، باعتبارها دالـــة لمتغير مستمر يقيس سمة في الفرد، فهو متغير القدرة في نظرية الاستجابة للمفردة، ومتغير العمر الزمني عند ثيرستون.

2. كلاهما يفترض أن النموذج يتضمن بارامترات تصف سمات المفردة، موقع المفردة على متصل العمر عند ثيرستون، وبارامترات المقياس لدالـــة التوزيع الطبيعي في نظرية الاستجابة للمفردة.

3. موضوع التحليل نفسه، حيث يستهدف كلا الموضوعين تمثيل موقع المفردة (Item Location)، و استجابات الفرد (Respondent values) باعتبارها نقاط على مقياس المتغير الكمـــي، وهذه تعتبر نقطة مهمة لأن تصحيح مقياس (بينيه) لا يعتمد على العدد الكلـــي والاستجابات الصحيحة، بل على عدد الاستجابات الصحيحة لمفردات التدريج العمري التي يمكن للطفل إعطاؤها .(ديانا، 2010، ص.37)

2- افتراضات نماذج نظرية الاستجابـــة للمفردة الاختبارية أحاديـــة البعد:

تستند نماذج الاستجابة للمفردة أحادية البعد إلى افتراضات قوية ينبغي أن تتوفر في البيانات المتعلقة بالاختبارات لكي تؤدي إلى نتائج يمكن الوثوق بها، حيث يتم اختبار النموذج الملائم للبيانات وفقاً لمدى تحقق هذه الافتراضات في البيانات، حيث أن ذلك يؤدي إلى دقة وإحكام القياس، وفيما يلي توضيح بشيء من التفصيل لهذه الافتراضات الأساسية، كالتالي:

2-1- أحاديـــة البعد(Unidimensionality):

تفترض نماذج الاستجابة للمفردة وجود قدرة واحدة تفسر أداء الفرد في الاختبار. فمن الممكن أن يمثل الاختبار قدرة أحادية البعد بالنسبة لمجموعة من المختبرين وفي الوقت نفسه لا يمثل القدرة على الأحادية نفسها لمجموعة أخرى من المختبرين، فمثلا في حالة المسائل اللفظية المتعلقة بالرياضيات، يتوقع أن يكون الفهم القرائي عاملاً هاماً في حل هذا النوع من المسائل، فإذا كان أفراد العينة (المختبرين) يفهمون الأسئلة فهما لفظياً جيداً ففي هذه الحالة تكون القدرة الرياضية هي السمة الوحيدة وراء أدائهم في الاختبار، وعلى هذا يكون الاختبار أحادي البعد، بينما لو كانت مجموعة المختبرين منخفضي الكفاءة في جانب الفهم القرائي،

ففي هذه الحالة تكون كل من القدرة الرياضية، والقدرة على الفهم القرائي هما معاً وراء الأداء في الاختبار، وبهذا لا يكون هناك بعدا وحيدا وراء الأداء في الاختبار.

استخدم كل من هامبلتون و تروب (Hambleton and Traub, 1973) إجراء التحليل العاملي للتحقق من افتراض أحادية البعد، واقترح كل من لورد و نوفيك (Lord and Novidke,1968) استخدام معاملات الإرتباط الرباعية لاختبار أحادية البعد أكثر من استخدام معاملات إرتباط (∅) (Phi)، بسبب وجود بعض المشكلات المتعلقة بها، كعدم ثابتها عبر العينـــات المختلفة. واقترح ريكاس (Reckase,1979) طريقة أخرى لاختبار أحادية البعد، وهي مصفوفة معاملات الارتباط للمفردة البينية التي تسمح لقيم الجذور الكامنة رسمها بيانياً من الأكبر إلى الأصغر. وتلك الرسومات البيانية يمكن أن تستخدم لتحديد ما إذا كان هناك عامل بارز من عدمه. كما استخدم ماكدونالد (McDonald,1980)، و هاتي (Hattie,1981) تحليل البواقـــي والتحليل العاملي غير الخطي لاختبار أحاديـــة البعد للمفردات الاختبارية. واعتبر هامبلتون و سوامينثان (Hambleton and Swaminathan,1985) أن تحليل البواقي يكون أوضح باستخدام اختبارات مميزة لإختلاف الكمية التي يتم تفسيرها.

ويرى بي هيوشين (Pei-Huachen, 2005)، أنه من الصعب تحقيق هذا الافتراض على أكمل وجه، ويرجع ذلك إلى أن هناك بعض العوامل الشخصية والمعرفية، وكذلك أيضاً ما يتعلق بالاختبار من تعليمات وتوجيهات كلها مجتمعة تؤثر في أداء الأفراد في الاختبار حتى ولو كان هذا التأثير بدرجة ضعيفة. وهذه العوامل تشمل مستوى الدافعية، قلق الاختبار، القدرة على الإجابة بسرعة، وكذلك المهارات المعرفية التي يمكن أن تقاس من خلال مفردات الاختبار. (ممي السيد، 2014، ص.56)

2-2- الاستقلال الموضعـــي(Local Independence):

ويقصد به أن احتمال الإجابة الصحيحة للفرد على مفردة اختبارية يكون مستقلا عن ناتج إجابته على أي مفردة أخرى في الاختبار، عند ضبط كل من القيمة التقديرية لقدرته، والقيمة التقديرية لصعوبة المفردة، لذلك فإن الإستقلال الموضعي يشير إلى أن مفردات الاختبار تكون غير مرتبطة. وربما يبدو هذا متعارضاً مع أحد مبادئ نظرية القياس الكلاسيكية التي تؤكد أهمية الإتساق الداخلـــي للاختبارات، أي أنه كلما كانت درجات المفردات مرتبطة زاد ثبات هذه الدرجات.

غير أن هذا التعارض ربما يزول إذا علمنا أنه في اطار النماذج المعاصرة أحادية البعد تكون إجابات الأفراد على المفردات مشروطة بمستوى قدرة كل منهم، وصعوبة المفردات. أي أن الاستقلال الموضعي يدل على أنه لا توجد ارتباطات باقية بين المفردات عندما يتم إشراط (Conditioned) القيم التقديرية لقدرة الفرد، وصعوبة المفردة. ولعل هذا يشير إلى أن مفردات الاختبار تقيس بعداً واحداً.

ولكي يتحقق هذا الافتراض ينبغي أن يكون نمط استجابـــة الفرد أي واحد (صحيح أو صفر) مساوياً حاصل ضرب احتمالات حصوله على درجة معينة في كل مفردة من مفردات الاختبار. وهذا يتحقق إذا كانت جميع المفردات تقيس سمة كامنة أو قدرة واحدة. فأحادية البعد تتطلب إذن وجود سمة كامنة واحدة بحيث يكون جميع أزواج المفردات تحقق افتراض الاستقلال، وهذا يعني أن المفردات يجب أن تكون مستقلة إحصائياً لأي مجتمع فرعي متجانس من الأفراد فيما يتعلق بالسمة الكامنة. لذلك فإن الاستقلال الموضعي يتعلق أيضاً بعدد السمات أو القدرات التي ينطوي عليها الأداء على المفردات، أي أن الاستقلال الموضعي يعد مؤشرا لأحادية البعد إذا كان النموذج المستخدم يحدد قيمة تقديرية لقدرة الفرد على بعد أحادي. غير أن الاستقلال الموضعي، وأحادية البعد، ليسا مفهوماً واحداً، فالاختبار يكون مقياساً لبعدين إذا وجدت سمتان كامنتان، بحيث أن المفردات تكون مستقلة بالنسبة للأفراد المتجانسين في كل منهما. ويمكن القول بعامة أن أبعاد اختبار معين يساوي عدد السمات الكامنة المطلوبة لتحقيق الاستقلال الموضعـــي. ومن هنا يتبين أن الاستقلال الموضعي يعد افتراضاً ربما يصعب تحققه في كثير من الاختبارات، وبخاصة الاختبارات العقلية. (علام، 2013، ص.63)

وأشار حمادنة (2011) في هذا السياق إلى العديد من المؤشرات الإحصائية التي تستخدم في الكشف عن الارتباط الموضعي بين المفردات التي تتطلب الاستجابة الثنائية، ومن بين أهم هذه المؤشرات:

1. مؤشر (G^2) للكشف عن الفروق بين ما هو متوقع وما هو ملاحظ لتكرار الاستجابات لزوج من المفردات.
2. مؤشر فيشر (Fisher Z) يتم من خلاله تحويل الأخطاء الملاحظة إلى أخطاء معيارية، بمتوسط مقداره صفرا، ومن ثم فإن أي قيمة من قيم مؤشر فيشر للمفردات إذا كانت أعلى بمقدار انحرافين معياريين يمكن اعتبارها مفردة غير مستقلة موضعياً.

ومن أبرز المؤشرات كذلك والتي تستخدم للكشف عن الارتباط الموضعي بين المفردات، هو مؤشر الإحصائي (Q3) الذي اقترحته ين (Yen) والذي تدل قيمته على قيمة عدم الاستقلالية المحلية في أنماط

الاستجابات التي يعطيها الأفراد. وتتلخص طريقته في إيجاد معامل الارتباط بين البواقي في اجابات الأفراد عن أي من أزواج المفردات بعد عزل النموذج المعتمد. وبطريقة أخرى يتم بعد تقدير كل من معالم المفردات وقدرات الأفراد إيجاد التوقع لاستجابات تلميذ محدد على كل من المفردات، ثم بعد ذلك يتم إيجاد قيمة الفرق بين توقع استجابة التلميذ والاستجابة الفعلية التي تم الحصول عليها ليشكل هذا الفرق ما يسمى بالباقي في استجابة التلميذ عن المفردة المحددة. بعد ذلك يتم إيجاد معامل ارتباط قيم البواقـــي لأي زوج من المفردات لجميع التلامذة. وبمعرفة أن التوقع لمعامل الارتباط بين البواقي يساوي $\frac{1-}{N-1}$ عند تحقق الاستقلالية في استجابات التلامذة عندما يكون حجم العينة كبيراً، فإن ذلك يعني استخدام هذه الحقيقة في فحص الاستقلالية من خلال قيمة (Q3) المستخرجة والذي يدل على عدم الاستقلالية المحلية في حال أن تكون قيمة هذا الإحصائي كبيرة (التقي، 2013، ص.147). ويعتبر مؤشر (Q3) معامل ارتباط للبواقي لزوج من المفردات بعد ضبط السمة المقدرة، وحسب زينسكي وهامبلتون وسيريسي (Zenisky and Hammbleton and Sireci, 2002)، و أيلا (Ayala, 2009, p.133)، أن هذا الاحصائي من أكثر الأساليب كفاءة في الكشف عن افتراض الاستقلال الموضعي والذي يختبر علاقة البواقي ما بين أزواج المفردات، ولحساب مؤشر (Q3) يجب تقدير القدر $\widehat{(\theta\imath)}$ لكل مفحوص، وذلك ليتم استخدام هذه القدرة لتقدير احتمال الإجابة الصحيحة للمفحوص الذي يمتلك هذه القدرة لكل المفردات. ويتم حساب الباقي ويرمز له بالرمز (dij)، وذلك بأخذ الفرق بين الأداء الملاحظ والأداء المتوقع للمفحوص، ويمكن التعبير عن ذلك رياضياً بالمعادلة رقم (08):

$$dij = xij - pj\widehat{(\theta\imath)} \quad \ldots\ldots\ldots\ldots\ldots\ldots(08)$$

باعتبار أن (xij) يمثل العلامة الملاحظة للشخص (i) على المفردة (j) وهي إما واحد عند الاستجابة الصحيحة أو صفر عند الاستجابة الخاطئة. و $pj\widehat{(\theta\imath)}$ تمثل احتمال اجابة المفردة (j) اجابة صحيحة من قبل المفحوص (i)، ويمثل مستوى بواقي الشخص والمفردة (k) و يعطي بالمعادلة رقم (09):

$$dik = xik - pk\widehat{(\theta\imath)} \quad \ldots\ldots\ldots\ldots\ldots\ldots(09)$$

وبالأخذ القيمة المطلقة لمؤشر |Q3|، تكون قيمة المؤشر ما بين الصفر والواحد الصحيح المطلقة. وكلما اقترب من الصفر فهو أفضل، وترى ين (Yen,1984) أن قيمة هذا المؤشر تميل لأن تكون سلبية، وخاصة في المقاييس والاختبارات القصيرة، وبالتالي لو أخذنا المفردتين (j,k)، فإن (Q3) يكون معامل الارتباط بين البواقي لهاتين المفردتين لجميع المفحوصين (p.125-145). ويمكن التعبير عنه بالمعادلة رقم (10):

$$Q3jk = rdjdk \ldots\ldots\ldots\ldots\ldots\ldots(10)$$

واعتبر المعيار 1-L/1- باعتبار (L) عدد المفردات داخل الاختبار، مؤشراً على تحقق افتراض الاستقلال الموضعي (Ayala, 2009, p.133). وقد اعتمد ريز (Reese,1999) في الحكم على مستويات الانتهاك بناءً على هذا الأسلوب من خلال القيم، التالية: 0.30؛ 0.05؛ 0.01؛ 0.000 بوصفها مستويات للانتهاك الصفري والمتدني والمتوسط و العالي على التوالـــي. وأشار زينسكي وهامبلتون وسيريسي (Zenisky and Hammbleton and Sireci, 2002) إلى أن قيم هذا الإحصائي عادة ما تأخذ قيماً صغيرة، و أن أعلى القيم له التي تم ملاحظتها في عديد الدراسات كانت تحوم حول القيمة 0.10 . (النعيمي، 2011 ، ص.162)

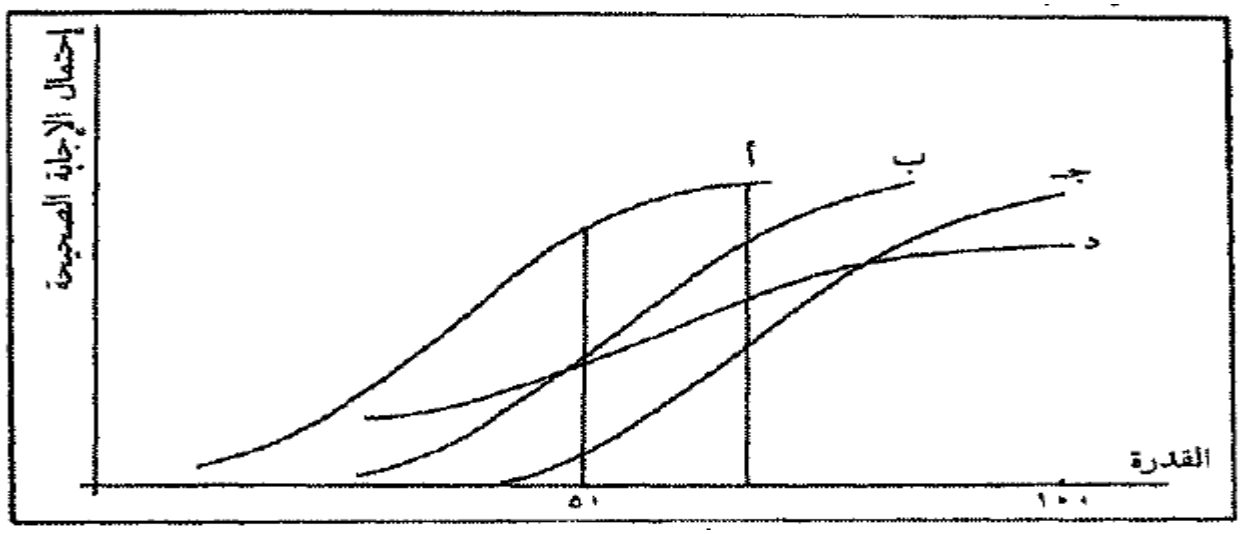
وكقاعدة عامة لتقييم مؤشر الإحصائي (Q3) للحكم على الاستقلال الموضعي أن لا يكون ارتباط أزواج الفقرات من خلال هذا المؤشر (Q3) أكبر من القيمة 0.2 أو أصغر من القيمة −0.2. (Chen ;1997 Thissen,)، أي اعتبار القيمة المطلقة للمؤشر Q3 أكبر من $|Q3| > 0.2$ مؤشراً للتبعية أو الاعتماد (Ayala, 2009, p.133). (Francis et al, 2001, pp.159-174) (Chrisitine, 2010, p.48)

2−3− توازي المنحنيـــات المميزة للمفردات (Item Characteristic Curve (ICC)):

تتفق كل من النظرية الكلاسيكية والحديثة في القياس على فرضية مؤداها وجود متصل للقدرة، وإذا علم موقع الفرد على هذا المتصل، فإنه يمكن تقدير احتمال إجابة فرد ما إجابة صحيحة على مفردة اختبارية تقيس هذه القدرة، إلا أن كلا منهما تختلفان في الدالة التي يتم بواسطتها تحديد موقع الفرد وارتباطه باحتمال الإجابة الصحيحة، حيث أن نماذج النظرية الكلاسيكية تفترض أن شكل الدالـــة يتخذ شكل منحنى التوزيع الاعتدالـــي في الاختبارات جماعية المرجع، وشكلاً ملتوياً إلتواء سالباً في الاختبارات محكية المرجع. في حين أن نماذج نظرية الاستجابة للمفردة تفترض أنها تتخذ شكل المنحنـــى اللوغاريتمي، وبالتالي وجود دالة مميزة خاصة بكل مفردة من مفردات الاختبـــار تعتمد في أساسها على موقع الفرد والمفردة معاً على متصل القدرة وتتخذ شكل المنحنى اللوغاريتمي التراكمي. (وليد، 2014، ص.17)

وترى كاظم (1988) أن تساوي قوة التمييز لمفردات اختبار ما يرتبط بتحقق شرط استقلالية القياس، حيث تتساوى الصعوبة النسبية للمفردات عبر مستويات القدرة و تعرف قوة تمييز المفردة بأنها معدل التغيير الذي يعرف بأنه الميل النسبي للمنحنى المميز للمفردة على محور القدرة وتقل قوة المفردة على التمييز بين

الأفراد على مستويات القدرة المختلفة أو تزيد تبعاً لإنحناء المنحنى المميز للمفردة، كما يوضحه الشكل رقم(05) لأربعة منحنيات مميزة لأربعة مفردات (أ، ب، جـــ، د) حيث يلاحظ ما يأتي:



شكل (05) يوضح المنحنيات المميزة لأربع مفردات. (كاظم، 1988أ، ص.72)

1. تتوازي تقريبا المنحنيات (أ)، (ب)، (جـــ)، أي أن لها شكلاً وإنحناء عاماً واحداً. ومعنى هذا أن الزيادة المتساوية في مستوى القدرة (مستوى الدرجة الكلية)، يقترن بزيادة متساوية تقريباً في احتمـــال الإجابة الصحيحة على أي من هذه المفردات. وبعبارة أخرى أن قوة تمييز المفردات بين قدرات الأفراد متساوية، وهذا ما يختلف بالنسبة للمنحنى (د).

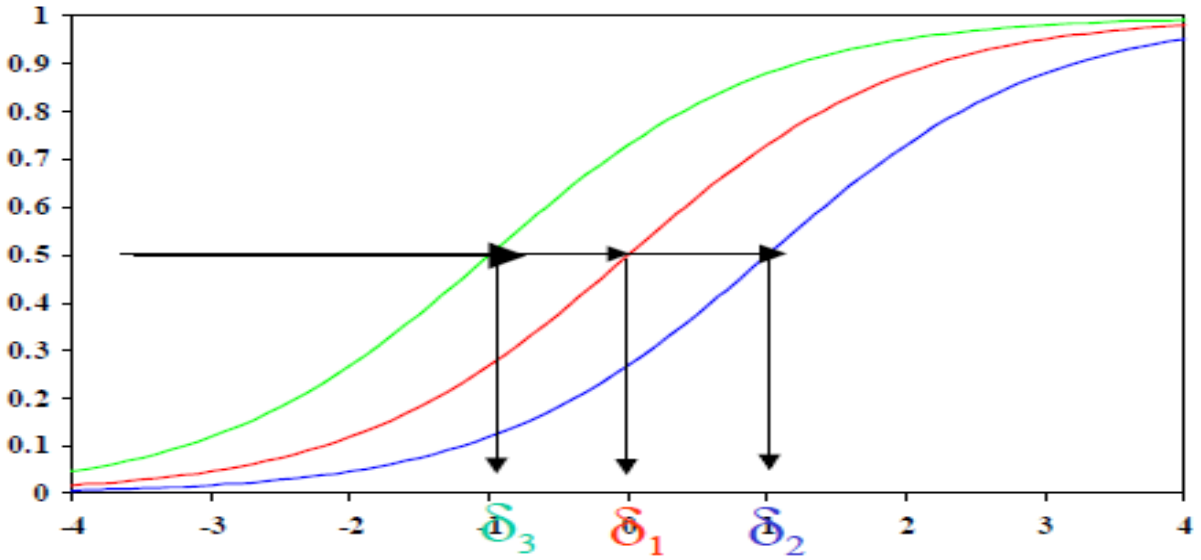
2. إن فاعلية المفردة (أ) تمتد من المستوى المنخفض من القدرة، وحتى المستوى فوق المتوسط مباشرة، وليس لهذه المفردة فائدة تذكر للأفراد ذوي المستوى العالي من القدرة ولا تستطيع أن تميز بينهما على هذه القدرة حيث يمكن لهؤلاء الأفراد جميعاً الإجابة الصواب على هذه المفردة . أما المفردة (جـــ)، فتمتد فاعليتها من المستوى العالي من القدرة، وليس لهذه المفردة فائدة تذكر للأفراد منخفضي القدرة . حيث يخفق هؤلاء جميعا في الإجابة على هذه المفردة، ومن ثم فهي لا تستطيع أن تميز بينهما. أما المفردة (ب)، فتمتد فاعليتها عبر المستويات المتوسطة من القدرة.

3. في المدى المشترك لفاعلية هذه المفردات الثلاثة (أ، ب، جـــ)، يكون احتمال الإجابة الصواب على المفردة (أ) أكبر دائما من احتمال الإجابة صواب على المفردة (ب)، وهذا يكون دائما أكبر من احتمال الإجابة الصواب على المفردة (جـــ)، وذلك عند أي مستوى من مستويات هذه القدرة، المحدود بالمدى المشترك لفاعلية هذه المفردات. أي أن المفردة (أ) تكون دائما أسهل من المفردة (ب)، وهذه تكون دائماً

أسهل من المفردة (جـــ)، وذلك عند المستويات المختلفة من القدرة في هذا المدى المشترك لفاعلية هذه المفردات، أي أن صعوبة هذه المفردات تكون مستقرة عبر مستويات القدرة، مما يعني ملاءمة هذه المفردات للنموذج.

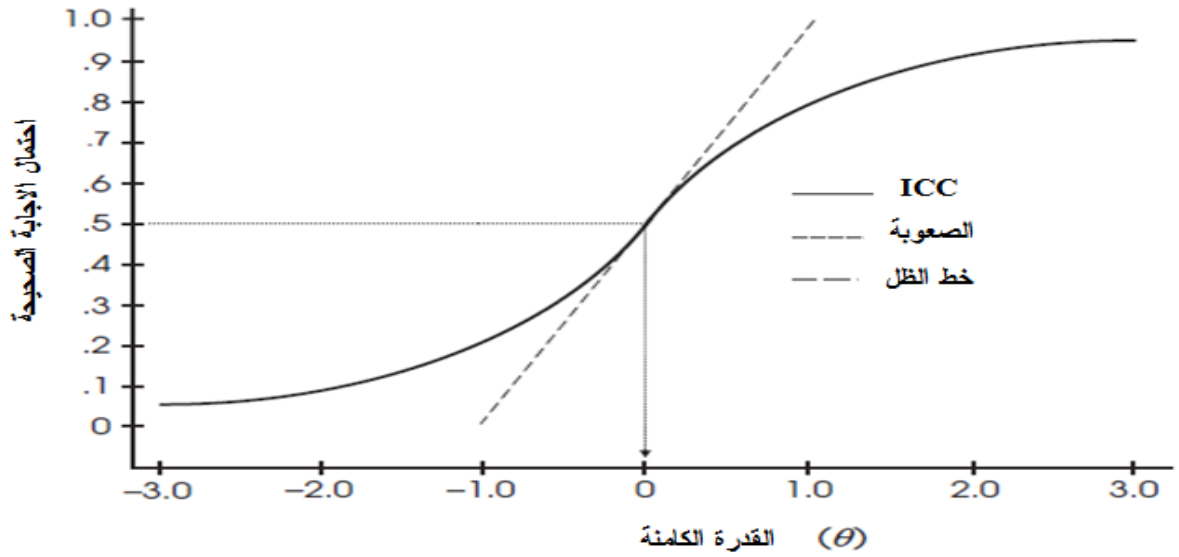
4. إن احتمال الإجابة الصواب على المفردة (د) يكون أكبر من احتمال الإجابة الصواب على المفردة (ب)، عند المستوى المنخفض من القدرة في حين تكون أقل منها عند المستوى العالي من القدرة. معنى هذا أن المفردة (د) تكون أسهل من من المفردة (ب) في المستويات المنخفضة من القدرة بينما يكون أصعب منه في المستويات العالية من القدرة . وهذا ما يمكن استنتاجه أيضا عند مقارنته بالمفردة (جـــ)، حيث تكون المفردة (د) أسهل من المفردة (جـــ) في المستويات المتوسطة من القدرة، في حين يكون أصعب منه عند المستويات العالية من القدرة. ومعنى هذا أن صعوبة هذه المفردة (د) غير مستقرة عبر مستويات القدرة المختلفة. وهو اخلال بفرض النموذج عن استقلال صعوبة المفردة عند قدرات العينة، مما يعني عدم ملاءمة هذه المفردة للنموذج.(كاظم، 1988أ، ص.72).

ويلاحظ من المنحنيات المميزة لثلاثة مفردات في الشكل رقم (06)، التي تم تدريجها باستخدام نموذج (راش) أحادي المعلم لا تتقاطع، ولكن تختلف عن بعضها في الصعوبة فقط وتتلاقى فقط عند المحور الأفقي الذي يمثل متصل القدرة أو السمة .



شكل (06) يوضح توازي المنحنيان المميزة للمفردات بحسب نموذج راش أحادي البارامتر. (Wu, M, 2007, p.31)

ومن ملاحظة الشكل رقم (06) يتضح أن القوة التمييزية لإحدى المفردات في إزدياد ميل المنحنى المميز لها بالقرب من الوسط، أي بالقرب من النقطة على متصل القدرة التي يتوقع عندها أن يكون احتمال الإجابة الصحيحة على المفردة 0.50. فالميل الحاد (Steep Slope) يدل على أن التغيير نسبياً في قدرة الفرد يجعل احتمال إجابته الصحيحة على المفردة يتباين تبايناً ملحوظاً (تمييز مرتفع)، أما الميل البسيط فيدل على أن التغيير الكبير نسبياً في قدرة الفرد لا يغير احتمال الإجابة الصحيحة تغييراً ملحوظاً (تمييز منخفض)، لذلك فإن تمييز المفردة يتناسب تناسباً طردياً مع ميل المنحنى المميز لها عند نقطة انقلابه. و يعد المنحنى المميز للمفردة الاختبارية العمود الفقري لنظرية القياس المعاصرة، كما يعد من الافتراضات الأساسية لنماذج الاستجابة للمفردة أحادية البعد، بينما يعد السطح المميز للمفردة من الافتراضات الأساسية لنماذج الاستجابة متعددة الأبعاد (علام، 2013، ص.62). ويمكن تمثيل منحنى خصائص المفردة (ICC) والتنبأ باحتمال الإجابة الصحيحة في الاستجابة الثنائية، في نموذج الاستجابة الثنائية للمفردة ثنائية المعلم، كما يوضحه الشكل رقم(07):



شكل (07) يوضح منحنى خصائص المفردة (ICC) والتنبأ باحتمال الاجابة الصحيحة في الاستجابة الثنائية، في نموذج الاستجابة للمفردة ثنائية المعلم. صعوبة المفردة عند القدرة صفر، وتمييز المفردة وهو ميل خط الظل عند القدرة صفر.

(Rex .B. Kline, 2015, p.95)

2-4- **الحد الأدنـــى من أثر التخمين (Minimal guessing):**

يفترض نموذج (راش) أحادي البارامتر إنعدام التخمين أو الحصول على أقل قيمة متدنية له، والذي يحدد من خلال ارتفاع المنحنى المميز للمفردة عند متصل القدرة عندما تكون القدرة أقل ما يمكن، وقيمه نظرياً حسب بيكر محصورة مابين 0 – 1، وعمليا مابين 0 – 0.35 وما فوق ذلك فهي قيم ليست مقبولة، ويتحدد بالخط المقارب الأدنـــى (Lower Asymptote) في المنحنى المميز للمفردة. (Baker, 2001, p.28). في حين يفضل وروم (Warm, 1978) أن تكون قيم التخمين 0.2 أو أقل من ذلك، بينما لاحظ هاريز (Harris,1989) أن قيمة معامل التخمين في الواقع العملي تتخذ قيما أقل من 0.3، وأوصى (Urry, 1977) بأن لا يزيد معلم التخمين المنخفض عن 0.3 في الاختبارات بشكل عام، وكلما كان معلم التخمين منخفض كلما كان أفضل، كما اقترح كل من ليناكر (Linacre, 2012)، وستيفن (Steven, 2003)، و (وليد، 2014) أن قيم تخمين المفردات يجب أن لا تتعدى القيمة 0.10.

2-5- **السرعـــة (Speediness):**

ميز كل من لورد و نوفيك (Lord and Novick,1968) بين اختبارات السرعة واختبارات القوة، فاختبار السرعة يعتمد على قدرة الممتحن للإجابة على مفردات عديدة خلال وقت محدد وثابت، أي أن تعتمد الدرجة في اختبار السرعة على معدل الاستجابـــة، أما اختبار القوة يعتمد على الإجابة على مفردات الاختبار دون زمن محدد. ويعرف فاروق عبد الفتاح موسى (1990) كما ورد عن ميمي السيد، أن سرعة الأداء في الاختبار، بأنها مدى اعتماد درجة الفرد في الاختبار على أدائه فيه، وتقاس سرعة الأداء في الاختبار عادة بنسبة الأفراد الذين لم يصلوا ولم يجيبوا عن آخر مفردة في الاختبار. وتفترض نماذج الاستجابة للمفردة أن عامل السرعة لا يلعب دوراً في الإجابة عن مفردات الاختبار، أي أن إخفاق بعض المختبرين في الإجابة على بعض مفردات الاختبار يرجع إلى إنخفاض قدرتهم وليس إلى تأثير عامل السرعة في إجابتهم. ويلاحظ أن افتراض السرعة ليس له تأثير مباشر على المنحنى المميز للمفردة، وإنما يكون تأثيره غير مباشر فعامل السرعة يجعل بعض التلامذة يلجأون إلى التخمين في الإجابة، وبالتالي فإن سرعة الأداء تصبح أحياناً إحدى القدرات التي يقيسها الاختبار.(2014، ص.59)

3- أنواع نماذج الاستجابـــة للمفردة أحاديـــة البعد:

توجد نماذج متعددة في اطار نظرية الاستجابة للمفردة، وذلك لإختلاف الافتراضات المتعلقة بالبيانات الاختبارية، إذ ربما يفترض أن الأداء في الاختبار ينطوي على سمة أحاديـــة البعد، أو على سمات متعددة. أو ربما يفترض أن المفردات تتباين في قدرتها على التمييز بين المستويات المختلفة للقدرة، أو أن الإجابات الصحيحة على مفردات الاختيار من متعدد تتأثر بالتخمين العشوائـــي أو لا تتأثر كما في الأسئلة المفتوحة، أو عما اذا كانت الدرجات ثنائية أم متعددة الأقسام، وهكذا.

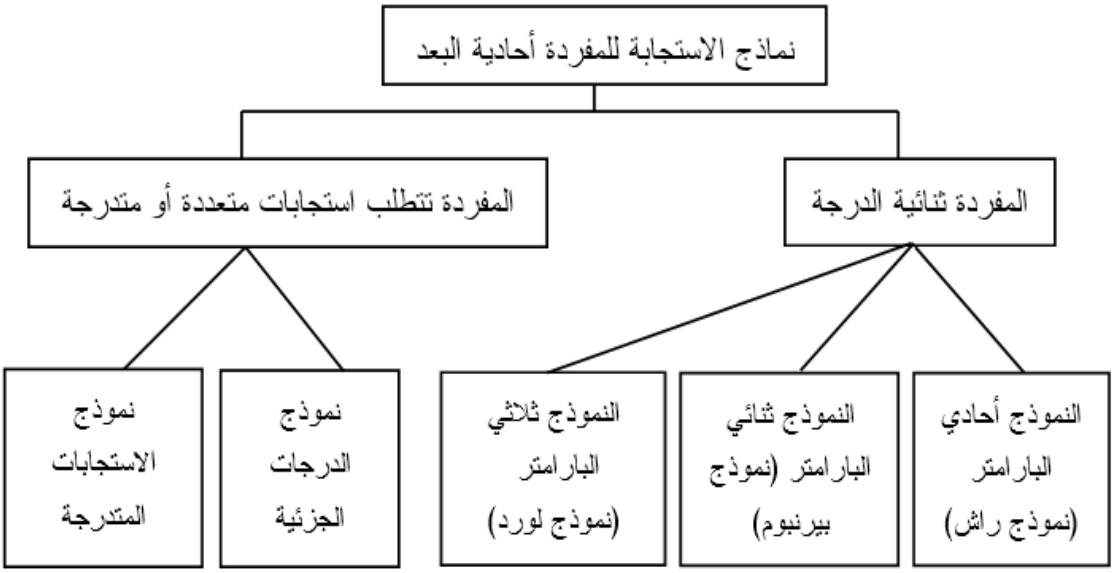
لذلك يمكن تقسيم هذه النماذج الى قسمين رئيسين: أحدهما يسمى نماذج الاستجابة للمفردة أحادية البعد (Unidimensional Models)، والآخر يسمى نماذج الاستجابة للمفردة متعددة الأبعاد(Multidimensional Models). وينقسم كل من نوعى النماذج إلى نماذج الإستاتيكية (Static Models)، ونماذج ديناميكية (Dynamic Models).(بوعلام، 2013، ص67)

3-1- النمـــاذج الإستاتيكيـــة:

يهتم هذا النوع من النماذج بالقياس في مدة زمنية واحدة، وكذلك بتحديد العمليات التي ينطوي عليها الأداء في الاختبارات السيكولوجية والتربوية. وتختلف هذه النماذج في صيغتها الرياضية التي تربط أداء الفرد في الاختبار بدرجات قدرته، كما تختلف في عدد الخصائص أو البارامترات المستخدمة في وصف كل مفردة من مفردات الاختبار. (ميمي، 2014، ص.61)

ويوجد ثلاثة نماذج أساسية شائعة الاستخدام، لكل منها صيغة رياضيـــة تمثلها المنحنيات المميزة للمفردات. وتطبق هذه الصيغ في حالة المفردات ثنائيـــة الدرجة التي لا تتطلب السرعة في الإجابة. وتقيس بعداً واحداً. كما توجد نماذج أخرى تتعلق بالاستجابات الأسمية (Nominal Responses)، والاستجابات متعددة الأقسام (Polytomous Responses)، وتشمل نموذج الدرجات الجزئية (Partial Credit Model)، ونموذج الاستجابات المتدرجة (Graded Response Model)، وغيرها. وهذه المجموعة من النماذج يمكن استخدامها في تحليل البيانات المستمدة من استبيانات الإتجاهات والشخصية، وموازين التقدير، والمفردات الاختبارية التي تتطلب خطوات متعددة، حيث تعين درجة جزئية للخطوات التي استكملت في الإجابة لحل

مشكلة معينة، مثل حل المسائل الرياضية. ويوضح الشكل رقم (08) تخطيط لهذه النماذج: (علام، 2013، ص.67-68)



شكل (08) يوضح أنواع نماذج الاستجابـــة للمفردة أحادية البعد.

-نماذج تتعلق بالمفردة ثنائية الدرجة:

3-1-1- النموذج أحادي المعلم (نموذج راش)(One Parameter logistic Model or Rasch Model):

هو أحد النماذج الرياضية الاحتمالية المرتبطة بنظرية الاستجابة للمفردة، والتي حققت نتائج جيدة في مجال تحقيق موضوعية القياس السلوكي. ويعد هذا النموذج من أشهر نماذج هذه النظرية، حيث يفترض أن السمة المقاسة أحادية البعد، وقد اعتبر كل من هامبلتون وسوامينثان (Hambleton and Swaminthan, 1985) أن افتراض أحادية البعد شيء مرغوب فيه عند جميع مطوري الاختبارات، وذلك لأجل تطوير وتعزيز عملية التفسير لدرجات الاختبار والمقصود بذلك الافتراض أن يكون هناك عامل واحداً سائداً على العوامل الأخرى، وسنناقش هذا النموذج بأكثر تفصيل لاحقاً في هذا الفصل.

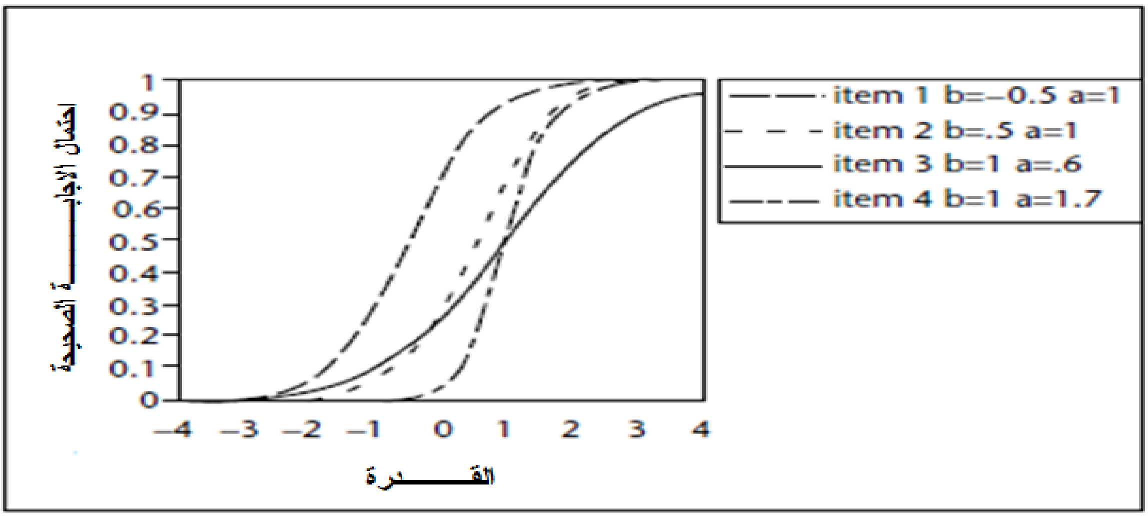
3-1-2- النموذج ثنائي البارامتر (نموذج بيرنبوم) (Birnbaum,1968):

اقترح عالم الإحصاء بيرنبوم (Birnbaum,1968) مع مجموعة من زملائه بجامعة كولومبيا الأمريكية هذا النموذج ((Tow-Parameter Logistic Model (2PL)، وهو على عكس نموذج راش يسمح للمفردات بأن تختلف في كل من بارامتري الصعوبة والتمييز. ويفتقر هذا النموذج إلى بعض الخصائص الإحصائية التي يتميز بها نموذج راش، لذلك فإن عملياته الحسابية أكثر صعوبة. وقد نشر هذا النموذج في كتاب لورد،

ونوفيك (Lord and Novick,1968) " النظريات الإحصائية لدرجات الاختبارات العقلية". لذلك تضمنت الصيغة الرياضية لهذا النموذج بارامتر تمييز المفردة، وأصبح يشتمل النموذج على بارامترين (الصعوبة، والتمييز) يمثلان خصائص المفردة، والذي يعطى بالمعادلة رقم (11):

$$P_i(\theta) = \frac{e^{a_i(\theta_i - b_i)}}{1 + e^{a_i(\theta_i - b_i)}} \qquad \text{.......................................(11)}$$

ويلاحظ من الصيغة رقم (11)، أن تمييز المفردة (a_i) يضرب في الفرق بين مستوى القدرة، وصعوبة المفردة. وهذا الفرق يكون له تأثير كبير في الاحتمـــالات للمفردات الأكثر تمييزاً. وتضمين بارامتر التمييز في هذا النموذج يجعل تقدير القدرة معتمداً على النمط العام لإجابات الأفراد الصحيحة، وإجاباتهم الخطأ على المفردة. إذ يمكن أن نحصل على تقديرات مختلفين لمستوى السمة من نمطين مختلفين درجاتهما الكلية واحدة. فتوصل الفرد إلى الإجابة الصحيحة على مفردات تمييزها مرتفع، واخفاقه في الإجابة على مفردات تمييزها منخفض، يؤدي إلى أعلى تقدير لمستوى السمة، لذلك فإن هذا النموذج لا يعين أوزاناً متساوية في تقدير مستويات السمة على عكس نموذج راش. ويترتب على ذلك امكانية تقاطع المنحنيات المميزة لمفردات الاختبار الذي يتم اعداده استناداً إلى هذا النموذج، كما يوضحه الشكل رقم (09):



شكل (09) التمثيل البياني لأربعة مفردات في النموذج ثنائي البارامتر.

يتضح من الشكل رقم (09)، أن المنحنيات المميزة لأربعة مفردات تختلف في الجزء الذي تلتقي فيه بالمحور الأفقى (وهذا يعني أن المفردات الأربعة تختلف في صعوبتها)، كما يختلف المنحنيات الأربعة في

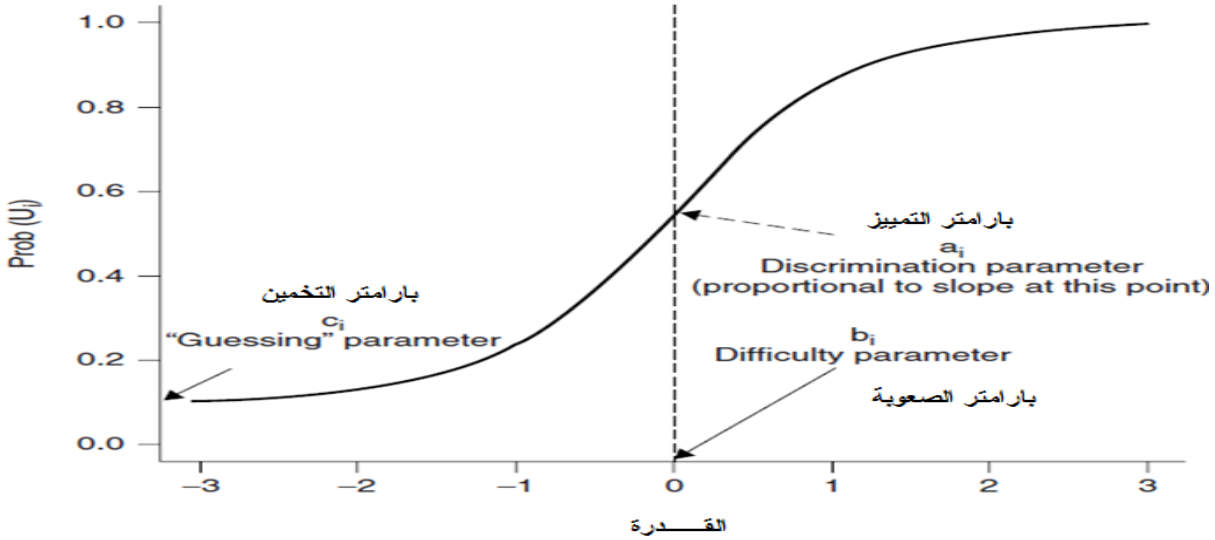
الميل (Slope) فميل المنحنى المميز للمفردة 3 أكبر من ميل منحنى بقية المفردات، مما يدل على أن المفردة 3 تميز بين الأفراد بدرجة أكبر من بقية المفردات 1،2،4.

3-1-3- النموذج ثلاثـــي البارامتر (نموذج لورد):

إن نموذج راش، والنموذج ثنائي البارامتر لا يأخذان بعين الاعتبار أن احتمال الإجابة الصحيحة على مفردة اختيار من متعدد يكون أكبر بالنسبة للأفراد من ذوي القدرة المنخفضة لأنهم ربما يلجأون إلى التخمين العشوائي. ففي نموذج راش تميل قيمة ($\varnothing$) Pi إلى الاقتراب من الصفر كلما صغرت قيمة($\varnothing$). ومع هذا يمكننا توقع أنه بالنسبة للأفراد من ذوى القدرة المنخفضة تكون نسبة من يجيبون منهم إجابة صحيحة على مفردة معينة سوف يكون أكبر من الصفر، وذلك لأنهم ربما يمكنهم تخمين هذه الإجابة، لذلك فإن النموذج ثلاثي البارامتر (Three-Parameter Logistic Model (3PL)) أضاف بارامتر ثالث يمثل احتمال توصل هؤلاء الأفراد إلى الإجابة الصحيحة على المفردة عن طريق التخمين، وعندئذ يجب أن يؤخذ ذلك بعين الاعتبار عند مطابقة البيانـــات المستمدة من الاختبار لهذا النموذج. والصيغة الرياضيـــة للنموذج ثلاثى البارامتر(Lord's Model) والذي يعطى بالمعادلة رقم (12):

$$P_i(\theta) = c_i + (1 - c_i)\frac{e^{a_i(\theta_i - b_i)}}{1 + e^{a_i(\theta_i - b_i)}} \quad \dots\dots(12)$$

يسمى البارامتر (C_i) الخط التقاربي الأدنـــى (Lower Asymtote) للمنحنى المميز للمفردة، ويمثل احتمال الإجابة الصحيحة للأفراد من ذوى القدرة المنخفضة. والهدف من إضافة هذا البارامتر في النموذج هو محاولة مراعاة عدم مطابقة المنحنيات المميزة للمفردة عند الطرف الأدنى لمتصل السمة، حيث يكون التخمين العشوائي أحد المتغيرات التي تؤثر في الأداء على الاختبار. ويمكن تمثيل هذه الصيغة الرياضية بالشكل رقم (10): (Charles et al, 2012, p.246)



الشكل (10) **التمثيل البياني للنموذج ثلاثي البارامتر.**

يتضح من الشكل رقم (10)، منحنى لمفردة اختبارية تمثل بثلاثة بارامترات وهي (الصعوبة، والتمييز، والتخمين)، حيث يمثل الخط التقاربي الأدنـــى، وهو الخط الذي يلتقى بالطرف الأسفل للمنحنى في اللانهاية (بارامتر التخمين). إلا أن لورد (Lord) لا يفضل تسمية هذا البارامتر ببارامتر التخمين، وإنما يقترح تسميته "بارامتر مستوى شبه الصدفة Pseudo-Chance Level Parameter". (علام، 2013، ص.74)

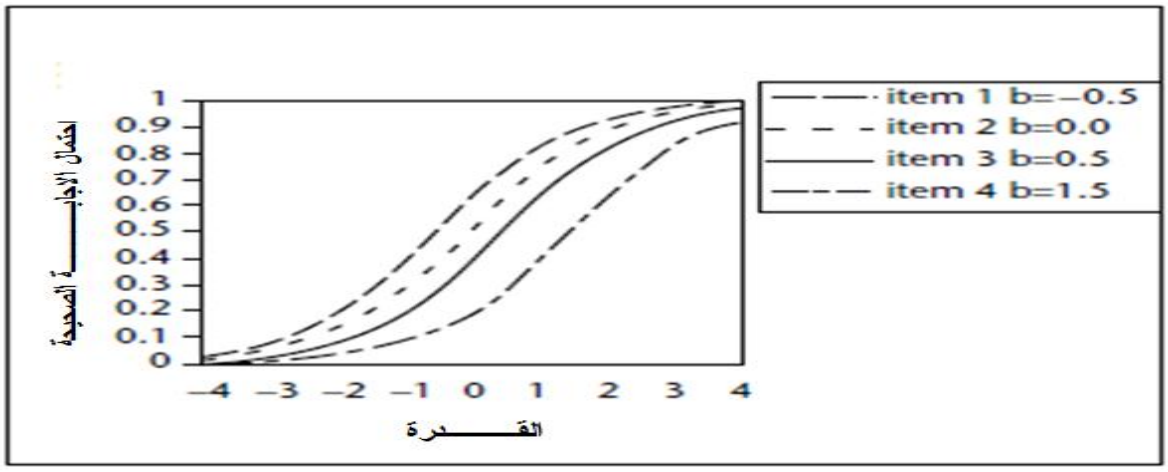
4- نموذج راش أحـــادي المعلم (البارامتر):

يعد هذا النموذج أبسط نماذج الاستجابة للمفردة أحادية البعد، ويعرف بإسم نموذج راش (Rasch Model) نسبة إلى عالم الرياضيات الدانماركي جورج راش بجامعة كوبنهاجن (Rasch,1960). وعندما كان راش (1901−1980) على قيد الحياة عارض بشدة تسمية هذا النموذج بإسمه، حيث أن هذا النموذج لا يعد نموذجاً واحداً بل عدة نماذج. والنموذج الذي يعتمد على الدرجة الثنائية هو واحد من مجموعة نماذج إحصائية أطلق عليها راش إسم "نماذج القياس Measurement Models"، وليس نموذج راش. وقد طور هذا النموذج بطريقة مستقلة عن غيره من نماذج الاستجابة للمفردة، ويمكن اعتباره نموذجاً يكون فيه المنحنى المميز للمفردة تمثله دالـــة ترجيح لوغاريتمـــي أحادية البارامترات (One-Parameter Logistic Function). والمتغير التابع في نموذج راش البسيط هو احتمال أن فرد (j) يجيب إجابة صحيحة على مفردة اختبارية (i). أما المتغيرات الكامنة المستقلة فهي درجة قدرة الفرد (∅)، وصعوبة المفردة (bi). وتضم المتغيرات المستقلة بعملية جمع، كما أن صعوبة المفردة (bi) تطرح من قدرة الفرد (∅). وعلاقة هذا الفرق بالاستجابة

للمفردة يعتمد على اختيار المتغيرات المستقلة التي يتم نمذجتها في صيغة رياضية h احتمالية، وهي دالـــة ترجيح لوغاريتمـــي غير خطية (Nonlinear Logistic Function)، والتي تعطى بالمعادلة رقم (13):

$$P_i(\theta) = \frac{e^{D(\theta - b_i)}}{1 + e^{D(\theta - b_i)}}, \quad i = 1,2,\ldots,n \qquad \ldots\ldots(13)$$

حيث أن :(i): المفردات وتساوي (1، 2، .....، n)؛ و Pi (∅): إلى احتمالية أن يجيب المستجيب الذي قدرته (∅)على المفردة ؛ و(bi): صعوبة المفردة (i)؛ و(e): إلى الأساس اللوغاريتمي الطبيعي وهو يساوي 20,718 تقريباً؛ و(D): إلى معامل القياس أو التدريج (*scaling factor*) وهو مقدار ثابت يساوي 1,7 أو 1,702، ويستخدم هذا المعامل ليؤكد أن الدالـــة اللوغاريتمية هي تقريب حقيقي للدالة التجميعية الطبيعية لهذا الأنمـــوذج، وبدون هذا المعامل تصبح الدالة اللوغاريتميـــة تحويلاً خطياً للدالة التجميعية الطبيعية. وتسمى هذه الصيغة "نموذج الترجيح اللوغاريتمي أحادي البارامتر 1PL"، وذلك لإنها تستخدم الدالة الأسية في التنبؤ بالاحتمالات، وتشتمل على بارامتر (معلم) واحد فقط، وهو بارامتر صعوبة المفردات لتمثيل الفروق بين المفردات (Mislevy, 1990, p.195) (علام، 2013، ص.69). ويمكن تمثيل هذه الصيغة الرياضية بالمنحنى المميز لأربعة مفردات تقيس سمة كامنة واحدة، كما يوضحه شكل رقم (11):



شكل (11) التمثيل البياني لصيغة نموذج راش لإربعة مفردات مختلفة الصعوبة.

يتضح من الشكل رقم (11) أن المنحنيات الأربعة متوازيـــة تختلف فقط في إزاحة الموقع أي في (الصعوبة)، حيث أن المفردة 4 تبدو أكثر صعوبة من بقية المفردات.

4-1-معنى الموضوعية في نموذج راش:

نعني بالموضوعية في نموذج راش أحادي البارامتر، موضوعية المقارنة بين نتيجة تفاعل قدرتي فردين مع صعوبة مفردة مناسبة، كما تعني أيضا موضوعية المقارنة بين صعوبة مفردتين استجاب لهما فرد مناسب، ولا يكون التقدير الكمـــي لهذه المقارنة دقيقا تماماً، بل تتراوح لمدى معين على جانبي هذا التقدير، أي تقل أو تزيد عنها نوعاً ماً، وتبدو هذه الموضوعية من ناحيتين:

4-1-1- استقلال بارامتر قدرة الفرد عن المفردة المستخدمـــة:

بالرغم من أن المقارنة بين قدرتي فردين يعتمد على استخدام مفردة ملائمة، إلا أن هذه المقارنة لا تتأثر باستخدام أي مفردة كانت من هذه المفردات الملائمة، وكما سبق ذكره، فإن المفردات الملائمة تعني المفردات الصادقة في تعريفها للمتغير موضوع القياس، أي الصادقـــة في قياسها للصفة وفي تدرجها عن متصلها، كما تعني أيضاً تساوي قوتها على التمييز بين مستويات الأفراد على الصفة. وهكذا فإن معنى موضوعية المقارنة بين قدرتي فردين، هو استقلال هذه المقارنة عن المفردة المستخدمة، أي أن استجابة الفردين لأي مفردة من المفردات الملائمة، ينبغي أن تؤدي إلى نفس التقدير الكمي للمقارنة بين قدرتي هذين الفردين (باعتبار الخطأ المعياري على جانبي هذا التقدير). وعلى هذا فإن ما نصل إليه ليس هو القدرة المطلقة للفرد (V) مثلاً، وإنما هو بعده عن قدرة فرد آخر وليكن (U)، وهذا الفرق يجعل قدرة الفرد (U) كنقطة أصل تقاس منها قدرة الفرد (V).

4-1-2- استقلال بارامتر صعوبـــة المفردة عن الفرد الذي يجيب عليها:

بالمثل ورغم أن المقارنة بين صعوبة مفردتين تعتمد على استجابة فرد ملائم، إلا أنها لا تتأثر باستجابة أي فرد يكون من الأفراد الملائمين. ويكون الفرد ملائماً، عندما تكون استجابته على المفردة استجابة صادقة، لا تختلف كثيراً عن الاستجابة المتوقعة لمعظم الأفراد في مستوى هذا الفرد. وهكذا فإن موضوعية المقارنة بين صعوبة مفردتين يعني استقلال هذه المقارنة عن الفرد الذي يجيب على هذين المفردتين، أي أن استجابة أي من الأفراد الملائمين على هذين المفردتين، ينبغي أن يؤدي إلى نفس التقدير الكمي للمقارنة بين صعوبتي هذين المفردتين (مع اعتبار الخطأ المعياري على جانبي هذا التقدير). وعلى هذا فإن ما نصل اليه ليس هو الصعوبة المطلقة للمفردة (i) مثلا، وإنما هو بعدها عن صعوبة مفردة أخرى ولتكن (C). وهذا الفرق يجعل

صعوبة المفردة (C) كنقطة أصل تقاس منها صعوبة المفردة (i)، وكلما توافقت صعوبة المفردات المستخدمة، مع مستوى قدرات الأفراد الذين يجيبون عليها، كلما توصلنا إلى هذه الموضوعيـــة في القيـــاس. (كاظم، 1988ب، ص.26)

4-2-الجانب السيكولوجـــي في نموذج راش:

يتضمن نموذج راش واقعاً سيكولوجيـــاً في صورة رياضيـــة، يعبر تعبيرا واقعيا عن ديناميكية التفاعل بين الأفراد ومفردات الاختبارات، فالمعادلة تفسر احتمال أن يجيب فرد إجابة صحيحة عن مفردة اختبار ما، وقد سبق أن ذكرنا أن نموذج راش يقوم على قضيتين أساسيتين تتلخصان، في أنه كلما زادت قدرة الفرد زاد احتمـــال إجابته إجابة صحيحة عن مفردة اختبار ما، وأن هناك احتمـــال كبير أن يجيب أي فرد إجابة صحيحة عن مفردة تتميز بالسهولة بالنسبة له. فإذا كانت هاتان القضيتان اللتان تقرران ديناميكية التفاعل بين الأفراد ومفردات الاختبارات صحيحتين، فإنه يصبح نموذج راش الذي تمثله هذه المعادلة بمثابة تعبير واقعي عن هذا التفاعل، كما يصبح النموذج الوحيد من بين النماذج التي عرضناها الذي يفسر أداء الأفراد في الاختبارات والمقاييس. فإذا حققت مجموعة من البيانات المستمدة من هذه الاختبارات متطلبات نموذج راش، فإن هاتين القضيتين يجب أن تكونان صحيحتين. أما إذا لم تحقق البيانات هذه المتطلبات، يكون معنى هذا أن البيانـــات تناقض الأساس المنطقي الذي بنى عليه النموذج.

ولتوضيح ذلك نفترض أن إحدى مفردات اختبار لم "تطابق fit" النموذج، عندئذ لا يكون صحيحاً أن الفرد الأكثر قدرة يحتمل بدرجة أكبر أن يجيب عن المفردة إجابة صحيحة من الفرد الأقل قدرة. لذلك يجب أن نتساءل عن سبب حدوث ذلك، أي نتساءل عما تقيسه هذه المفردة. وبالمثل إذا لم "يطابق fit" الفرد نموذج راش، فإنه لا يكون صحيحاً أيضاً أن احتمـــال إجابته إجابة صحيحة عن مفردة تتميز بالسهولة، أكبر من احتمـــال إجابته إجابة صحيحة عن مفردة تتميز بالصعوبـــة. وهنا يجب كذلك أن نتساءل عن سبب حدوث ذلك، أي عما حدث لهذا الفرد أثناء إجابته عن المفردة. إذ ربما لم تميز المفردة بين مستويات قدرة الأفراد، وبذلك لا تصلح لقياس هذه القدرة. (علام، 2001، ص.211) (علام، 2013، ص.58)

وعطفاً على ما سبق، يتضح أن النموذج الذي لا يساعد في التفسير السيكولوجي للسمات المراد قياسها لا يصلح للقيـــاس النفسي والتربوي. ولكن لا يجب أن ننكر التطبيقات العملية لنموذج راش التي تجعل منه

نموذجا رياضياً يشتمل على أقل عدد ممكن من البارامترات، و في الوقت نفسه له أساس سيكولوجي. وربما يكون هذا هو السبب الذي جعل علماء القياس النفســي والتربــوي المعاصر يهتمون بهذا النموذج اهتماماً خاصاً.

4-3- وحدات القيــاس المستخدمة في نموذج راش:

تتدرج صعوبة المفردات وقدرات الأفراد على ميزان مقياس واحد، وإن وحدة قياس كل من قدرة الفرد وصعوبة المفردة تقدر بوحدة قياس واحدة مشتقة مباشرة من نموذج راش، وهي وحدة اللوجيــت (Logit)، ويطلق عليه أيضاً الترجيح اللوغاريتمــي، ويعرف باللوغاريتم الطبيعي لمرجح نجاح الفرد على المفردات التي تعبر نقطة صفر التدريج عن صعوبتها عندما يساوي هذا المرجح مقداراً ثابتاً هو الأساس الطبيعي (e)، أي 2,718، ويكون عندئذ احتمــال نجاحه 0,73. (كاظم ، 1988أ، ص.57).

ويمكن تلافي التقديرات السالبة لكل من القدرة والصعوبة، وذلك بتعديل نقطة أصل التدريج على اعتبار أن صفر التدريج صفراً اعتبارياً، كما يمكن أيضاً تغيير حجم أو سعة وحدة القياس بحيث تتلافى التقديرات الكسرية لكل من القدرة و الصعوبة، وذلك على أساس أن الفرق بين القدرة والصعوبة هو الذي يحدد احتمال الاستجابة الصواب في ضوء نموذج راش، لذلك فإن اضافة أو طرح قيم ثابتة إلى قيم كل من القدرة والصعوبة لا يغير قيم هذا الاحتمــال. ومن ثم يمكن تحديد نقطة أصل تناسب هدف القياس. ويتم تحويل وحدة القياس باستخدام معادلة التحويل الخطــي رقم (14):

$$y = \alpha + \beta\chi \quad \text{.....(14)}$$

حيث أن: (y) هو التدريج الجديد؛ و(a) هي الإزاحة المطلوبة لنقطة أصل التدريج وتحدد نقطة الأصل في التدريج الجديد؛ و(β) هو عامل المسافة الذي يحدد سعة وحدة القياس الجديدة؛ و(χ) هو تدريج اللوجيت. (كاظم، 1988أ ، ص.101). وتتنوع تدريجات المفردات وتختلف تبعاً لأهداف القياس المختلفة. وفيما يلي بعض الأنواع المختلفة من التدريجات:

1. وحدات تدريج الجماعية (نيت).(Normative Scaling Units Nits)
2. وحدات التدريج المعتمد على محك مستقل (سيت) ((Substantive Scaling Units (sits).

3. وحدات التدريج الخاصة باحتمال الاستجابة الصواب (شيب) (Response Probability Scaling Units (chips)).

4. وحدة التدريج المستخدمة في المقاييس البريطانية (BAS) (ميمي السيد، 2007، ص.47).

5. وحدة المنف: إذ تحول تدريج اللوجيت إلى تدريج له خصائص تدريج الواط نفسها (تدريج مئوي نقطة أصل تساوي 50)، إلا أن طول أو سعة وحدة القياس(منف) يختلف عن سعة وحدة الواط. وتحصل على تدريج المنف (α=50) (β=5) في معادلة التحويل الخطي السابقة الذكر (كاظم، 1994، ص.126).

6. وحدة الواط (WATT): إذ يتم تحويل تدريج اللوجيت إلى تدريج يمتد من صفر إلى 100، ومتوسط صعوبة المفردات يساوي 50، ويمتاز هذا التدريج المئوي بسهولة التفسير لتقديرات قدرة الأفراد. (Masters, 1984, p.140).

5- تقدير بارامترات النماذج أحاديـــة البعد:

يعتمد احتمال استجابة فرد على مفردة اختبارية في نماذج الاستجابة للمفردة أحادية البعد على قدرة الفرد ($\varnothing$)، والبارامترات المتعلقة بالمفردة. ولكن هذه القيمة الاحتماليـــة وهذه البارامترات تكون غير معلومة، بينما يكون المعلوم لدينا إجابات الأفراد على المفردات، لذلك فإن أساليب تقدير البارامترات تهدف لتحديد قيمة ($\varnothing$) لكل فرد مختبر، وكذلك قيم بارامترات المفردة من إجابات الأفراد عليها.

5-1- تقديرات الأرجحية القصوى(Maximum Likelihood Estimation):

نالت تقديرات الأرجحية القصوى، اهتمام خبراء القيـــاس المعاصر، لذلك نجد العديد من الأساليب والمنهجيات التي ابتكرها هؤلاء العلماء لتيسير تقديرات بارامترات نماذج الاستجابة للمفردة. وتتباين هذه الأساليب في تحقيقها للمحكات التي ينبغي الاستناد إليها في تقييم مدى جودة أسلوب إحصائي استدلالي معين، بحيث يساعدنا في التوصل إلى قيم تقديرية للبارامتر في المجتمع المستهدف. (علام، 2013، ص.93)، ومن بين أهم المحكات، نجد: (أ). **الإتساق الداخلي** ويعني أن قيمة هذا الإحصاء تقترب تدريجياً من قيمة معلم المجتمع كلما زاد حجم العينة؛ **(ب). عدم التحيز** وهذا يعني أن القيم التقديرية المتوقعة لقياس إحصائي بعيد، أي قيمة متوسط جميع العينات العشوائية الممكنة ذات حجم معين، ينبغي أن تساوي قيمة معلم المجتمع؛ (ج). **الكفاية** وهي تعني أن توزيع العينـــة لا يعتمد على المعلم المراد تقديره، لذلك فإن معرفة الأسلوب

الإحصائي وتوزيع معايناته يكون كافياً لتزويدنا بهذه المعلومات؛ **(د). الفعالية النسبية** وتعني أنه إذا توافر مقياسان إحصائيان غير متحيزين لتقدير معلم المجتمع فإن أفضلهما هو الذي يكون أكثر فعالية بالنسبة للآخر، أي يكون الخطأ المعياري لتوزيع معايناته أقل. (ميمي السيد، 2014، ص.87)

5-2- أسلوب الأرجحية القصوى المشتركة((Joint Maximum Likelihood Procedure (JML):

نظرا لأن بارمترات الأفراد (∅)، وبارامترات المفردات تكون غير معلومة في معظم الأحيان، فإن تقديرها يمثل أحد المشكلات الرئيسية في نماذج الاستجابة للمفردة. وعند استخدام أسلوب الأرجحية القصوى المشتركة (JML) يمكن إجراء ذلك في خطوتين، كالتالي:

أولاً: يتم اختيار قيم مبدئية لبارامترات الأفراد، ويفضل أن تحسب هذه القيم المبدئية بقسمة لوغاريتم النسبة بين عدد الإجابات الصحيحة، وعدد الإجابات الخطأ لكل فرد مختبر. وتحول القيم الناتجة إلى قيم معيارية للتخلص من مشكلة عدم التحديد (Indeterminacy)، والتعامل مع قيم بارامترات الأفراد (بارامترات القدرة) على أنها معلومة.

ثانياً: يتم التعامل مع بارامترات المفردات على أنها معلومة، ويجري تقدير بارمترات الأفراد.

وتتكرر هاتان الخطوتان حتى نجد أن القيم التقديرية للبارامترات لا تتغير بعد الإنتهاء من الخطوتين في مرتين متتاليتين (Hambleton et al,1986). (علام، 2013، ص.96)

هذا الأسلوب الاحصائي في التقدير يعتمد عليه البرنامج الإحصائي وينستبس، والمصمم للتحليل وفقا لنموذج (راش) أحادي البارامتر.

5-3- أسلوب الأرجحية القصوى الهامشية((Marginal Maximum Likelihood (MML):

طور بوك و أيتكن (bock and Aitken,1981) هذا الأسلوب مع مستويات القدرة غير المعلومة، وذلك عن طريق تنفيذ مرحلتين للتقدير إحداهما تسمى مرحلة التوقع (Expectation)، والأخرى تسمى مرحلة التعظيم (Maximaization)، ففي مرحلة التوقع يتم حساب العدد المتوقع من الأفراد الذين يجيبون إجابة صحيحة على كل مفردة، واستخدام هذه التوقعات في معـــادلات التقدير. أما مرحلة التعظيم فالهدف منها الحصول على معالم مفردة معدلة استناداً إلى التوقعات المعدلة.

ويتميز هذا الأسلوب بمميزات متعددة، لعل أهمها أنه يمكن استخدامه في تقدير معالم جميع النماذج أحادية البعد، وكذلك النماذج متعددة الأبعاد، كما أن هذا الأسلوب يتميز بالفعالية سواء كان الاختبار عدد

مفرداته قليلاً أو كثيراً. وعلى الرغم من هذه المميزات، إلا أن هذا الأسلوب يشوبه أيضاً بعض العيوب من أهمها ما يلي: (أ). لا يؤدي في بعض الأحيان إلى نتائج مرضية حتى بعد عمليات التكرار المتعددة؛ (ب). يحتاج إلى توافر بيانات اختبارية لعينة كبيرة الحجم نسبياً من أجل الاقتراب من توزيع مستويات القدرة، والحصول على دالـــة الأرجحية الهامشية لمعالم المفردات.(ميمي السيد، 2014، ص.89)

5-4- أسلوب الأرجحية القصوى المشروطة((Conditional Maximum Likelihood (CML):

كان هذا الأسلوب حتى وقت قريب يمكن إجراؤه لتقدير بارامترات مفردات الاختبارات التي لا يزيد عددها عن 30 أو 40 مفردة (Wainer, et al, 1980). غير أن التطورات التالية جعلت من الممكن عملياً استخدامه إذا بلغ عدد المفردات 80 أو 100 مفردة. وترجع أهمية هذا الأسلوب إلى أن التقديرات الناتجة للبارامترات تكون متسقة. ولكن يقتصر استخدامه على النموذج أحادي البارامتر والنماذج المنبثقة عنه، حيث تشرط دالـــة الأرجحية بعدد الإجابات الصحيحة للأفراد على المفردات الاختبارية.

ومن الجدير بالذكر أن هذا الأسلوب لا يصلح في حالة النموذجين ثنائي، وثلاثي البارامتر، حيث تختلف قيم بارامترات تمييز المفردات. وبذلك تعتمد الدرجات الكلية للأفراد في الاختبار على بارامترات تمييز غير معلومة وغير منظورة. فهذه الدرجات الكلية لا تعد إحصاء كافياً لتقدير مستويات قدرة الأفراد، حيث أن هذه التقديرات سوف تعتمد على أى المفردات أجاب عليها الفرد إجابـــة صحيحة. ويتميز هذا الأسلوب بعدة مميزات من أهمها، ما يلي:(أ). لا يتطلب افتراض توزيع لمستويات القدرة كما في الأسلوب الثاني، فبارامترات القدرة غير المعلومة لا تؤخذ بعين الاعتبار في معادلات تقدير صعوبة المفردات؛ (ب). التقديرات الناتجة تتميز بخصائص مطلوبة في ظروف متعددة، مثل: الاتساق، والتوزيع الاعتدالي، والكفاية. غير أن هذا الأسلوب يشوبه أيضاً بعض العيوب من أهمها، مايلي: (أ). اقتصار استخدامه على النموذج أحادى البارامتر؛ (ب). يؤدي استخدامه إلى فقدان بعض المعلومات، حيث أنه لا يستطيع التوصل إلى تقديرات للمفردات والأفراد الذين أجابوا على جميع المفردات إجابة صحيحة، أو إجابة خطأ؛ (ج). توجد صعوبات في التحليل العددي عند استخدام اختبارات عدد مفرداتها كبير، مما يؤدي إلى فشل برنامج الحاسوب في استكمال إجراء التقديرات (Baker,1987 ; Embertson et al, 2000).

5-5- أساليب أخرى تعتمد على نظريـــة بييز (Bayes Theorem):

من الجدير بالذكر أنه توجد أساليب أخرى لتقدير بارامترات نماذج الاستجابة للمفردة، تعتمد على نظرية بييز التي تربط بين الاحتمالات الشرطيـــة (Conditional) ، والاحتمالات الهامشيـــة (Prior Probabilities). وتتطلب هذه الأساليب افتراض احتمالات قبلية للبارامترات في ضوء اعتبارات نظرية أو إمبيريقية، وكذلك إجراءات إحصائية متقدمة. (علام، 2013، ص.101)

5-3- الاختبارات التحصيلية المعتمدة على نظرية الاستجابة للمفردة:

وهي الاختبارات التحصيلية التي تحقق مطالب الموضوعية في القياس، وتتوافق درجات التلامذة عليها مع خصائص أسئلتها (مفرداتها) ، وذلك من خلال الاعتماد في بنائها على نظرية السمات (Trait Theory)، حيث تتمكن مثل هذه الاختبارات من قياس القدرات الفعلية للتلامذة.

تفترض نظرية السمات الكامنة، أنه يمكن التنبؤ بالأداء التحصيلي للتلميذ أو تفسير أدائه على الاختبار التحصيلي في ضوء خصائص مميزة لهذا الأداء تسمى السمات، وتحاول نظرية السمات الكامنة تقدير درجات الأفراد في هذه السمات من خلال أداء الأفراد على مجموعة من مفردات الاختبار. (أمين، 2009، ص.236).

6- صدق الاختبـــار وثباته في ضوء نموذج راش اللوغاريتمـــي الاحتمالـــي:

6-1- صـــدق الاختبار:

يعني صدق الاختبار في ضوء نموذج راش، بأن تعرف المفردات فيما بينها متغيراً واحداً أي أن مفردات المقياس تتدرج من حيث صعوبتها بحيث تعرف متغيراً واحداً، كما يعني تدرج قدرات الأفراد على المتغير محددة في تقديرات أدائهم على هذا المقياس وهذا يوضح أن كل من صعوبة المفردات و قدرات الأفراد تتدرج على متصل واحد ، لتمثل متغيراً واحداً (كاظم، 1988 أ ، ص.98).

6-2- ثبـــات الاختبار:

يعني ثبات الاختبار في ضوء نموذج راش بتحقيق مطالب الموضوعية في القياس، أي عندما تستخدم أداة للقياس تم بناؤها وفق هذا الأنموذج يجب أن يتحقق أن ما يأتي:

1. استقلال القياس عن المقياس المستخدم.

2. استقلال القياس عن مجموعة الأفراد التي طبق عليها المقياس.

أي أن استقلالية القياس وتحرره الذي يوفره أنموذج راش تتيح الفرصة لثبات القياس بحيث لا يختلف القياس سواء لقدرة الفرد أو صعوبة المفردة، باختلاف عينة التدريج أو باختلاف المقيـــاس المستخدم لقياس قدرة ما (الكحلوت، 1991، ص.131). كما أن نموذج راش لا يعطي فقط معامل ثبـــات للمقياس الكلي، ولكنه يعطي أيضاً معامل ثبات لكل من الفرد والمفردة (El- Korashy, 1995 , p.760).

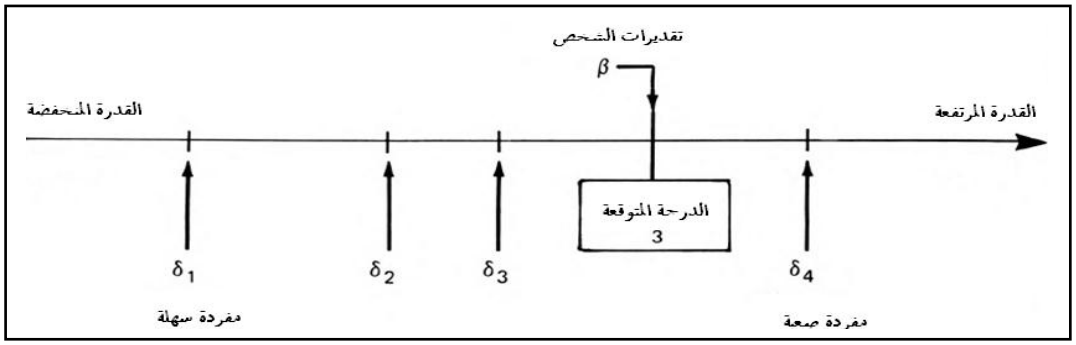
7- تدريـــج الاختبار وفق نموذج راش:

إن الحديث عن مفهوم تدريج الاختبار يتطلب الوعي بكل من مفهوم منحنى خصائص المفردة، ومنحنى خصائص الاختبار، واقتران المعلومات للمفردة والاختبار، وكذلك مواصفات متصل القدرة فمن المفترض مثلاً أن يؤخذ بعين الاعتبار أن متصل القدرة تتوسطه القيمة (صفر) كما أن وحدة القياس فيه هي الواحد الصحيح، ويمتد ضمن المدى النظري ($\pm \infty$)، كما أن قيم تقديرات معلمات المفردات والمفحوصين تقع على المتصل ضمن المدى المشار إليه(محاسنة، 2013، ص.271).

ويقصد بتدريج الاختبار، هو إنشاء ميزان لتدريج الصفة السلوكية الذي يحدد مستوى الأفراد على هذه الصفة مقدراً بوحدة معروفة، ويتمثل في هذا انتظام مواقع المفردات على متصل صعوبتها حول صفر تدريج بحيث يعبر موقع كل مفردة عن صعوبتها (ميمي السيد، 2007، ص.53).

وتتم عملية التدريج هذه من خلال تطبيق الاختبار على مجموعة من المفحوصين بحيث تأخذ إجاباتهم على مفردات الاختبار نمطاً ثنائياً (1- صفر)، وتجرى مجموعة من العمليات الرياضية للبيانات المستمدة من إجابات المفحوصين على مفردات الاختبار، وذلك من خلال البرامج الإحصائية بالحاسوب، للحصول على متصل القدرة وبشكل عملـــي، وموحد لكل من المفردات والمفحوصين، بعد ذلك يتم تقدير معلمات المفردات والمفحوصين على متصل القدرة. إذ إن تدريج الاختبـــار (قدرة، صعوبة، تمييز) يعد جزءاً مهم من إنجاز الاختبار ووضعه قيد الاستخدام، وهذا يدعم تفسير وتوضيح البناء النظري لنظرية القياس الحديثة. لذلك فإن النتيجة النهائية لعملية تدريج الاختبار تتمثل في تحديد متصل القدرة وفي ظل أنموذج راش، فإن وحدة القياس لمتصل السمة هي 1 وله نقطة منتصف تتمثل في النقطة صفر. وفي كل عملية تدريج فإن المفحوصين الذين يحصلون على نفس العلامة الخام، سوف يحصلون على نفس تقدير القدرة في ظل نموذج

راش لكن ذلك لا يحدث في ظل الأنموذجين ثنائي وثلاثي المعلمات (محاسنة، 2013، ص.281). كما يتضح في الشكل رقم (12):



شكل (12) يوضح تعريف متغير القياس من خلال عملية تدريج أو معايرة مفردات الاختبار.

(Benjamin and Wright, 1979, p.02**)**

8- ملاءمة البيانات للنموذج:

8-1- كيفية التحقق من حسن مطابقـــة البيانات الاختباريـــة الافتراضات نموذج راش:

سبق وأن أوضحنا أن نماذج الاستجابة للمفردة تستند إلى افتراضات معينة بشأن العلاقة بين الاستجابات للمفردات الاختبارية والسمة أو السمات الكامنة التي تنطوي عليها هذه الاستجابات. ومن بين أهم الافتراضات الرئيسة، هو: افتراض أحادية البعد، أي أن هناك سمة كامنة واحدة كافية لتفسير التباين المشترك بين الاستجابات للمفردات. كذلك افتراض الاستقلال الاحصائي الموضعي الذي يعني أن المفردات الاختبارية غير مرتبطة (مستقلة) عند ضبط مستوى أو مستويات قدرة الأفراد المختبرين. وبعبارة أخرى فإن احتمال الاستجابة على إحدى المفردات يتحدد فقط بمستوى قدرة الفرد المختبر، وليس باستجاباته لمفردات أخرى. وافتراض الاستقلال الموضعي يماثل افتراض "الأخطاء غير المرتبطة المشروطة بالدرجات الحقيقية للفرد المختبر" في اطار نماذج القيـــاس الكلاسيكية. وقد سبق أن أوضحنا كذلك أن افتراض البعد الأحـــادي، والاستقلال الموضعي مرتبطان، على الرغم مما يبدو أنهما متمايزان. فمجموعة من البيانات الاختبارية تكون أحادية البعد عندما تكون الاستجابات للمفردة مستقلة إحصائياً إستناداً إلى سمة كامنة واحدة.

وقد قام كل من هامبلتون وروجرز (Hambleton and Rogers,1988)، وهاتي (Hattie,1984-1985) كما جاء في أبو علام، باجراء مراجعة شاملة للمداخل والأساليب المستخدمة في تقييم أبعاد مصفوفة بيانات معينة، نذكر منها مايلي: (2013، ص.104)

1. التحقق من اتساق أنماط استجابات الأفراد للمفردات الاختبارية.

2. تحليل المكونات الأساسية لمصفوفة البيانات الاختبارية، وتحديد عدد قيم الجذور الكامنة الأكبر من الواحد الصحيح، أو عمل رسم بيانـــي لهذه القيم بعد ترتيبها من الأكبر إلى الأصغر، وتمثيل هذا الترتيب على المحور الأفقي، وقيم الجذور الكامنـــة على المحور الرأسي، وذلك لمصفوفة معاملات الارتباط الرباعي (Tetrachoric Correlation)، ودراسة هذه القيم لتحديد ما إذا كان هناك عامل أول أكثر تشبعا يفسر أكبر قدر من التباين المشترك في المصفوفة.

3. التحليل العاملي غير الخطي (Nonlinear Factor Analysis)، وهذا يتطلب إجراء مطابقة لنموذج التحليل العاملـــي غير الخطي لمصفوفة معاملات الارتباط الداخلي بين المفردات (Ineritem Correlation Matrix)، وفحص بواقي تحليل التغاير (Residual Covariance) بعد إجراء مطابقة لنموذج معين من نماذج التحليل العاملي غير الخطي.

4. التحقق من المفردات التي يبدو احتمال مخالفتها لافتراضات نموذج راش، وذلك لمعرفة ما اذا كانت هذه المفردات تسلك بطريقة مختلفة عن المفردات الأخرى في المجموعة.ويمكن إجراء ذلك بتعيير هذه المفردات كمجموعة فرعية، ومقارنة نتائج التعيير (بارامترات الصعوبة)، بنتائج تعيير المجموعة الكلية للمفردات.

2-8- أساليب تقييم حسن مطابقـــة البيانـــات لنموذج راش:

بعد التحقق من خاصيتي البعد الأحـــادي، والاستقلال الموضعي في مصفوفة البيانات الاختبارية، والتوصل إلى القيم التقديرية لبارامترات المفردات والأفراد باستخدام أحد النماذج المناسبة، ينبغي تقييم حسن مطابقة النموذج للبيانات. وقد أوضح كثير من خبراء القيـــاس المعاصر أنه لا يوجد أسلوب وحيد يمكن استخدامه في هذا الشأن بحيث يكون كافياً، وأكدوا ضرورة الحصول على أدلـــة متعددة للمراجعة الشاملة لمدى ملاءمة النماذج، والحكم على ملاءمة النموذج الذي يتم اختياره واستخدامه في تطبيق معين.

وللتحقق من مطابقة مجموعة المفردات التي تشكل الاختبار كوحدة واحدة لنموذج معين، وكذلك التحقق من مطابقة كل مفردة على حدة للنموذج، وبخاصة إذا كان الاختبار يشتمل على أنواع مختلفة من المفردات

التي تتطلب نماذج مختلفة لتقدير بارامتراتها. ونقصد بمطابقة المفردات لنموذج معين إمكانية النموذج تفسير كيفية استجابات الأفراد لإحدى المفردات، أو التنبؤ بهذه الاستجابات. (علام، 2013، ص.107).

وقد اقترح خبراء القيـــاس كثيراً من الأساليب أو المقـــاييس والمؤشرات لتقييم مطابقة المفردات لهذا النموذج ومن، أهمها:

8-2-1- أسلوب اختبار مربع كاي (Chi-Square χ^2):

يستخدم الأسلوب الإحصائي مربع كاي(χ^2)، لحساب مدى ملاءمة المفردة للنموذج المستخدم في كثير من البرامج الإحصائية على غرار برنامج (RUMM2020, BILOG-MG ,PARSCALE)، إلا أن لهذا الإحصائي بعض المحاذير عند استخدامه، حيث أنه شديد الحساسيـــة لحجم العينة، كما يجب ألا تقل عدد مفردات الاختبـــار عن 20 مفردة. ولهذا ينصح خبراء القيـــاس بعدم الإعتماد عليه كمحك وحيد لحذف المفردات، بل يجب أن يكون للباحث نظرة متفحصة بالإضافة إلى نتائج مربع كاي (Embretson & Reise, 2000, pp.230-238)، وينصح فيكميروفتش(Vukmirovic) ،أنه في حالة استخدام العينات الكبيرة يجب ألا يلتفت الباحث إلى دلالة (χ^2) نفسها، والاتساق بين القيم المختلفة في الاختبار، حيث تشير القيم الكبيرة الشاذة غالباً إلى وجود مشكلة ما بالمفردة وعلى الباحث أن يكون حذراً عند حذف المفردات. ومن الجدير بالذكر كذلك أن هامبلتون، وسواميناثان (Hambleton and Swaminathan,1985) قد أوضحا أن كثيراً من قيم مربع كاي الدالـــة إحصائياً تتأثر كثيراً بكبر حجم عينة الأفراد، مما يضفي بعض الشك على هذه القيم، وتفسيرها على أن البيانـــات لا تطابق النموذج، لذلك إقترحا إستخدام القيم النسبية لمربع كاي عند تقييم حسن مطابقة المفردات بدلاً من الاعتماد على الدلالـــة الإحصائية. كما اقترح كل من ميسلفاي، وبوك (Mislevy and Bock,1996)، استخدام نسبة أرجحية مربع كاي(Likelihood Ratio Chi-Square) في حالة الاختبارات التي تشتمل على أكثر من 20 مفردة. (علام، 2013، ص.108)

8-2-2- أسلوب حساب الملاءمـــة عن طريق مقياسى الملاءمة التقاربـــي والتباعـــدي:

يستخدم هذان الأسلوبان الإحصائيان في برنامج مايكروسكال وبرنامج وينستبس، وتحدد جودة المفردات بدرجة ملاءمتها للنموذج. وذلك من خلال نوعين من إحصاءات الملاءمة أولهما، هو إحصـــاء الملاءمة التقاربية(Infit)، وثانيهما إحصاء الملاءمة التباعدية (Outfit). وتعتمد الملاءمة التقاربيـــة على أداء الأفراد الذين تقترب قدراتهم من صعوبة مفردات المقياس، بينما تعتمد الملاءمة التباعديـــة على أداء الأفراد الذين

تبتعد قدراتهم (أي تزيد كثيراً أو تقل كثيراً) عن صعوبة المفردات. وتعتبر المفردة ضعيفة من حيث مستوى الملاءمة (Misfit) عندما تزيد قيمة إحصاءات الملاءمة الموجبة (سواء في إحصاءات الملاءمة المتقاربة أو المتباعدة) عن القيمة 2 أي تتجاوز حدود دلالة توزيع (ت) عند مستوى 0.05. وقد يشير ذلك إلى عيب في صياغتها أو عدم صدقها في قياس ما تقيسه باقي المفردات. أما إذا كانت إحصاءات الملاءمة السالبة تقل عن القيمة −2 فتعتبر متجاوزة لحدود الملاءمة (Overfit) وهذا يعني أنها شديدة التشابه أو الاعتماد على بعض مفردات المقياس، أي لا تكون مستقلة عنها وبالتالي يجب حذفها، لأنها لا تحقق افتراضات النموذج.

أما بخصوص اختبار مدى ملاءمة الأفراد للنموذج، فإذا كانت إحصاءات الملاءمة التقاربية (Infit) موجبة بحيث تتجاوز القيمة 2، فإن هذا يعني أن الأفراد يجيبون بصورة خطأ عن مفردات تقترب في مستوى صعوبتها من مستوى قدراتهم. وإذا كانت إحصاءات الملاءمة المتباعدة (Outfit) موجبة بحيث تتجاوز القيمة 2، فإن هذا يعني أما أن الأفراد يجيبون صواباً عن المفردات التي تزيد درجة صعوبتها عن مستوى قدراتهم أو أنهم يفشلون في الإجابة عن مفردات أقل من درجة صعوبتها من مستوى قدراتهم. أما إذا كانت الملاءمة تقل عن −2 فقد يعني هذا تشابه إجابات الأفراد. (وليد، 2014، ص.30)

واشار كل من رايت ولينكر (Wright and Linacre) كما ورد عن كاظم، إلى أنه ليس هناك تقدير مطلق في تقييم مدى الملاءمة، وبالتالي ينبغي أن يكون الحكم على الملاءمة متوافقا مع أهداف الموقف، واقترحا حدود الدلالة ± 2 حيث يقوم الباحث بحذف الأفراد غير الصادقين أولا ثم يعاد التحليل مرة أخرى. أما المفردات غير الملاءمة، فيقوم الباحث بفحصها وحذفها ثم إعادة التحليل بعد ذلك مرة ثالثة. (1994، ص.169)

والخلاصة أن تقييم حسن مطابقة البيانات الاختبارية لأحد نماذج الاستجابة للمفردة يتطلب التحقق من افتراضات النموذج المستخدم، وبخاصة افتراض البعد الأحادي، والاستقلال الموضعي، والتحقق من تنبؤات النموذج من خلال فحص البواقي المعيارية لمطابقة البيانات للنموذج، وتحديد طبيعة عدم المطابقة(Misfit) .

وعلى الرغم من تعدد أساليب هذا التقييم، إلا أنه يمكننا تصنيفها في ثلاث مجموعات و بخاصة فيما يتعلق بنموذج راش، وهي كالتالي:

1. قيمة مربع كاي (χ^2)، وذلك بمقارنة تكرارات الاستجابات الملاحظة بالاستجابات المتوقعة لمجوعات فرعية من الأفراد المختبرين المحددة مسبقا.

2. قيم المعيارية (Z) إستناداً إلى دالـــة الأرجحية للاستجابات أو لنمط المفردات.

3. قيم بواقي الدرجات(Score Residuals) استناداً إلى متوسطات انحرافات الدرجات الملاحظة عن الدرجات المتوقعة لاستجابـــات الأفراد المختبرين. (علام، 2013، ص.112)

9- دالـــــة المعلومـــات (Information Function):

يعد مفهوم دالـــة المعلومات من المفاهيم الأساسية في نظرية الاستجابة للمفردة الاختبارية سواء كانت المفردة ثنائية الدرجة، أو تتطلب استجابات متدرجة متعددة. وتساعد هذه الدالـــة في بناء أدوات قياس تجعل دقة القيـــاس أو المعلومات المطلوبة أكبر ما يمكن. وعلى الرغم من أن معظم تطبيقات دوال المعلومات اقتصرت على نماذج الاستجابة للمفردة ثنائية الدرجة (إجابة صحيحة أو خطأ)، إلا أن الباحثين بدأو في الآونة الأخيرة دراسة خصائص هذه الدوال في حالة استخدام نماذج الاستجابـــة متعددة الأقسام، ونماذج الاستجابة للمفردة متعددة الأبعاد. لذلك يعد مفهوم دالـــة المعلومات من المفاهيم السيكومترية التي تتميز بالعمومية، ويتعلق بتباين تقدير الأرجحية القصوى، فعندما نحصل على تقدير الأرجحية القصوى لبارامتر القدرة، فإن تباين هذا التقدير يكون مقلوب دالـــة المعلومات المناظرة. وبالمثل عندما نحصل على تقديرات الأرجحية القصوى لبارامترات المفردات، فإن مصفوفة التباين- التغاير (Variance-Covariance Matrix) للتقديرات تكون معكوساً لمصفوفة المعلومـــات (Information Matrix) المتعلقة بتقديـــرات بارامترات المفردات. ويرى بيرنبوم (Birnbaum,1968) أن مفهوم "المعلومات Information" بأنه مقدار يتناسب تناسباً عكسياً مع مربع سعة فترة الثقـــة حول تقدير قدرة فرد معين، والخطأ المعياري لهذا التقدير يساوي $1/\sqrt{(\text{المعلومات})}$. فعندما يكون مقدار المعلومـــات عند مستوى قدرة معين كبيراً تصبح فترة الثقة (Confidence Band) حول تقدير القدرة ضيقة، بينما تتسع هذه الفترة عندما يكون مقدار المعلومات عند هذا المستوى للقدرة أكثر إتساعاً. ونظراً لأن دالة المعلومات تتباين بتباين مستوى القدرة، فإن خبراء القياس اقترحوا أن منحنيات معلومات الاختبارات يمكن أن تكون بديلا لتقديرات الثبات المستخدمة في نماذج القياس

الكلاسيكيـــة، والأخطاء المعياريـــة للقيـــاس المستخدمة في تفسير درجات الاختبارات. (علام، 2013، ص.113)

9-1- دالـــة معلومـــات المفردات (Item Information Functions):

يمكن لأي منحنى مميز لمفردة اختبارية معينة (ICC) لأي من نماذج الاستجابة للمفردة ثنائية الدرجة (Dichotomous Item)، أو متعددة الأقسام (Polytomous Item)، إجراء تحويل عليه ليصبح منحنى معلومات المفردة ((Information Curve (IIC)، وهذا المنحنى يبين كمية المعلومات السيكومترية التي تشتمل عليها مفردة إختبارية عند جميع النقط على متصل السمة الكامنة. والصيغة الرياضية العامة لمنحنى معلومات المفردة (ICC) تعطى بالمعادلة رقم (15):

$$I(\theta) = \frac{[P\prime i(\theta)]^2}{[Pi(\theta)][1-Pi(\theta)]} \ldots\ldots\ldots\ldots\ldots\ldots (15)$$

حيث يشير الرمز $P'i(\theta)$ إلى المشتقة الأولى للدالـــة المميزة للمفردة عند مستوى قدرة أو سمة معين(θ) . وقيمة هذه المشتقة عند مستوى معين من مستويات القدرة (θ) يمثلها ميل المماس للمنحنى المميز للمفردة عند النقطة$P(\theta)$. والرمز $Pi(\theta)$ إلى الاحتمال المشروط لإجابة المفردة (i) إجابة صحيحة المستمد من المنحنى المميز للمفردة (ICC). (علام، 2013، ص.113)

يتضح للباحث بأن الدالـــة المعلوماتية للمفردة، هي مؤشر على دقة كل مفردة في تقدير قدرة التفكير الهندســـي لدى التلامذة، أي كلما تزداد معلومات المفردة تزداد الدقة في التقدير وينخفض الخطأ المعياري للقياس. لهذا فإذا كنا نستطيع أن نقدر البارامتر بدقة فهذا معناه أننا نملك معلومات أكثر عن هذا البارامتر. وقيمة الدالة المعلوماتية للمفردة مقدار غير ثابت يتوقف على مستويات القدرة التي نتعامل معها وفقاً لنظرية الاستجابـــة للمفردة وهي عكس النظرية الكلاسيكية للقيـــاس، والتي ترى بأن الثبات مقدار ثابت عند كل مستويات قدرات الأفراد. لهذا فإن الدالـــة المعلوماتية للمفردة مهمة جداً في بنوك الأسئلة، والاختبارات الموائمة المحوسبة ((Computerized Adaptive Testing (CAT)).

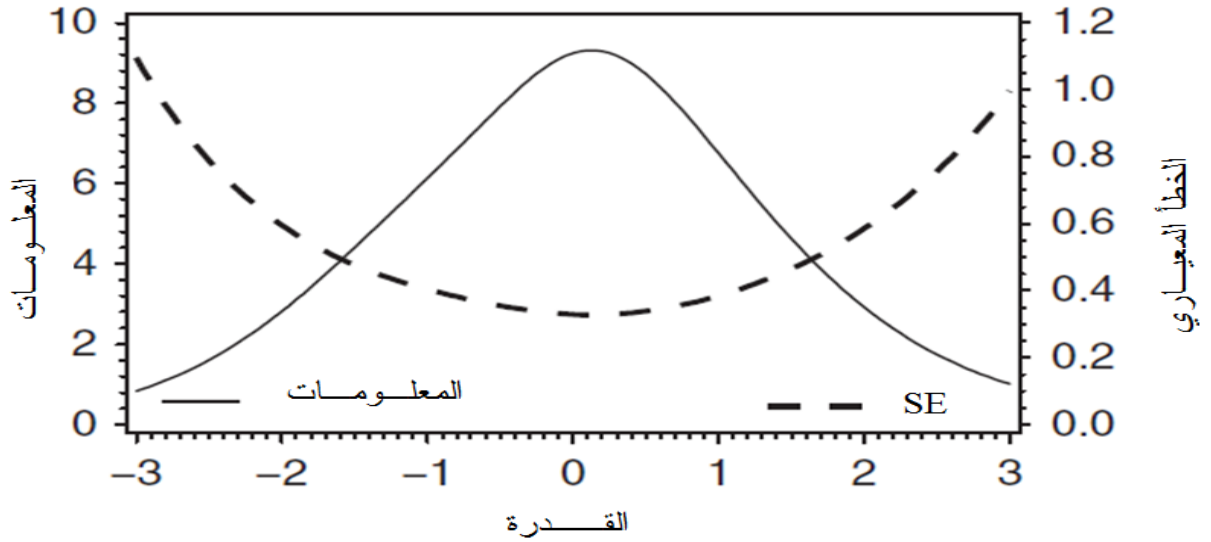
2-9- دالـــة معلومـــات الاختبار(Test Information Function):

يعرف لوجيمان (Logemanm,2006) دالـــة معلومـــات الاختبار، على أنها مجموع دوال معلومات المفردات التي تشكل الاختبار، وتعد دوال معلومات المفردات حجر الأساس لدوال معلومات الاختبار، والتي تحددها المعادلة رقم (16):

$$I(\theta) = \sum_{i=1}^{n} \frac{[P\prime i(\theta)]^2}{[Pi(\theta)][1-Pi(\theta)]} \ldots\ldots\ldots\ldots\ldots (16)$$

ويلاحظ أنه تم إضافة علامة (∑) إلى الطرف الأيمن من معادلة دالة معلومات المفردة، وذلك لأن دالة معلومات الاختبار تعتمد على تجميع إسهامات كل مفردة من مفردات الاختبار من المعلومات عند مستوى قدرة معين (θ). (ميمي السيد، 2014، ص.93) (علام، 2013، ص.117)

وتقدم دالـــة معلومات الاختبار تقديراً عن مدى جودة المفردات كلها في حدود دالة الاختبار على مقياس القدرة (θ)، وكلما زادت دالـــة المعلومـــات للاختبار عند درجة قدرة معينة (θ)، كان التمييز أكبر وخطأ القياس الذي يحتوى الاختبار أقل عند ذلك المستوى (θ). (ميمي السيد، 2014، ص.93). وفيما يلي مثال على دالـــة معلومـــات الاختبار في الشكل رقم (13):



شكل (13) العلاقة بين معلومات الاختبار والخطأ المعياري لتقدير الكفاءة (الموضع الذي يكون فيه أكبر معلومات يكون فيه الخطأ المعياري للقياس أقل). (Christine, 2010, p.07)

يرى الباحث بأن دالـــة المعلوماتية للاختبار، تعني مجموع الدوال المعلوماتية لمفرداته، لهذا فإن دالة معلومات المفردة هي مؤشر للدقة في تقدير قدرة التفكير الهندســـي لدى التلامذة، ونظراً للأهمية الكبيرة

للدالة المعلوماتية، إلا أن الباحث لاحظ غياب توظيف هذا المفهوم في أغلب الدراسات التي تم عرضها، والتي تناولت موضوع نظريـــة الاستجابة للمفردة، ما عدا دراسة (أيمن القهوجي، 2015) والتي اظهرت نتائجها، بأن متوسط القدرة يساوى صفراً عند أقصى دالة للمعلومات، فالاختبار يعطى أقصى دالة للمعلومات عند مستوى هذه القدرة. ويرجع الباحث هذا العزوف ربما إلى صعوبة الحصول على إحصاءات حول هذه الدوال بدون برامج إحصائيـــة متطورة سابقاً والتي تقوم بالتحليل وفق نظرية الاستجابة للمفردة، والتي اصبحت اليوم تمدنا باحصاءات للدوال المعلوماتية للمفردة، والاختبار ككل.

10- معادلـــة أو تكافؤ الاختبارات (Test Scores Equating):

تشير المعادلة إلى مجموعة من الأساليب في تعديل تقديرات مواقع أشخاص على مقاييس مختلفة لوضع تقديرات الأفراد على مقياس عام واحد. حيث قدم ليفنجستون (Livingston, 2004) تعريفاً بسيطاً وعاماً للمعادلة:

" تعتبر الدرجة على صورة جديدة والدرجة على صورة أخرى متكافئتين في مجموعة من الأفراد طبقت عليهم الصورة إذا كانوا يمثلون"الموقع النسبي" نفسه في المجموعة ". (دي إيالا، 2017، ص.375)

ويشير مبدأ ثبات المفردات لقدرة المفحوص إلى أنه ينبغي على المفحوص الحصول على نفس تقدير القدرة بغض النظر عن مجموعة المفردات المستخدمة. ويفترض مبدأ الثبات أن قيم معالم المفردة لمجموعات مختلفة من المفردات تعبر عن نفس قيـــاس مقياس القدرة.(بيكر، 2010، ص.142)

كما تشير أدبيات القياس المعاصر إلى تعريفات متعددة لعملية تكافؤ درجات الاختبارات. إذ يرى هولاند، وروبين (Holland and Rubin, 1982) أن عملية تكافؤ الاختبارات تعني إجراء تعديلات عددية على درجات كل صيغة من صيغ الاختبار للتعويض عن اختلاف الصعوبة من صيغة إلى أخرى. ويرى هيلز، وهيرش (Hills and Hirsh, 1988) أن التكافؤ (Equating) هو أسلوب لعزل التأثير على درجات الاختبار من الفروق الناجمة عن الاختلافات في متوسط مستويات صعوبة المفردات، ومتوسط مستويات تمييزها، وبين صيغتين أو أكثر من صيغ الاختبار التي تبنى من أجل أن تكون متكافئة. بينما يرى لورد وستوكنج (Lord and Stocking, 1990) أن تكافؤ درجات الاختبارات هو عملية إيجاد درجات متناظرة للصيغ المختلفة من الاختبار. كما يرى (Lord, 1989) أن التكافؤ يتحقق إذا لم يكن هناك اختلاف في تقدير مستوى قدرة فرد، سواء على الاختبار (س)، أو الاختبار(ع).(علام، 2013، ص.213)

10-1- أنواع معادلـــــة الاختبارات:

يوجد نوعان رئيسان من أنواع تكافؤ الاختبارات، هما:

10-1-1- التعــادل الأفقـــي:

وفيه تكون الصيغ المختلفة للاختبار متماثلة في المحتوى، ومتقاربة في مستوى صعوبة مفرداتها، وتقيس السمة نفسها في مجتمع معين من الأفراد. والهدف من إجراء عملية المعادلة أو التكافؤ، هو تعديل الفروق في صعوبة المفردات، أو الفروق في توزيعات درجات الأفراد المختبرين. ويكون هذا النوع ضرورياً عندما تكون هناك حاجة لتوافر صيغة متعددة من اختبار معين.

10-1-2- التعــادل الرأســـي:

وفيه تكون صيغ الاختبار مختلفة في مستوى صعوبتها ولكن تقيس نفس النطاق أو المحتوى، وكذلك تكون عينات الأفراد التي تطبق عليها هذه الصيغ مختلفة في مستوى قدرتها، والهدف من إجراء عملية التكافؤ الرأســـي هو تحويل درجات صيغ الاختبارات المصممة لمجتمعات من الأفراد في مستويات تعليمية مختلفة إلى ميزان واحد. والتكافؤ الرأســـي يكون مطلوباً عندما نود إجراء مقارنات بين الأفراد المختبرين من ذوي مستويات متباينة.

10-2- متطلبات إجراء معادلـــة الاختبارات:

يتطلب إجراء معادلة الاختبارات، أن تقيس الصيغ المتعددة من اختبار معين نفس القدرة أو السمة أو المهارة بدرجة كافية، ولذلك ينبغي أن تبنى الاختبارات المراد إجراء تكافؤها وفقاً لنفس المواصفات، وأن تكون متقاربة أو متماثلة في صعوبتها، وثبات درجاتها. وقد حاول لورد (Lord, 1980) تحديد عددٍ من المتطلبات التي ينبغي تحققها لإجراء تكافؤ الاختبارات سواء باستخدام أساليب نماذج الاستجابة للمفردة أو أساليب نظرية القياس الكلاسيكية، وهذه المتطلبات، هي: (أ) قياس نفس الخاصة؛ (ب) التعادل؛ (ج) استقلالية دالة التحويل عن البيانات المستمدة من مجتمع معين (Invariance)؛ (د) التماثل.

والخلاصة أنه يمكن النظر إلى عملية المعادلة في حد ذاتها، بأنها مؤلفة من عدة مراحل، وهي مرحلة جمع بيانات تتبعها، مرحلة التحويل بعد التوصل إلى مطابقة بين النموذج والبيانات وتقديرات لمعالم المفردات من أجل تطبيق وربط الصور الاختبارية في تقدير مواقع الأفراد.

11- الأداء التفاضلـــي (التمايزي) للمفردة الاختبارية (Differential Item Functioning):

إن كون مؤشري الصعوبـــة والتمييز لاختبار ما مقبولين، لا يعني أن هذا الاختبار يعمل بشكل نفسه مع كافة المجموعات البشرية (من عرق أو جنس أو عمر معين) التي يتم اختبارها. ولهذا السبب يتطلب وضع الاختبـــارات التحصيلية القياسيـــة وغيرها من اختبارات القدرة الحصول على معلومات تحيز (Biasedness) مفردات الاختبار، أو ما يسمى بفعل المفردة التباينـــي (Differential item functioning (DIF)). (لويس، 2007، ص.94)

نال تحليل "تحيز المفردات" اهتماماً متزايداً كأحد مظاهر بحث تحيز الاختبـــارات ضد مجموعات الأقليات، ومثل ذلك التحليل يتعلق بدرجة أساسية "بالصعوبة النسبية" لكل مفردة اختبارية فيما يتعلق بالمجموعات ذوي خلفيات ثقافية وخبرية غير متشابهة، وهذا المجال لتحليل المفردات يطلق عليه "الأداء التمايزي للمفردات (DIF)"، وهو يسعى إلى تحديد المفـــردات التي تكون هناك احتمالات مختلفة للنجاح للأشخاص من مجموعات ثقافية مختلفة وقدرتهم متساوية. ولهذا الغرض، فإن القدرة المتساوية تعني التساوي فيما يتعلق بالتكوين الفرضـــي الذي يصمم الاختبار لتقييمه، أو السلوك المحك الذي يصمم للتنبؤ به. وأحد الأسباب المحتملة لشذوذ المفردة، هو أنها لا تقيس نفس التكوين الفرضـــي في مجموعات مختلفة. فمثلاً، ربما يقيس مصطلح "متشابهات analogies" الاستدلال اللفظي في مجموعة معينة، ولكنه يقيس معرفة المفردات اللغوية في مجموعة أخرى، وذلك لأنه يحتوي على كلمة مفتاحية غير مألوفة لكثير من أعضاء أقلية معينة. وبالمثل، ربما تقيس مسألة حسابية القدرة الرياضية في مجموعة معينة، ولكنها تقيس القدرة على فهم عبارات لفظية معقدة في مجموعة أخرى. وفي هذين المثالين، كان الافتقار إلى المعرفة موائما للتكوين الفرضي الذي يقيسه الاختبار ككل. غير إننا إذا افترضنا أن المفردات الرياضية التي تتضمن كسوراً عشرية، تكون أكثر صعوبة نسبياً لأعضاء مجموعة معينة، فإن هذا الفريق يكون موائماً للتكوين الفرضـــي المتعلق بالقدرة الرياضية. ولذلك ، فإن مثل هذا الشذوذ – بهذا المعنى – لا يدل على أن المفردات متحيزة. وقد تم تطوير العديد من الطرق لتحديد مثل تلك المفردات ذات الأداء التمايزي، وتشمل كلاً من الإجراءات التي تعتمد على الأحكام، وكذلك الإجراءات الإحصائية .(أناستازي، 2015، ص.251-253)

وأحد تفسيرات سبب ظهور الأداء التفاضلي للمفردة يشير إلى تعددية البعدية، أي يمكن تصور الأداء التفاضلي شكلاً من أشكال تعددية البعدية يحدث عندما تقيس المفردة عدة أبعاد، وعندما تختلف المجموعات الظاهرة عيانياً في مواقعها بالنسبة لبعضها بعضاً على المتغير/ أو المتغيرات الثانوية الكامنة. ومن منظور نظرية الاستجابة للمفردة، وجود الأداء التفاضلـــي في مفردة يعني أن تقديرات معالم تلك المفردة لا توصف باللاتغاير في المجموعات الظاهرة (أي مطابقة في بيانات المفردة)، أي أن هذا الافتقار إلى "اللاتغاير" الذي يفسر كدليل على وجود الأداء التفاضلـــي للمفردة.(دي إيالا، 2017، ص.398-399)

11-1- **الإجـــراءات التي تعتمد على الأحكـــام:**

لا توجد طريقة واحدة تعد الأفضل لتحليل تحيز المفردات تصلح لجميع الأغراض. فنظراً لأن الطرق المختلفة تقدم أنواعا مختلفة إلى حد ما من المعلومات، فمن المرغوب فيه استخدام مزيج من الطرق. والمزيج المناسب يعتمد على الاستخدامات المتوقعة للاختبار وعلى نوع الاستدلالات المراد التوصل إليها من الدرجات. وأفضل مزيج بعامة يشمل بعض الإجراءات الإحصائية، والإجراءات التي تعتمد على الأحكام – إذا طبقت تطبيقاً مناسباً – يمكن أن تقدم معلومات مفيدة يصعب الحصول عليها بإجراء أخر والتحليل الذي يعتمد على الأحكام يكون مناسباً بخاصة في المراحل الأولى والأخيرة في بناء الاختبارات، وذلك لإجراء ومتابعة التحليلات الإحصائية. وهو يبدأ عادة مبكراً في بناء اختبار معين لفحص محتواه الذي يحتمل أن يكون متحيزاً أو يقلل من شأن الأقليات أو يكون مؤيداً للأنماط الثقافية فيما يتعلق بالأدوار المهنية أو الاجتماعية. ولهذا الغرض، يستخدم الناشرون الرئيسيون للاختبارات عادة مراجعين للمفردات من فريق الخبراء لديهم، ومن ممثلين لمستشارين خارجيين. وتساعد مثل تلك المراجعات أيضا في تحديد محتوى الاختبار الذي ربما يقتصر على ثقافة معينة، وبذلك يكون غير مألوف بالنسبة لمجتمعات معينة من المختبرين. غير أنه ينبغي ملاحظة أن المراجعات التي تعتمد على الأحكام لم تبين بعامه نجاحها في التنبؤ بالصعوبة أو القوة التمييزية للمفردات بالنسبة لمجتمعات مختلفة. والتحليلات الإحصائية والنتائج الإمبريقية تكونان مطلوبتين لهذا الغرض. (أناستازي، 2015، ص.253)

11-2- الإجـــراءات الإحصائيـــة:

11-2-1- طريقـــة تمييز المفـــردة:

تستخدم هذه الطريقة معامل الارتباط الثنائي النقطي (Point biserial correlation coefficient,r_{pbis}) كمؤشر على قدرة المفردة التمييزية. وهو معامل ارتباط المفردة مع الدرجة الكلية للاختبار، ويستخدم كمؤشر دال على العلاقة بين متغيرين أحدهما ثنائي (طبيعي أو محول) والآخر متصل، حيث أن المفردات ذات الارتباط الضعيف مع الاختبار لا تساهم في اضافة تباين حقيقي إلى تباين الاختبار (jensen, 1980).

ويرى ثورندايك (Thorndike, 1982) أن اختلاف معامل الارتباط الثنائي النقطي (r_{pbis}) بين المجموعتين قد يفسر على أن المفردة تختلف في فعاليتها لقياس السمة التي يهدف الاختبار لقياسها في المجموعتين ويحسب مؤشر تمييز المفردة لمجموعة من الأفراد وفق المعادلة رقم (17) (Croker and Algina,1986):

$$r_{pbis} = \frac{u_+ - u_x}{\sigma_x} \cdot \sqrt{\frac{p_1}{1 - p_1}} \quad \ldots\ldots\ldots\ldots\ldots\ldots (17)$$

حيث أن: (u_+) الوسط الحسابي لعلامات الأفراد الذين أجابوا على المفردة إجابة صحيحة؛ (u_x) الوسط الحسابي لعلامات جميع الأفراد الذين استجابوا على المفردة؛ (σ_x) الانحراف المعياري لعلامات المجموعة.(p_1) صعوبة المفردة.

وبعد إيجاد معاملات ارتباط المفردة مع العلامة الكلية، يتم ترتيب معاملات الارتباط تصاعدياً أو تنازلياً لكل مجموعة، والمفردة التي يكون معامل ارتباطها مع الاختبار في النصف العلوي للمجموعة الأولى وفي النصف السفلي للمجموعة الثانية تصنف على أنها مفردة متحيزة، أي أنه إذا اختلف ترتيب موقع معامل الارتباط للمفردة في المجموعتين بصورة واضحة، فإن ذلك يدل على تحيزها.

11-2-2- طريقة كاي تربيع (χ^2 The Chi-Square Test):

ينظر للمفردة غير المتحيزة وفق هذه الطريقة، على أنها المفردة التي يكون لجميع الأفراد في مستوى القدرة الواحدة نفس احتمالية الإجابة الصحيحة عليها بغض النظر عن المجموعة الفرعية التي ينتمون إليها. وتقوم على عدد من افتراضات النظرية الحديثة في القيـــاس مثل: أحادية البعد، والاستقلال الموضعي

للمفردات. وتتميز طريقة كاي تربيع عن طريقة فرق المساحة بين منحنيات خصائص المفردة أن القدرة تتوزع فيها إلى فئات وليست نقاط، وهي أسهل حسابياً، ولا تحتاج لعينة كبيرة من المفحوصين.

ويمكن تلخيص إجراءات تطبيق هذه الطريقة حسب الخطوات التالية:

1. توزيع علامات المفحوصين على الاختبار في مستويات قدرة مختلفة بوضعها في فئات.
2. يتم تحديد عدد الأفراد ضمن كل فئة ولكل مجموعة من المجموعات.
3. إيجاد عدد الأفراد الذين أجابوا على المفردة إجابة صحيحة في كل فئة ولكل مجموعة من المجموعات.
4. ثم تحسب القيم المتوقعة لكل مجموعة ومن ثم حساب قيمة كاي تربيع (χ^2) والتي تمثل إحصائي شوينمان (Scheuneman) من خلال القانون رقم (18):

$$\chi_s^2 = \sum_{j=1}^{j} \frac{(O_{1J} - P_J N_{1J})^2}{P_J . N_{1J}} + \sum_{j=1}^{j} \frac{(O_{2J} - P_J N_{2J})^2}{P_J . N_{2J}} \quad \text{.....(18)}$$

حيث أن: (O_{1J}) و (O_{2J}) عدد المفحوصين الذين أجابوا إجابة صحيحة على المفردة من المجموعتين الأولى والثانية ممن هم ضمن المستوىJ؛ و(N_{1J}) عدد مفحوصي المجموعة الأولى الذين لهم علامات في المستوىJ؛ و(N_{2J}) عدد مفحوصي المجموعة الثانية الذين لهم علامات في المستوىJ؛ و(P_J) نسبة المفحوصين من المجموعتين الذين أجابوا بشكل صحيح على المفردة ممن هم ضمن المستوىJ؛ و($P_J N_{1J}$) و ($P_J N_{2J}$) القيم المتوقعة.

وللكشف عن التحيز يتم مقارنة قيم (χ^2) المحسوبة مع قيمة (χ^2) الحرجة عند درجة حرية (J-1)(K-1) حيث أن:(K) تمثل عدد المجموعات؛ و(J) تمثل عدد مستويات القدرة. فإذا كانت قيمة(χ^2) دالة إحصائياً، دل ذلك على تحيز المفردة، وبإزدياد قيمة كاي تربيع بعد أن تكون دالـــة إحصائياً تكون مؤشراً أكبر على تحيز المفردة.

11-2-3- طريقة مانتل- هانزل (Mantel-Hanzel Method,MH):

تعد طريقة مانتل – هانزل من أكثر طرق النظرية الكلاسيكية انتشاراً في الكشف عن تحيز المفردة، حيث تقوم هذه الطريقة بإستقصاء التحيز بين مجموعتين فرعيتين من مجتمع الدراسة، بشرط أن تتكون العينة على الأقل من (200) شخص في كل مجموعة، إحداهما تسمى المجموعة المرجعية (Reference

Group)، وتسمى الأخرى المجموعة المستهدفة (Focal Group) وهي المجموعة التي يعتقد بوجود مفردات تتحيز ضدها، والفرضية الصفرية (H_0) هنا تنص بعدم وجود فروق في نسب الإجابة الصحيحة بين المجموعتين المرجعية والمستهدفة في مستوى القدرة الواحدة (t). ولتطبيق هذه الطريقة تنظم البيانات في جدول توافق (Contingency Table) ثنائي البعد (2*2) كما هو مبين في الجدول رقم (02):

جدول (02)

يوضح توافق ثنائي البعد (2*2) **لنوع المجموعة (مرجعية ومستهدفة) وطبيعة الإجابة (صح وخطأ).**

	إجابة صحيحة (1)	إجابة خاطئة (0)	المجموع
المجموعة المرجعية (R)	A_t	B_t	NR_t
المجموعة المستهدفة (F)	C_t	D_t	NF_t
المجموع	$N1_t$	$N0_t$	N_t

حيث إن: (A_t) عدد المفحوصين من المجموعة المرجعية الذين أجابوا على المفردة إجابة صحيحة في مستوى القدرة (t)؛ و(B_t) عدد المفحوصين من المجموعة المرجعية الذين لم يستطيعوا الإجابة على المفردة إجابة صحيحة في مستوى القدرة (t)؛ و(C_t) عدد المفحوصين من المجموعة المستهدفة الذين أجابوا على المفردة إجابة صحيحة في مستوى القدرة (t)؛ و(D_t) عدد المفحوصين من المجموعة المستهدفة الذين لم يستطيعوا الإجابة على المفردة إجابة صحيحة في مستوى القدرة (t)؛ و(NR_t) مجموع المفحوصين الذين أجابوا على المفردة من المجموعة المرجعية ممن هم في نفس مستوى القدرة (t)؛ و(NF_t) مجموع المفحوصين الذين أجابوا على المفردة من المجموعة المستهدفة ممن هم في نفس مستوى القدرة(t) ؛ و($N1_t$) مجموع المفحوصين الذين استطاعوا الإجابة من كلا المجموعتين في مستوى القدرة (t)؛ و($N0_t$) مجموع المفحوصين الذين لم يستطيعوا الإجابة من كلا المجموعتين في مستوى القدرة (t)؛ و(N_t) عدد المفحوصين في مستوى القدرة (t). واعتماداً على الجدول يتم تطبيق معادلة مانتل– هانزل كاي تربيع رقم (19):

$$MH_{\chi^2} = \frac{\left[|\sum A_t - \sum E\,(A_t)| - \frac{1}{2}\right]^2}{\sum var\,(A_t)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(19)$$

إذ تمثل $E(A_t)$ القيمة المتوقعة لعدد المفحوصين الذين أجابوا على المفردة إجابة صحيحة من المجموعة المرجعية في مستوى القدرة (t).

$$E(A_t) = \frac{(\mathrm{NRt.\ N1t})}{\mathrm{Nt}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (20)$$

والقيمة ($\frac{1}{2}$) في المعادلة رقم (19)، هي معامل تصحيح "ييتس" (Yates) للمتصل. أما $\mathrm{var}(A_t)$ فتمثل تباين (A_t) ، وكلاهما يمكن حسابه من جدول التوافق، ويتبع الإحصائي ($\mathrm{MH}\chi^2$) توزيع (χ^2) بدرجات حرية تساوي 1، حيث يحسب التباين بالمعادلة رقم (21) التالية: (Abedalaziz, 2010, p.106)

$$var(A_t) = \frac{NR_t NF_t N1_t N0_t}{N_t^2 (N_t - 1)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (21)$$

ويتم تقويم إحصائي مانتل–هانزل ($\mathrm{MH}\chi^2$) مقابل القيم الحرجة (الجدولية) لـــ (χ^2)بدرجات حرية تساوي 1. وتوجد طريقة أخرى في تفسير الفرضية الصفرية أن احتمالات أعضاء المجموعة المرجعية للاستجابة 1 هي نفسها لأعضاء المجموعة المستهدفة. ويشار إلى هذه بفرضية نسبة الأرجحية الثابتة أو المشتركة، ويستخدم مؤشر نسبة الأرجحية المشتركة (Combined Odds Ratio, αMH) للدلالة على إتجاه التحيز، والذي يحسب وفق المعادلة رقم (22):

$$\alpha\mathrm{MH} = \frac{\sum \frac{A_t D_t}{N_t}}{\sum \frac{B_t C_t}{N_t}} \quad \ldots\ldots\ldots\ldots\ldots\ldots (22)$$

ويتم تحديد الأداء التفاضلي من عدمه، من خلال قيمة نسبة الأرجحية المشتركة (αMH) فإذا كانت ($\alpha\mathrm{MH} = 1$) فالمفردة لا تظهر أداءً تفاضلياً، وإذا كانت ($\alpha\mathrm{MH} < 1$) فالمفردة تظهر أداءً تفاضلياً لصالح المجموعة المستهدفة، وإذا كانت ($\alpha\mathrm{MH} > 1$) فالمفردة تظهر أداءً تفاضلياً لصالح المجموعة المرجعية (kanata and Vaughn,2004). وتجدر الإشارة إلى أن طريقة مانتل – هانزل، تعطي تقريباً مناسباً لطريقة منحنى خصائص المفردة في الكشف عن الأداء التفاضلـــي، وذلك من خلال النتائج التي توصلت إليها كل من الطريقتين، بالرغم من إحصائيات طريقة مانتل–هانزل مختلفة جداً عن طريقة منحنى خصائص المفردة. (دي إيالا، 2017، ص.401)

11-2-4- فرق المساحـــة بين منحينـــات خصائص المفردة:

نظراً للاستخدام المتزايد لنظرية استجابة المفردة، انبثقت عنها طريقة منحنى خصائص المفردة (ICC) كأبرز تطبيق في الكشف عن الأداء التفاضلـــي للمفردة (DIF) (Hambleton and Swaminathan,1985)، ومنحنى خصائص المفردة هو دالـــة تربط بين احتمال إجابة المفحوص على المفردة $Pi\,(\theta)$ من جهة، وخصائص المفردة وقدرة المفحوص (θ) التي يقيسها الاختبار الذي يحتوي على تلك المفردة من جهة أخرى. ويعتبر المعلم (b) مؤشراً لتباين الأداء التفاضلـــي، والذي هو ببساطة الفرق بين معلمتي صعوبة المفردة في المجموعتين (منحنى خصائص المفردة في المجموعتين) (Linacre and Wright, 1987 ; Draba,1977)، ويستخدم مؤشراً الفرق بين صعوبتي المفردة في المجموعتين على نطاق واسع بسبب بساطته، كما إنه مؤشر جيد لحجم التأثر، ومع ذلك، فإنه ليس مرغوب فيه عندما يكون الأداء التفاضلـــي من النوع غير المنتظم والناجم عن الاختلاف في معلمتي التمييز والتخمين للمفردة.(yoke, 2010, p.60)

كما قدم راجو (Raju, 1988) طريقة أخرى يمكن من خلالها حساب المساحة بين منحنيات خصائص المفردة ذات المعلمة الواحدة أو المعلمتين أو الثلاث معالم. حيث يعبر عن المساحة في النموذج اللوجيستي ثلاثي المعلم بالمعادلة رقم (23):

$$Area = (1-c).\left|\frac{2(a_2-a_1)}{Da_1.a_2}\,In\left(1+exp\left(\frac{Da_1.a_2(b_2-b_1)}{a_2-a_1}\right)\right)-(b_2-b_1)\right| \quad \ldots\ldots\ldots\ldots\ldots\ldots (23)$$

حيث أن: (a_1) تمييز المفردة للمجموعة الأولى؛ و (a_2) تمييز المفردة للمجموعة الثانية؛ و(b_1) صعوبة المفردة للمجموعة الأولى؛ و (b_2) صعوبة المفردة للمجموعة الثانية؛ و (c) قيمة التخمين؛ و(D) Scaling Factor قيمة ثابتة وتساوي 1.7. وإذا طبقنا نموذج المعلمتين فإنه لا وجود لمعلم التخمين (c) وعليه تقل متغيرات المعادلة رقم (23)، لتصبح معادلة رقم (24):

$$Area = \left|\frac{2(a_2-a_1)}{Da_1.a_2}\,In\left(1+exp\left(\frac{Da_1.a_2(b_2-b_1)}{a_2-a_1}\right)\right)-(b_2-b_1)\right| \ldots\ldots\ldots\ldots (24)$$

وعند استخدام نموذج المعلمة الواحدة فإن المساحة تكون القيمة المطلقة للفرق بين معلمتي الصعوبة للمجموعتين (Hambleten et al,1991). وحسب (راجو) أن المساحة بين المنحنيين في النموذج ثلاثي المعلم قد تصل إلى ما لانهاية أو أنها خارج مدى القدرة، لذلك إقترح تساوي معلمة التخمين في المجموعتين، واشتق تعبيراً رياضياً تظهر فيه الأخطاء المعيارية للمساحة التي تتوزع توزيعاً طبيعياً، وباعتبار معلمة التخمين ثابتاً، وفي حال عدم تساوي معلمة التخمين للمجموعتين، فإنه عادة ما يتم تثبيت هذه المعلمة عند (C=0.2) لحساب فرق المساحة بين منحنيات خصائص المفردة، وإيجاد نقطة قطع للمساحة الإحصائية التي تعتبر معياراً للحكم على وجود الأداء التفاضلي للمفردة، أو عدم وجوده، ويتم حساب القيمة الحدية، كالتالي:

1. تقسيم المجموعة المرجعية إلى مجموعتين عشوائيتين متكافئتين.
2. يتم تقدير منحنيات خصائص المفردة لكل مجموعة لوحدها.
3. يمكن تكرار عملية اختيار المجموعة وتقسيمها لمجموعتين عشوائيتين متكافئتين لعدد من المرات، وفي كل مرة يتم تقدير معالم المفردة لكل مجموعة منفصلة.
4. يتم حساب المساحة بين تلك المنحنيات في المجموعات ويتم إعتماد أكبر قيمة للمساحة بين المنحنيين في المجموعات المختارة، كنقطة قطع تقارن بها المساحة الفعلية.
5. **القرار:** إذا كانت المساحة الفعلية بين المنحيين أكبر من القيمة العظمى لفرق المساحات بين المنحنيين في المجموعات المختارة عشوائياً (نقطة القطع)، فإن ذلك يدل على الأداء التفاضلي للمفردة، وإذا كانت أقل من قيمة نقطة القطع، فإن ذلك يدل على عدم وجود أداء تفاضلـــي للمفردة. (حمادنة، 2007، ص.24)

11-2-5- طريقة اختبار "ت" راش- ويلش (Rasch-Welch t test):

وهي طريقة قام بها ويلش (welch) لتحسين اختبار "ت" ستودنت (Student's *t*-test)، ويعرف باختبار (Aspin-Welch Test) أو (Welch's t-test) أو (Satterthwaite method) ويلش – ساترثويت، هو طريقة إحصائية تستخدم لاختبار الفرضية الصفرية فيما يتعلق بالفروق بين عينتين، عندما يكون حجم العينتين، والانحراف المعياري، والتباين غير متساوي. وهو حل مقترح لمشكلة (بهرنز- فيشر Behrens-Fisher). كما يشترط التوزيع الطبيعي للعينة، أي إذا ثبت بأن تجانس التباين غير محقق لدى العينتين المستقلتين، فإنه يتم اللجوء إلى الاختبـــار الذي يعتمد على تعديل عدد درجات الحرية، بما يتوافق مع تباين العينات غير المتساوي، ويعتبر اختبار "ت" راش ويلش أشهر هذه الاختبارات. (Sheldem , 2014, p.14)

يتم احتساب الأداء التفاضلـــي للمفردة باستخدام اختبار "ت" راش– ويلش من خلال برنامج وينستبس، ويتم من خلالها طرح صعوبة المفردة في المجموعتين (المرجعية و المستهدفة)، ليتم تحويل التباين لتوزيع طبيعي معياري، باستخدام الخطأ المعياري للقياس، كما توضحه صيغة المعادلة رقم (25):

$$t_{DIF} = \frac{b_{ref} - b_{foc}}{\sqrt{SE^2(b_{ref}) + SE^2(b_{foc})}} \quad \text{.....(25)}$$

حيث (b_{ref}) صعوبة المجموعة المرجعية؛ و(b_{foc}) صعوبة المجموعة المستهدفة؛ (SE^2) الخطأ المعياري للقياس.(Cauffman, 2006, pp.502-521)

ومن خلال المعادلة رقم (26)، يتم استخراج درجة حرية معدلـــة وهي غير درجة الحرية في الحالات الطبيعية وباستخراجها يتم استخراج القيمة الجدولية ليتم مقارنتها بالقيمة المحسوبة لـــ (ت) المستخرجة من المعادلة (25). وتعطى درجة الحرية لـــ"ويلش–ساتيرثويت" (Welch-Satterthwaite) بالمعادلة رقم (26):

$$d.f=(EV1+EV2)^2 /(EV1^2/(N1-1) +EV2^2/(N2-1)) \text{.....(26)}$$

حيث تمثل (EV1) الخطأ المعياري لتباين المجموعة الأولى، (EV2) الخطأ المعياري لتباين المجموعة الثانية، (N1) عدد العينة في المجموعة المرجعية، (N2) عدد العينة في المجموعة المستهدفة. وبطريقة أخرى يمكن أن يعطى بالمعادلة رقم(27):

$$d.f= (SE1^2+SE2^2)^2 / (SE14/(N1-1) +SE24/(N2-1)) \text{.....(27)}$$

حيث تمثل (SE1) الخطأ المعياري لمتوسط المجموعة الأولى، (SE2) الخطأ المعياري لمتوسط المجموعة الثانية، (N1) عدد العينة في المجموعة المرجعية، (N2) عدد العينة في المجموعة المستهدفة . ويمكن كذلك حساب درجة الحرية للاختبار بالاعتماد على الصيغة المصححة لـــ Satterthwaite's، والتي تمكن من الحصول على نفس القيمة، ويتم ذلك باستخدام صيغة المعادلة رقم (28):

$$m = \frac{(n_A -1)(n_B -1)}{(n_A -1)C_B^2 + (n_B -1)C_A^2} \text{.....(28)}$$

(Satterthwaite, 1946, pp.110-114) (Welch, 1947, pp.28-35)

11-3- القرار الإحصائي وتفسير الأداء التفاضلـــي للمفردة:

يفسر الأداء التفاضلـــي وتحديد ما إذا كان هناك تحيز لمفردات الاختبـــار أو المقياس طبقا لاختبار (ويلش)، من خلال ثلاثة مؤشرات أو محكات، وهي كالتالي :

1. قيمة اختبار "ت" يجب أن تكون محصورة ما بين 2.0 > t <2.0-.
2. قيمة التباين (DIF) يجب أن تكون محصورة ما بين 0.5 > DIF <0.5- لوغريتم.
3. قيمة مستوى الدلالة (p) يجب أن تكون محصورة ما بين 0.05 > p <0.05-.

(Linacre, 2012, p.659) (Juhaida et al, 2016, p.183)

ويمكن كذلك استخدام المعايير المطورة من طرف مؤسسة خدمات الاختبـــارات التربويـــة الإمريكية (Education Testing Services -ETS) للحكم على قيم الإحصائي (DIF contrast)، من خلال مؤشر لوغريتم نسبة الأرجحية التركمية لـــ مانتل – هانزل (Mantel-Haenzel log Odds-Ratio MH-LOR)، باستخدام طريقة الإنحدار اللوجستي (اللوغاريتمي)، اعتماداً على نظرية الاستجابـــة للمفردة الاختبارية، كالتالي:

1. النمط (A) أداء تفاضلي ضئيل: القيمة المطلقة لـــمؤشر $|DIF|$ = 0 وغير دال إحصائياً، إما: $|DIF|$ < 0.43.
2. النمط (B) أداء تفاضلي معتدل: القيمة المطلقة لـــمؤشر $|DIF|$ = 0 وغير دال إحصائياً، إما: $|DIF|$ ≥ 0.64 و $|DIF|$ > 0.43 و غير دال إحصائياً.
3. النمط (C) أداء تفاضلي كبير: القيمة المطلقة لـــمؤشر $|DIF|$ دال إحصائياً و $|DIF|$ > 0.43 أو $|DIF|$ ≥ 0.64. (Zwick, 1999, pp.1-28) . (Linacre, 2008)

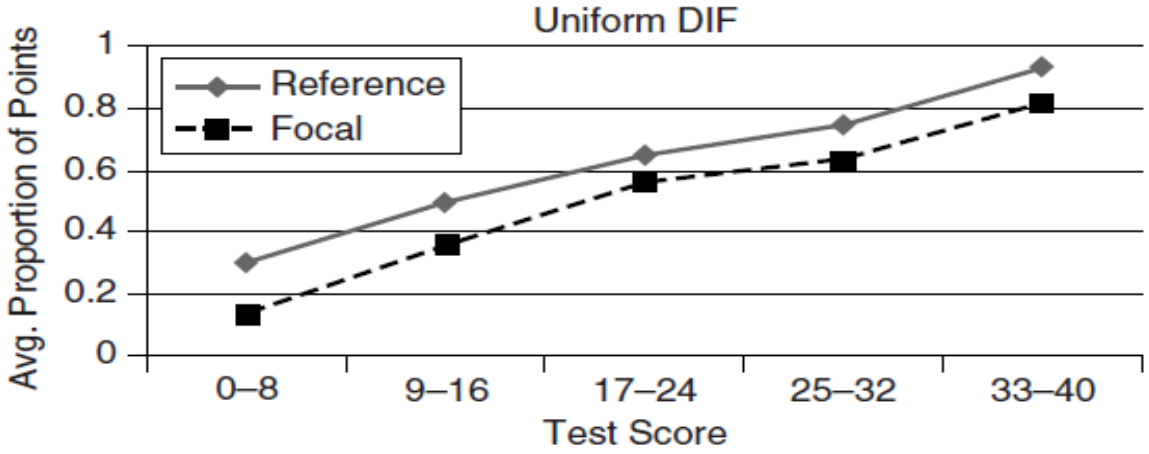
ويستدل على إتجاه الأداء التفاضلي من خلال إشارة قيمة الإحصائي (MH-LOR)، فالإشارة الموجبة تشير إلى أن المفردة تظهر أداء تفاضلياً لصالح المجموعة المرجعية (الأغلبية)، والإشارة السالبة تشير إلى أن المفردة تظهر أداء تفاضلياً لصالح المجموعة المستهدفة (الأقلية). (Penfield and Camrilli, 2000)

11-4- أنواع الأداء التفاضلـــي للمفردة الاختبارية:

بناء على نتائج العديد من الدراسات حول الأداء التفاضلي للمفردة (Hambleton and Rogers,1989 ; Pae,2004 ; Swaminthan and Rogers,1990) ظهر نوعان من الأداء التفاضلـــي للمفردة، وهما:

11-4-1- الأداء التفاضلي المنتظم (Uniform DIF):

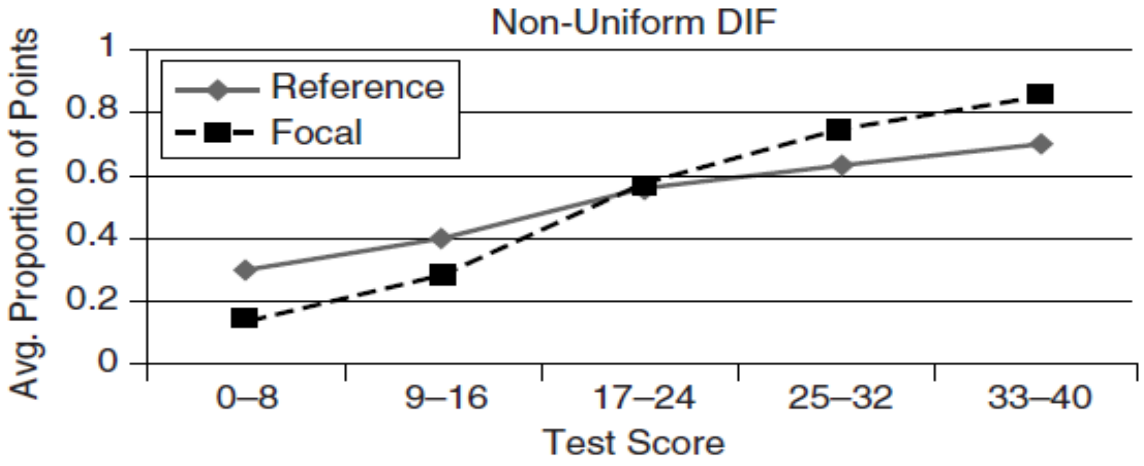
وهو التحيز الذي لا ينتج عن تفاعل مستوى القدرة مع عضوية المجموعة، أي أن احتمال الإجابة الصحيحة للمفردة يكون أكبر لمجموعة دون أخرى في جميع مستويات القدرة. و من خلال الشكل رقم (14) يلاحظ عدم وجود تفاعل بين مستوى القدرة وعضوية المجموعة، وإنما احتمال الإجابة الصحيحة على المفردة دائما أكبر للمجموعة المرجعية Reference . (Charles et al, 2012, p.305)



الشكل (14) يوضح الأداء التفاضلـــي المنتظم.

11-4-2- الأداء التفاضلي غير المنتظم (Non-uniform DIF):

ويظهر هذا النوع من التحيز عندما يحدث تفاعل بين مستوى القدرة (الأداء) وعضوية المجموعة، أي يظهر التحيز لصالح المجموعة المستهدفة (Focal) عند مستوى معين من القدرة، و عند مستوى قدرة آخر يظهر لصالح المجموعة المرجعية (Reference). ومن خلال الشكل رقم (15)، يلاحظ الأداء التفاضلي لصالح المجموعة المستهدفة عند المستويات الدنيا من القدرة، أما عند المستويات العليا فكان لصالح المجموعة المرجعية، فعند تقاطع المنحنيين يكون هناك تفاعل بين مستوى القدرة والمجموعة، وبالتالي ظهور الأداء التفاضلـــي غير المنتظم (Gruijter and Kamp, 2005).(حمادنة ، 2007، ص.08) (Charles et al, 2012, p.305)



الشكل (15) يوضح الأداء التفاضلـــي غير المنتظم.

11-5- كيفية التعامل مع المفردات التي تظهر الأداء التفاضلـــي:

عندما يتم التحديد الإحصائي للمفردات الشاذة، فإن مصدر الشذوذ ربما يجري توضيحه جزئياً من خلال مزيد من الإجراءات الإحصائية، مثل: تحليل البدائل الخطأ التي تم اختيارها في مفردة اختيار من متعدد. وهذه التحليلات التكميلية إلى جانب المراجعات التي تعتمد على الأحكام، ينبغي أن تقدم الأساس لأجراء لما هو مناسب. فالمفردة الشاذة ربما تحذف، أو تراجع ككل، أو تستبدل في أي جزء ضعيف فيها، أو ربما توسع تعليمات الاختبار أو توضح، أو ربما تستبقى المفردة الأصلية بعد إعادة تقييم مواصفات الاختبار ذاتها. ويمكن أن تكون النتيجة أما إجراء تغييرات في مواصفات الاختبار، أو توضيح ما يتعلق بالاستدلالات المعينة التي يمكن التوصل إليها من درجات الاختبار. (أناستازي، 2015، ص.254) .

كما يطرح بعض الباحثين تساؤلاً مهما وهو: ماذا يفعل الباحث في حالة العثور على مفردة يحتوي على الأداء التفاضلـــي؟ هل يقوم بحذف هذه المفردة؟ وهذا ما لا ينصح به، أم يقوم بالتحفظ عليها إلى غاية التحقق منها عن طريق المحكمين وعمل المزيد من الدراسات حول صلاحية المفردة. إن جزءاً من الإجابة على هذه التساؤلات يتعلق بمدى أهمية القرارات التي يتم اتخاذها على أساس عملية القياس هذه، أو مدى خطورة الوقوع في الخطأ جراء إتخاذ قرار خاطئ. غير أن بعض الباحثين يتحدثون عن إمكانية إعادة صياغة المفردات التي تحتوي على الأداء التفاضلي للمفردة، أو حذف تلك التي يكون حجمها كبير جداً، لكن قبل الحذف يجب الأخذ بعين الاعتبار تأثير هذا الحذف على صدق المحتوى للاختبار، والفقدان المحتمل للمعلومات من قياس مستويات مختلفة من السمة الكامنة. على العموم حذف هذه المفردات قد يكون غير مقبول مع الاختبارات التي لا تحتوي على عدد كبير من المفردات. ويشير أورلندو (Orlando, 2002) إلى

أن تأثير وجود الأداء التفاضلـــي للمفردة في مفردات اختبار ما يختلف حسب حجم الأداء التفاضلي للمفردة، وعدد المفردات التي تحتوي عليه في الاختبار، والغرض من استخدام الاختبار. (بوسالم وكريش، 2015، ص.74)

ملخص الفصل:

نتيجة للمشكلات المصاحبة للقيـــاس الكلاسيكي في العلوم النفسية والتربوية، والتي تتسبب في عدم دقة النتائج التي يمكن أن تسفر عنها أساليب وأدوات القياس، ظهرت الحاجة إلى تطوير أساليب القيـــاس والتقويم بشكل يتوافق مع أساليب القياس الفيزيقـــي. لهذا تعددت الكثير من البحوث والدراسات من قبل المهتمين بالقياس النفســـي والتربوي، والتي حاولت تقديم حلول لمعالجة مشكلات القياس. مما اسفر عن ظهور بعض الإتجاهات الحديثة في القياس أبرزها نظرية الاستجابة للمفردة، والتي تفترض إمكانية التنبؤ بأداء الأفراد أو تفسير أدائهم في ضوء خاصية أو خصائص مميزة لهذا الأداء تسمى السمات، وهي تفترض وجود واحد أو أكثر من هذه السمات تكمن خلف استجابات الفرد لمفردات الاختبار. وقد بنيت هذه النظرية وغيرها من النماذج اللوغاريتمية، ومنها نموذج (راش) والذي يرى بأن النموذج يعتمد على قدرة الفرد وصعوبة المفردة وأن احتمال حدوث هذه الاستجابة للمفردة، دالـــة للفرق بين قدرة الفرد وصعوبة المفردة، كما يعتمد على عدة فروض، وهي: أحادية البعد، الاستقلال الموضعي، المنحنيات المميزة للمفردة، التخمين.

لهذا تطرقنا في هذا الفصل، إلى كيفية توظيف نظرية الاستجابة للمفردة في بناء الاختبارات وذلك بتوضيح نماذج نظرية الاستجابة للمفردة أحادية البعد و افتراضاتها، ومن ثم عرض تصنيف لنماذج نظرية الاستجابة للمفردة، ثم التوسع في تحديد مفهوم نموذج راش اللوغاريتمـــي أحادي البارامتر، وإلى كيفية تقدير بارامتراته من حيث صعوبة المفردات، وقدرات الأفراد، ثم إلى صدق وثبات الاختبار على ضوء نموذج (راش)، ومنه إلى كيفية إجراء اختبار حسن مطابقة البيانات لنموذج راش، وأخيراً كيفية الكشف عن الأداء التفاضلـــي للمفردة الاختبارية، وأنواعه وإتخاذ القرار بخصوص المفردات التي تبدي أداء تفاضلياً.

الفصـــل الرابـــع

التفكيـــر الهندســـــي

تمهيـــد:

بالرغم من أن عمليات التفكيـــر مازالت غامضة ولم تدرك بوضوح أعماقها بعد، إلا أنها تتحكم في كثير من وظائف الفرد وأفعاله، على المستويين الشعوري واللاشعوري. فالتفكير متضمن في كل مهمة يقوم بها الفرد، على أي مستوى من المستويات. وبالنسبة لعملية التدريس، فإن القرارات الخاصة بتصميم المواقف التدريسية من قبل المعلم، والقرارات الخاصة بتنفيذ المتعلم للمهام التي تتطلبها المواقف، تعتمد بالدرجة الأولى على آليات التفكير عند المعلم والمتعلم. لذا تعد تنمية التفكير من أهم أهداف تدريس الهندســـة، والتي تعتبر أحد المجالات المهمة في منظومة تعليم الرياضيات وتعلمها، باعتبارها مدخلاً لتطوير قدرات التلامذة ومهاراتهم والتي تمكنهم من اتقان كثير من الموضوعات التعليمية الأخرى، وكذلك تطوير مهاراتهم الحياتية، كما تشكل أحد أهم الإستراتيجيات التدريسية التي تساعد التلامذة في مواجهة الصعوبات التي تعترضهم أثناء تعلم الهندســـة. لذلك سيستهل الباحث هذا الفصل أولا بمفهوم التفكير؛ ثم مفهوم الهندســـة؛ وأخيراً التفكير الهندســـي ونظرية (فان هيل) لتطور التفكير الهندســـي على التوالي.

1- التفكيـــر:

1-1-ماهية التفكير:

التفكير نشاط عقلي به نكتسب المعارف ونحل به المشكلات والذي يظهر سلوكناً على أكثر ما يكون منطقية ومعقولية، وبه نكتشف من المعارف ما يؤمن لنا المزيد من السيطرة على العالم الذي نعيشه، والتفكير كما هو مستخدم في الحياة العامة مفهوم يغطي اضرباً متعددة من النشاطات العقلية (راضي، 2003، ص.437). ويعد التفكير من الظواهر النمائية التي تتطور عبر مراحل العمر المختلفة، حيث أن الأفراد ومنذ سن الطفولة، يدركون بسرعة بأننا نفكر، وأن لديهم سرعة البديهة لإبداء آراءهم حول ما نفعله عندما نفكر. كما يمارس الأطفال ومنذ ولادتهم ما سماه (بياجيه) التفكير الحس- الحركي، وتفكير ما قبل العمليات في الطفولة المبكرة، ثم التفكير المادي في مرحلة الطفولة المتأخرة، وأخيراً التفكير مع بداية مرحلة البلوغ (عدنان، 2009، ص.17). غير أن ما هية عملية التفكير لا تزال من الظواهر النفسية غير المقطوع بها في علم النفس، فقد وجد في تاريخ الفكر السيكولوجي من يراه فعالية عقلية رمزية بينما يراه آخرون فعالية عقلية تصوريـــة، ثم أنه برأي آخرين حركة عضليـــة تقوم بها أعضاء النطق بحيث يبدو

التفكير أشبه ما يكون بلغة صامتة، في حين يراه آخرون أشبه ما يكون بعملية اختزال (راضي، 2003، ص.438).

ويرى الباحث بأن التفكير هو عملية عقلية تصورية، بمعنى أن الذهن يبني بواسطة مصادره ومبادئه البنيوية الخاصة تمثيلاً ذهنياً، انطلاقاً من الصور الذهنية التي يكونها التلميذ، وتأويلها من خلال خبرات التعلم والتعليم التي يمر بها، وهو ما تدعمه أفكار (بياجيه) و (فان هيل) في تطور التفكير الهندســي.

1-2- تعريف التفكير:

تباينت وجهات نظر العلماء والباحثين التربويين حول التعريف العام للتفكير، إذ أنهم قدمو تعريفات مختلفة إستناداً إلى أسس وإتجاهات نظريــة متعددة، فمنهم من يعرفه على أنه عملية سلوكيــة خارجية، وآخرون يرون أنه عملية معرفيــة داخلية (عدنان، 2009، ص.18) . لذا يمكن استعراض عدد من أهم التعريفات التي تناولت التفكير بهدف مقارنتها والوصول إلى تعريف توفيقي للتفكير.

تعرفه موسوعة علم النفس التربوي، بأنه: " كل نشاط ذهني أو عقلي يتضمن سيلاً من الأفكار تبعثه وتثيره مشكلة أو مسألة تحتاج إلى حل، فهو لا يحدث إلا إذا سبقته مشكلة تتحدى عقل الفرد. فالتفكير مفهوم افتراضي يتضمن سيلاً أو تواردا غير منظم من الأفكار والصور والذكريات والانطباعات العالقة في الذهن"(عدنان، 2009، ص.19). ويعرفه (جون ديوي)، بأنه: " العملية التي يتم بها توليد الأفكار عن معرفة سابقة، ثم إدخالها في البنية المعرفية للفرد، وهو أيضاً معرفة العلاقة التي تربط الأشياء ببعضها، والوصل إلى الحقائق والقواعد العامة" (دياب، 2000، ص.25). وعرفه الشرقاوي (1989) نقلاً عن فارس، بأنه: " نشاط عقلــي واعي يعكس فيه الإنسان الواقع الموضوعي بطريقة مختلفة تماماً عما يحدث في الإحساس والإدراك، وهو عملية عقلية معرفية يتم من خلالها إنعكاس العلاقات والروابط بين الظواهر والأشياء في وعي الانسان" (فارس، 2011، ص.15). و عرفه مطر (2004)، بأنه:"عملية عقلية معرفية وجدانية عليا تبنى وتؤسس على محصلة العمليات الأخرى، كالإدراك والإحساس والتخيل وكذلك العمليات العقلية كالتذكر والتجريد والتعميم والتمييز والمقارنة والاستدلال فكلما اتجهنا في تفكيرنا من المحسوس إلى المجرد كلما كان أكثر تعقيداً " (2004، ص.05).

و يعرفه الباحث بأنه عملية عقليـــة معرفية مركبة، وبأنه نشاط واعي يقوم به الفرد عند الاستثارة كردة فعل على مشكلة تتحدى عقله، وبالتالي تجنيد معرفته السابقة، ثم إدخالها في بنيته المعرفية للتكيف مع واقع بيئته الجديد.

1-2-نظريات التفكير:

تطرقت غالبية النظريـــات والإتجاهات المختلفة في علم النفس إلى مفهوم التفكير، وحاولت تفسيره وفق مبادئها ومفاهيمها. لذلك يمكن تلخيص أهم هذه النظريات، على النحو التالي:

1-3-1- النظرية السلوكية:

لم تركز المدرسة السلوكيـــة على تفسير التفكير بشكل مباشر، وإنما اعتبرت أن الخبرة أو التعلم الذي يتشكل نتيجة العلاقة بين المثير والاستجابـــة هي بمثابة التفكير.(فارس، 2011، ص.33)

1-3-2- النظرية المعرفيـــة:

تعد النظرية المعرفيـــة من أهم النظريـــات التي فسرت التفكير، حيث تبلور ذلك من خلال دراسة الأسس الفسيولوجية للمعرفة؛ وإتجاه معالجة المعلومات؛ ونظرية (بياجيه):

أ. الإتجاه الفسيولوجي: حاول أصحاب هذا الإتجاه تفسير السلوك الانسانـــي بشكل عام، والتفكير بشكل خاص، من خلال ربط سلوك الإنسان لما يجري داخل الجسم من عمليات فسيولوجية عديدة في الجهاز العصبي داخل الدماغ، بدلاً من التركيز على محاولة فهمها كعملية معرفيـــة مجردة . فإذا أردنا أن نفهم كيف يحل التلميذ مسألة في الرياضيات، فلا بد من دراسة الدماغ وتتبع التغيرات التي تطرأ على دماغه خلال حل المسألة الرياضية.

ب. إتجاه معالجة المعلومات: تبلور هذا الإتجاه مع تطور نظم الحواسيب والإتصال، وبدأ العلماء بدراسة الخطوات والمراحل التي من خلالها معالجة المعلومات وفق نظام معالجة يتسم بالتسلسل والتنظيم ويحاكي نظم معالجة المعلومات في الحاسوب.(عدنان والعتوم وأخرون، 2009، ص.32)

ج. نظرية (جون بياجيه) في النمو المعرفـــي: يعتقد (بياجيه) أن هناك وظيفتان أساسيتان للتفكير، هما التنظيم والتكيف، وتتمثل وظيفة التنظيم من خلال نزعة الفرد إلى ترتيب وتنسيق الأنشطة المعرفية بشكل متكامل، بينما تشير وظيفة التكيف إلى نزعة الفرد إلى التلاؤم والتآلف مع البيئـــة الخارجية. ويتحقق

التكيف من خلال عملية التمثل والاستيعاب، حيث يتحقق التمثل عن طريق دمج الفرد للمعلومات داخل البناء المعرفـــي حتى يتحقق الفهم والإدراك، بينما يتم الاستيعاب من خلال نزعة الفرد لتغيير تراكيبه المعرفية لتواجه مطالب البيئة الخارجية. وقد أكدت نظرية (بياجيه) على أهمية تطور التفكير وفقا لمراحل النمو المعرفية المعروفة لديه. (عدنان والعتوم وأخرون، 2009، ص.33)

وتتفق أفكار (بياجيه) وخاصة مراحل النمو المعرفي لتطور التفكير، مع تصورات وأفكار (فان هيل) لمستويات التفكير الهندسي الهرمية، بالرغم أن (فان هيل) لم يحدد مراحل عمرية معينة لكل مستوى.

1-3-3- **النظرية الجشطالتية كوهلر(Kohler):**

ترى هذه النظرية أن التفكير يجب أن يتم بصورة كلية، من خلال النظرة الكلية للموقف، وإدراك العلاقات القائمة بين عناصر الموقف، مما أدى إلى تحديد ما عرف بالتعلم بالتبصر أو الاستبصار الذي يعتمد على الربط بين عناصر الموقف للوصول إلى الحل. وحددت هذه النظرية عدد من العوامل التى ترى بأنها تؤثر على التفكير والإدراك، عرفت بقوانين الإدراك مثل: قانون الصورة والخلفية؛ والإغلاق؛ والتشابه؛ والتقارب، والتي تشير جميعها إلى امكانية تحقيق الفهم من خلال السياق والمجال الذي يحدث فيه الإدراك، ومما يوجه التفكير وفق هذه القوانين.(فارس، 2011، ص.33)

ويعد مفهوم الاستبصار من أهم المفاهيم التي اعتمد عليها (فان هيل) في نظريته لمستويات التفكير الهندســـي، والتي أخذها من النظرية الجشتطالتية، ويقول (فان هيل) أن الاستبصار موجود عندما يقوم شخص بموقف جديد على نحو كاف، ويضيف أنه وعلماء الجشتالت يقولون الشيء نفسه لكن بكلمات مختلفة.

1-3-4- **نظرية ليف فيجوتسكي(Lev Vygotsky,s):**

يعتقد (فيجوتسكي) أن هناك تطوراً من الأشكال الدنيا إلى الأشكال العليا من التفكير خلال عملية النمو والتطور، ولذلك يكون هناك إنتقال للتحكم من البيئة الى الفرد. ويؤكد (فيجوتسكي) أن التفكير له أصل اجتماعي، حيث ينمو مع التطور النفسي الاجتماعي.(عدنان، 2009، ص.34)

وعبر عن ذلك (فيجوتسكي) بمفهومه المعروف: منطقة النمو المحيط بالمركز The Zone of Proximal Development (ZPD)، والتي عرفها بأنها المسافة بين مستوى نمو التلميذ الفعلي كما يحدده

أداؤه في حل مشكلة ما بمفرده، وبين المستوى الأعلى للنمو كما يحدده أداؤه في حل مشكلة تحت توجيه المعلم، أو المشاركة مع أقران أكثر تعلماً وخبرة.(مجدي عزيز، 2005، ص.455)

4-1-تصنيفات التفكير:

صنف التفكير إلى عدة مستويات حسب درجة تعقيد كل نمط من أنماط التفكير المختلفة، وقد ميز نيومان (Newmann,1990) بين مستويات التفكير العليا والدنيا، حيث استخلص أن مهارات التفكير الدنيا (الأساسية) تتطلب فقط التطبيق الآلي الروتيني للمعلومات المكتسبة سابقا، مثل: إسترجاع المعلومات المخزونة في الذاكرة، والإهتمام بالأرقام في القوانين المتعلمة سابقاً، وعلى العكس فإن مهارات التفكير العليا تتطلب حث التلميذ على الاستنتاج وتحليل المعلومات. كما أوضح (نيومان) نقطة مهمة هي أن تحديد مستويات التفكير الدنيا أو العليا يعتبر أمراً نسبياً، فالمهام أو الأعمال التي تتطلب مستويات دنيا من أحد الأشخاص قد تتطلب مستويات عليا من شخص أخر. وتبعاً لذلك فتحديد المدى الذي يحتاجه الفرد للتفكير في مشكلة ما يتطلب تحديداً لمستوى ذكائه. (ابرهيم، 2005، ص.111) (عدنان، 2009، ص.25)

وصنف جاكوبسن وآخرون (Jacobsen and Others) مهارات التفكير إلى ثلاثة مستويات رئيسة تتمثل، في: (أ). العمليات المعرفية الأساسية وتشمل الملاحظة، المقارنة، الاستنتاج، التعميم، فرض الفروض، الاستقراء، والاستدلال؛ (ب). العمليات المعرفية العليا وتشمل حل المشكلات، وإصدار الأحكام، التفكير الناقد، والتفكير الابداعي؛ (ج). ما وراء العمليات المعرفية أو التفكير من أجل التفكير.

فمهارات التفكير لا يمكن تعلمها لذاتها وبصورة مباشرة، ولكن يتم تعلمها من خلال مواقف التعلم المباشرة، التي يمكن أن تتوافر عند استخدام الأنشطة الإثرائية في التدريس. إذا، التفكير هو أكثر النشاطات المعرفية تقدماً، ينجم عن قدرة الكائن البشري على معالجة الرموز والمفاهيم، واستخدامها بطرق متنوعة تمكنه من حل المشكلات التي يواجهها في المواقف التعليمية والحياتية المختلفة.

حيث ركز (جان بياجيه) في كتابه "التطور المعرفـــي"، على ارتباط التفكير بموضوعه من ناحية، وطبيعة المرحلة التي يمر بها الطفل من ناحية أخرى بحيث يتحدد ذلك من التفكير البسيط إلى المعقد الذي يرتبط ارتباطاً وثيقاً في طبيعة المرحلة المعرفية التي يمر بها الفرد. (مجدي، 2005، ص.04)

ويرى الباحث أنه على الرغم من الاختلاف بين مستويات التفكير العليا ومستويات الدنيا، إلا أنه يمكن تدريسهما داخل القسم الدراسي الواحد، فحاجة التلميذ لاستخدام مستويات التفكير العليا أو الدنيا يتوقف على طبيعة المهمة أو المشكلة وكذلك ذكاء التلميذ نفسه.

1-5-أدوات أوعنـــاصر التفكيـــر:

توصف اللغة والمفاهيم والصور بأنها اللبنات الأساسية في التفكير. وتقدم هذه العناصر منفردة ومجتمعة في تارات أخرى العناصر الأساسية التي يبني بها التفكير ويتيسر بها إجراء العمليات الفكرية (راضي، 2003، ص.481).

1-5-1- **اللغـــة:** هي وسيلة التخاطب، وتشمل على مجموعة من الرموز المعرفية التي تمكن الفرد من التعبير عن خبراته ومعارفه وغالباً ما تكون اللغة صامتة لتدل على ما يدور بين الفرد ونفسه (حوالة، 2005، ص.213).

1-5-2- **المفاهيـــم:** المفاهيم عنصر للتفكير فلا يمكن أن نفكر دون استخدام ألفاظ ومعاني تعبر عن مفاهيم معينة، والمفهوم تكوين عقلي ينشأ عن تجريد خاصية أو أكثر من مواقف متعددة تتوفر فيها هذه الخاصية حيث تعزل هذه الخاصية عما يحيط بها من أي من المواقف المعنية وتعطى إسماً أو رمزاً للتعبير عنها، وتختلف المفاهيم في مدى تعقدها، كما أنها تعمل على إرتقاء مستوى التفكير.(عبيد وعفانة، 2003، ص.28) (فارس، 2011، ص.31)

1-5-3- **الصور الذهنيـــة:** هي رموز عقلية يستحضر بها الفرد صور الأشياء عند التفكير في موضوع ما، ويستخدم الفرد في تفكيره جميع قدراته الحسية (بصرية وسمعية ولمسية وعقلية وحركية) لربط صور الأشياء ببعضها عن طريق الحواس، مثل تخيل الفرد لملمس الشيء أو تخيل حركة ما، وغيرها من الصور الحسية. (حوالة، 2005، ص.213).

1-6- دور المفاهيـــم في التفكيـــر:

تكوين المفاهيم من الأمور الأساسية في تسهيل التفكير والتعامل مع الأشياء الموجودة في العالم، ومن المعروف أن الأشياء التي تحيط بنا تتعدد وتتنوع إلى مستوى غير محدود، ولكنها مع ذلك تقبل أن تصنف إلى أجناس وفصائل ومراتب وفقاً لصفات جوهرية تلتقي عليها مجموعة من هذه الأشياء يطلق

عليه إسم ما، تندرج تحته جميع هذه الأشياء، وعلى هذا فتكوين المفهوم يبدأ عندما يكون باستطاعتنا وضع مجموعة من الأشياء في وحدة متميزة من غيرها. ومن ذلك مثلاً إننا عندما نواجه مجموعة من الأشياء الملونة مختلفة ونجمع الأشياء الملونة بلون واحد ندرج فيه كل لون تنطبق عليه الصفات الأساسية لهذا اللون، وقد نلاحظ أشكالاً هندسيـــة تكون فيها السطوح محصورة بخطوط مستقيمة ثلاثة فتكون من ذلك مفهوم المثلثية. ويختبر مدى تكون المفهوم بالقدرة على تمييز الأفراد الذين يمكن أن يدرجوا تحته، فإذا رأيت سطحاً محصوراً بين ثلاثة أضلاع استطعت أن تميزه بأنه مثلث بالرغم من شكل الزوايا الناتجة عن التقاء الأضلاع أو سعة المساحة المحصورة بينهما. إذاً المفهوم بهذا المعنى يمثل نسقاً Schema عقلياً أو نموذجاً معرفياً مثالياً نحاكم في ضوئه الأشياء والحوادث ونحكم عليها بأنها تتصل بهذا المفهوم أو ذاك.(راضي، 2003، ص.488) (فارس، 2011، ص.86).

وتشير اللجنة القومية لمعلمي الرياضيات بأمريكا (NCTM ,1989)، إلى أن المفاهيم هي جوهر العملية الرياضية، وأن الرياضيات تصبح ذات معنى وأكثر فهماً ووضوحاً إذا أدرك التلامذة المفاهيم الرياضية ومعناها وتفسيرها .(ماضي، 2011، ص.102)

2- الهندســـة:

تعتبر الهندســـة أحد فروع الرياضيات الأساسية، التي إجتذبت مؤرخي العلم والتربية أكثر من أي فرع آخر، نظراً للأهمية التي وضعها الإغريق القدماء للهندســـة معياراً للتفكير السليم، والدور الأساسي الذي قامت به في التطور التاريخي لعلم الرياضيات. فهي تعمل على توسيع قدرات التلامذة العقلية وتنمية أساليب التفكير المختلفة لديهم، وتتيح الفرص لهم لإكتشافات منظمة ومتتابعة تساعدهم على تمثيل وشرح ووصف وفهم العالم المحيط بهم وتحليل المشكلات أو حلها .(سعيد، 2007، ص.168)

وتعود كلمة هندســـة في اللغة الأجنبية (geometry) في أصلها الى الكلمة اليونانية (geo)، إذ يقصد بها الأرض، و (metry) ويقصد بها القياس، وعلى هذا الأساس يكون المقصود من الكلمة قياس الأرض (أبو ملوح، 2002، ص.20). وقد نشأت الهندسة في مصر القديمة كوسيلة عملية لقياس الأطوال والزوايا والمساحات والحجوم وذلك لحاجتهم لعمل القياسات الخاصة بالفيضانات وإيجاد مساحة الأراضي للإستفادة من مياه الفيضانات وتحديد الضرائب وإيجاد أحجام الأشكال لحاجتهم إلى بناء المعابد والأهرامات. ولو أن

المصريين القدماء إكتشفوا بعض العلاقات بين الأشكال الهندسية إلا أنهم لم يستطيعوا وضع تعميم لهذه العلاقات فمثلاً إستطاعوا إكتشاف العلاقة التي تربط أضلاع المثلث القائم أي المثلث الذي أطوال أضلاعه 5،4،3 يكون قائم الزاوية إلا أنهم لم يستطيعوا تعميم ذلك إلى صيغة نظرية فيثاغورث. (نظلة، 1985، ص.57). وكانوا يستخدمون قواعد بدائية وتقريبية، كقولهم أن مساحة شكل رباعـــي أطوال أضلاعه d.c.b.a هي $\frac{a+c}{2} x \frac{b+d}{2}$، مع الملاحظة أن هذه القاعدة صحيحة عندما يكون الشكل الرباعي مستطيلاً .(سعد الله، 2009، ص.163). وتوصل البابليون كذلك إلى مفهوم التشابه وإيجاد علاقات (قوانين) المساحة، والحجوم . كما درسوا بعض العلاقات الهندسيـــة مثل المثلث المرسوم في نصف دائرة يكون قائماً، والعلاقات بين المضلعات المنظمة. كما يعتبر الأغريق (اليونان القدماء) أول من قدم فكرة البرهان (البرهان المباشر والبرهان بالتناقض أو الحذف)، ووضعوا الهندسة في القالب المنطقي الذي نألفه اليوم، وقد أسهم كثير من الرياضيين في تطور الهندسة في ذلك العهد ومنهم فيثاغورث (500 ق.م). إلا أن الهندســـة حتى هذا الوقت كانت محتاجة إلى نوع من الربط والتنسيق، ويعزى إلى إقليدس (300 ق.م) عمل هذا الربط وخلق نظام البديهيات.(سلامة، 1995، ص.92) (نظلة، 1985، ص.64). وكانت أعمال هذه المدرسة قد توجت بدراسة هندسية بالغة الأهمية هي المعروفة بأصول إقليدس (les Elements) التي ظلت خلال 20 قرناً منطلق كل عمل هندســـي. وتضم أصول إقليدس 15 كتاباً منها 13 كتاباً من تأليف إقليدس، وتتناول هذه الكتب الـــ 13 الأشكال الهندسية والمضلعات المحيطة بالدائرة والمحاطة بالدائرة والنسب والتشابه والهندســـة الفضائية ونظرية الأعداد القابلة للقياس .(سعد الله، 2009، ص.165).

لقد عرف إقليدس بطريقته النقطة (ما انعدم جزؤه)، الخط المنتهي (طول بدون عرض طرفاه نقطتان)، والمستقيم والسطح (ما له عرض وطول فقط)، والمستوي (سطح موضوع بين مستقيماته)، والزاوية (باستخدام ميل المستقيمين المشكلين للزاوية)، والدائرة و المستيقمات المتعامدة و المثلث، وبعدئذاً وضع مسلماته الخمس، كالتالي:

I. مسلمـــات اقليـــدس في الهندســـة المستويـــة:

1. لتكن A و B نقطتين. يوجد مستقيم يمر عبر كل نقطتين A و B.

2. يمكن تمديد كل قطعة مستقيمة [AB] إلى مستقيم يمر بالنقطتين A و B.

3. من أجل كل نقطة A و نقطة B مختلفة عن A، يمكن وصف دائرة مركزها A وتمر بالنقطة B.

4. كل الزوايا القائمة متساوية فيما بينها.

5. من نقطة خارج مستقيم معلوم يمكن تمرير مستقيم واحد يوازي المستقيم المعلوم.

وقد وجد في بعض اصدارات أصول اقليدس مسلمة سادسة تقول، بأنه لا يمكن لمستقيمين أن يحيطا بالفضاء . لكن بعض المؤرخين يرون في هذه المسلمة إضافة لم يكن اقليدس صاحبها. وقد بنى اقليدس هندسته على 5 مسلمات سابقة الذكر وأضاف إليها 9 بديهيات ذات صلة بالهندسة والحساب ومنها:

1. الأشياء المتساوية لنفس الشيء تكون متساوية.

2. إذا اضيفت متساويات إلى متساويات كانت النواتج متساوية.

3. إذا طرحت متساويات من متساويات كانت النواتج متساوية.

4. الأشياء التي تطابق أحدها الآخر تكون متساوية.

5. الكل أكبر من الجزء. (سعد الله، 2009، ص.165) (سلامة، 2008، ص.24)

بالإضافة إلى ذلك فإن المسلمات تحدد العلاقات بين المفردات غير المعرفة فتساعد على فهم معناها، فالمسلمات تعتبر الأساس الذي تشتق منه عبارات صحيحة أخرى يتم البرهنة على صحتها باستخدام قواعد المنطق الفرضي وتسمى مبرهنات. وقد استطاع اقليدس أن يبرهن 465 نظرية باستخدام هذه المسلمات الخمس، مما كون ما سمي بالهندسة الإقليدية التي تتكون منها غالبية كتب الهندسة في المرحلتين المتوسطة والثانوية (سلامة، 2008، ص.192)

وأخذ العرب كتاب اقليدس الأصول وترجموه إلى لغتهم وتفهموه جيداً وزادوا على نظرياته، فقد قدم إبن الهيثم نظريات ومسائل منها : " كيف ترسم مستقيمين من نقطتين مفروضتين داخل دائرة معلومة إلى أي نقطة مفروضة على محيطها بحيث يضعان مع المماس المرسوم من تلك النقطة زاويتين متساويتين". كما قدم البيروني (938– 1048) برهان لمساحة المثلث بدلالة أضلاعه وهو غير البرهان الذي أتى به هيرون من قبل. واستخدم ابن الهيثم والذي توفي سنة 1038م وغيره من العرب الهندسة المستوية والمجسمة في أبحاث الضوء وتعيين نقطة الانعكاس في أحوال المرايا الكروية، والاسطوانية، و المخروطية، المحدبة منها والمقعرة.

ونحب أن نشير هنا إلى أن الغرب عرفوا هندســـة اقليدس عن طريق العرب فترجموها عن العرب وأخذت تدرس في مدارسهم حتى سنة 1583م حينما إكتشف أصل هندسة اقليدس اليوناني(نظلة حسن، 1985، ص88). ويعرف أسلوب اقليدس بأسلوب البديهيات أو المسلمات، والذي بني على مسلماته الخمسة المشهورة، والذي يبدأ بمصطلحات غير معرفة مثل: النقطة والمستقيم والمستوى، والذي يسعى إلى بناء التعليل الاستنتاجي من خلال دراسة الهندسة. ولكن الدراسات الحديثة أوجدت بدائل لأسلوبه أكثر نجاحاً في تنمية القدرة لدى التلامذة على تبرير مقتراحاتهم (Bell,1976)، حيث أخذت العديد من دول العالم في الثمانينات ببناء مناهج الهندســـة في مدارسها على أسس نظرية فان هيل(van hiele)، حيث قدم (فان هيل) في 1957 أنموذجاً لنمو وتطور التفكير في الهندسة. يؤكد فيه على أن تطوير المفاهيم والأفكار الهندسية يمر بمستويات ذات طبيعة هرمية، تبدأ بملاحظة الأشكال، ثم تحليل خواصها، ثم إدراك العلاقات بين الأشكال المختلفة، وبالتالي صياغة استنتاجات منطقية تتعلق بها.(clements and Battista,1995) (عبد القوي، 2007، ص.166). والتي جسدت ميدانياً من خلال دراسة كلٍ من يوسيسكين (1982) في مشروع جامعة شيكاغو والتي من أهم نتائجها تصميم اختبار يقيس مستويات التفكير الهندسي الخمسة لفان هيل، ودراسة (Burger & Shaugnessy, 1986) بمشروع جامعة أوريغون، ودراسة (Fuys, Geddes & Tischler, 1988) بمشروع كلية بروكلين. والتي استطاعوا فيها تصنيف التلامذة وفقاً لهذه المستويات.

2-1- ماهية الهندســـة وتعريفها:

إذا أردنا تعريف الهندسة فإنه يصعب علينا تقديم تعريف دقيق لها، حيث أن الهندسة تنقسم إلى عدة فروع نذكر منها الهندســـة الإقليدية والناقصة والزائدة والكروية والتحليلية و الاسقاطية ويعد التوبولوجي أحد تلك الفروع. وتورد (الطنة) نقلاً عن (أبو ملوح)، بأن الهندسة هي ذلك الفرع من فروع الرياضيات الذي يبحث خصائص الأشكال الهندسية في المستوى والمجسمات في الفراغ والعلاقات القائمة بينهما وذلك اعتماداً على المسلمات وما يشتق منها من نظريـــات (الطنة، 2008، ص.27). وقد حاول الرياضي الألماني فيلكس كلاين (Felix klein) أن يعرف الهندســـة باستخدام فكرة التحويلات فذكر أن الهندســـة هي دراسة نوعية الأشكال (أو الفئات) ودراسة الخواص التي لم تتغير نتيجة هذه التحويلات (كاظم، 1970، ص.20).

2-2- البنية الهندسيـــة:

يذكر (أبو ملوح) نقلاًعن (القاطوني) في هذا السياق بأن التركيب البنيوي للهندســة يتألف من المكونات الخمس التالية:

أ. المفاهيم الأولية أو المفاهيم غير المعرفة: ويقصد بها مجموعة من المصطلحات الهندسيــة الأساسية التي يسأل عن تعريفها أبداً، ومنها النقطة والخط المستقيم والسطح.

ب. المفاهيم المعرفة: وهي تلك المفاهيم التي تعرف من خلال المفاهيم الأولية (اللامعرفات) فعلى سبيل المثال يعرف متوازي الأضلاع بأنه شكل رباعي فيه ضلعين متقابلين متوازيين.

ج. البديهيات والمسلمات: ويقصد بها تلك العبارات الهندسيــة التي بلغت من الوضوح حداً لا تحتاج معه إلى برهان أو دليل لإثبات صحتها.

د. النظريات: النظرية عبارة عن جملة تربط بين عدد من المفاهيم بعلاقة معينة يمكن البرهان على أنها صحيحة اعتماداً على مسلمات النظام فقط.

هـــ. البرهان: المقصود بالبرهان أنه عبارة عن مجموعة من الخطوات المرتبة بصورة متتالية منطقياً، والصادقة بناءً على مسلمــات النظام، التي من خلالها نستطيع اثبات صحة نظرية أو خطئها.(أبو ملوح، 2002، ص.21)

2-3- المفاهيم الهندسيــة وطبيعتها:

المفاهيم الهندسيــة أفكار مجردة يمكن وصفها أو تعريفها و لا يمكن إدراكها بالحواس، فما من أحد رأى الخط المستقيم، ولكننا نرى أشياء نصفها بأنها مستقيمة، فنقول أن حافة الورقة مستقيمة وحافة المكتب مستقيمة وهكذا، ويمكن أن نرسم أشكالاً على ورقة بيضاء ونقول أن هذه نقطة وهذه قطعة مستقيمة وهذه زاوية وهكذا. و تصنف المفاهيم الهندسيــة (والرياضية بشكل عام) إلى نوعان، هما:

1. مفاهيم أولية غير معرفة يدرك معناها ونصفها ولا نستطيع تعريفها.
2. مفاهيم معرفة وهي مفاهيم يمكن تعريفها بعبارة تحدد معناها بدقة ووضوح.

والمفهوم الرياضي بناء عقلي، وهو تجريد عقلي لخواص مشتركة ومميزة لمجموعة من الأشياء أو الأحداث التي يمكن ملاحظتها، تسمى هذه المجموعة مجموعة المرجع للمفهوم، وعناصر هذه المجموعة

تسمى أمثلة المفهوم، وأمثلة المفهوم يمكن أن تكون أشياء مدركة بالحواس(مفاهيم المجموعات)، أفعال (مفاهيم عملياتية)، مقارنات (مفاهيم علائقية)، أو تنظيمات (مفاهيم بنائية).

كما تسمى الخواص المشتركة والمميزة لأمثلة المفهوم بالخواص الجوهرية للمفهوم أو مسلمات المفهوم. وعندما نعرف مفهوماً ما فإن التعريف يجب أن يتضمن هذه الخواص مفصلة أو مختصرة. وبتحليل هذا التعريف للمفهوم سنجد أن للمفهوم خمسة أركان، وهي:

1. **الخواص الجوهرية (الأساسية):** وهي الخواص المشتركة بين الأشياء أو الأحداث التى تكون المفهوم، والمميزة لها عن غيرها . وتعتمد هذه الخواص لتصنيف الأشياء إلى أمثلة إنتماء للمفهوم أو عدم إنتماء.
2. **مصطلح المفهوم:** وهو الإسم أو الرمز الذي يطلق على المفهوم بعد تحديد خواصه الجوهرية.
3. **أمثلة المفهوم:** وهي كافة الأشياء أو الأحداث التي تتوفر في كل منها خواص المفهوم الجوهرية وكل واحد منها يكون مثالاً على المفهوم.
4. **تعريف المفهوم:** وهو تجميع أو تلخيص للخواص الجوهرية في عبارة بهدف تحديد المعنى الدقيق والواضح للمفهوم.
5. **الخواص الثانوية للمفهوم:** وهي خواص يمكن استنتاجها والبرهنة على صحتها اعتماداً على الخواص الجوهرية للمفهوم والمعرفة الرياضية التي سبقت هذا المفهوم في تسلسل عناصر البنية التي ينتمي إليها المفهوم. وتصنف الخواص الثانوية للمفهوم إلى ثلاثة أصناف أو أنواع، وهي: (أ) خواص تتوفر في جميع أمثلة المفهوم وتسمى خواص ثانوية ثابتة؛ (ب) خواص تتوفر في بعض أمثلة المفهوم وتسمى خواص ثانوية متغيرة؛ (ج) خواص تتناول علاقة بين مثالين من أمثلة المفهوم وتسمى خواص ثانوية علائقية.(سلامة، 2008، ص.17)

وقد يكون ممكناً تناول هذه الأركان الخمسة دفعة واحدة في المنهاج الدراسي ولذلك يقدم منها في مرحلة ما، ما هو مناسب لتلك المرحلة العمرية للمتعلم وبصورة تتمشى مع درجة نضجه وإستعداده. وفي مرحلة لاحقة يتم التوسع فيما أعطي في المرحلة السابقة وتقديم معلومات أخرى حول المفهوم وهكذا، إلى أن يكتمل تقديم المفهوم بأركانه الخمسة وباللغة الرياضية الدقيقة، وهو ما يتماشى مع مفهوم المنهاج الحلزوني. وفي المرحلة التي تستوجب على المتعلم إدراك الخواص الجوهرية للمفهوم، توجه نشاطات المتعلم نحو إدراك الخاصة أو الخواص الجوهرية من خلال عدد محدود من أمثلة المفهوم وعندما يكون

المتعلم قد جرد تلك الخاصة أو الخواص يقوم بعدها بتعميمها لتشمل جميع أمثلة المفهوم. (سلامة، 2008، ص18)

2-4-النظرة الحديثة للهندســـة:

ليست الهندســـة بمعزل عن الرياضيات، فينظر إليها رياضياً على أنها:

1. طريقة في التفكير وإثارته حيث يقوم التلميذ بعمل استنتاجات خاصة به من خلال ضم أشكال هندسية لبعضها البعض ومعرفة ماهو الشكل الناتج وكذلك يمكن استخدام أشكال هندسية كرتونية مختلفة يقوم التلميذ بتركيبها على صورة ألعاب رياضية حيث يقوم باستخدام الاكتشاف والاستنتاج في ذلك عند قص أو قطع جزء من شكل هندســـي ليصبح شكل هندسيـــاً آخر ومعرفة ماهو الجزء الذي نقطعه أو نقصه من المستطيل ليصبح شبه منحرف أو كيفية تحويل بعض الأشكال الهندسيـــة إلى أشكال أخرى مثل تحويل المربع إلى مثلثات متطابقة وإستخدام الورق والمقص في ذلك.

2. معرفة منظمة تتسم بالتنظيم والتسلسل فتتكون أصلاً من التعابير غير المعرفة وتصل في النهاية إلى التعميمات والمهارات الرياضية الهندسيـــة.

3. فن تتسم بالجمال والتناسق وتسلسل أفكارها والاستمتاع في عملها و مشاهدتها فرسومات أشكال وعمل مجسماتها يعتبر فناً راقياً ومتميزاً يظهر بوضوح فن الفنان الرياضي في ذلك. (أبو لوم، 2005، ص.19)

2-5- الهندســـة المدرسيـــة:

نتيجة للمستجدات والإتجاهات التي طرأت على الجوانب المختلفة لطبيعة الهندسة وطرائق تدريسها، فقد جعل المجلس القومي لمعلمي الرياضيات في الولايات المتحدة الأمريكية (National Council of Teachers NCTM)، موضوع الهندســـة من أبرز معايير الرياضيات المدرسية وتقويمها، وذلك لما تقدمه الهندسة للمتعلمين من معارف وعلاقات وبصيرة هندسيـــة مفيدة في مواقف الحياة اليومية، فضلاً عن كونها السياق المثالي لتنمية مهارات المتعلمين في الاستـــدلال والتبرير وأعمال البرهنـــة سواء أكانت استقرائية أم استنتاجيه. وفي مبادئ ومعايير الرياضيات المدرسية (Principles and Standards for School Mathematics,2000) الصادرة عن المجلس القومي لمعلمي الرياضيات في الولايات المتحدة (NCTM)، والتي اعتمدت في أساسها على معايير عام (1989) وبنت عليه، وحددت ما يتوقع من التلميذ

تعلمه من الرياضيات في المراحل الدراسية المختلفة، وقدمت هذه التوقعات في خمس معايير للمحتوى. وتحددت معايير محتوى الهندسية للصفوف الثلاثة الأخيرة من التعلم الأساسي، كالتالي:

1. تحليل خصائص ومزايا الأشكال الهندسيـــة في بعدين أو ثلاثة أبعاد وتنمية الحجج الرياضية حول العلاقات الهندسيـــة.
2. تحديد مواقع العلاقات المكانية ووصفها باستخدام الهندســـة الإحداثية وأنظمة التمثيل الأخرى.
3. تطبيق التحويلات واستخدام التناظر لتحليل المواقف الرياضية.
4. استخدام الإبصار والاستدلال المكاني والنمذجة الهندسية لحل المشكلات. (سعيد، 2007، ص.169)

لهذا استرعى نموذج (فان هيل) أنظار المشتغلين في التربية في كل من هولندا والإتحاد السوفياتي السابق ودول أوربا، فقامت هذه الدول بمراجعة مناهجها التقليدية في ظل مبادئ هذه النظرية، فأظهرت هذه المراجعة أن هذا النموذج يتمتع بقابلية عالية للتطبيق في نطاق واسع غير محدود. أما في الولايات المتحدة فإنه على الرغم من أن نموذج (فان هيل) قد درس عن كثب خلال العقود الثلاثة الأخيرة من القرن العشرين، فإن أفكار (فان هيل) لم تجسد إلا مؤخراً، حيث أوصى التقرير المعلن من المجلس القومي لمعلمي الرياضيات(NCTM) عام (1989) بادخال نظرية (فان هيل) للممارسة الفعلية ووضعها محل التنفيذ في أمريكا، كما أوصى كذلك الكونجرس العالمي لتعليم الرياضيات (ICME) في مؤتمره السابع المنعقد في عام (1992) بمدينة كيبيك الكندية بتدريس الهندسة في ضوء نظرية (فان هيل). (عبيد، 1993، ص.198)

كما حدد المجلس الوطني لمعلمي الرياضيات في الولايات المتحدة الأمريكية (NCTM) عدداً من المهارات أو الموضوعات التي تعتبر ضرورية لكل مواطن مثقف، وكانت إحداها الهندسة التي تضمنت: (أ) تصنيف الأشكال الهندسية المستوية من خلال صفاتها وخصائصها؛ (ب) حساب محيطات المضلعات؛ (ج) حساب المساحة للأشكال الرباعية والمثلثات والدوائر؛ (د) إدراك مفاهيم التشابه والتطابق والتعرف على الأشكال المتشابهة و المتطابقة (رباب، 2008، ص.28).

وتربط هندسة المرحلة المتوسطة بين الهندسة الملموسة الخاصة بوصف الأشكال الهندسية والعلاقات بينها في المرحلة الإبتدائية وبين هندسة المرحلة الثانويـــة التي تختص بالمعالجة التجريبية، ولما كانت البرامج الحديثة في تدريس الرياضيات تولى إهتماماً برياضيات جميع المراحل خاصة المرحلة الإبتدائية

فإن تدريس الهندســـة للمرحلة الإبتدائية طرأ عليه تطوير كبير في المادة والطريقة (نظلة، 1985، ص.217). فمن ناحية ادخلت بعض موضوعات ثبت من التجارب (خاصة تجارب بياجيه) ضرورتها في نمو المفاهيم الهندسية، مثل بعض المفاهيم التوبولوجية (كالمنحنيات المقفولة والمفتوحة، الخارج، الداخل، الحدود، الإتصالية، التوجيه)، والمجسمات (كالصندوق، الهرم، الكرة، الاسطوانة، المخروط، كعكة السميط) والتطابق والقيـــاس. ومن ناحية أخرى فإن طرق التدريس أصبحت تركز على إتاحة الفرصة للتلميذ للإكتشاف واستخدام الطرق التي تسهل التعميم بعد ذلك في المراحل الأخرى، كاعطاء التطابق ليس عن طريق حالات خاصة كتطابق المثلثات مثلاً، ولكن تعميم ذلك لأي شكل هندســـي عن طريق أفكار بسيطة للتحويلات الهندسية، أو في نمو مفاهيم القياس. لهذا يجب على المدرس أن يراعي عند تدريس هندسة المرحلة المتوسطة الإنتقال التدريجي من المعالجة الحدسية لمفاهيم الهندسة إلى المعالجة التجريدية وأن ينمي في التلميذ تقدير وفهم الأسس المنطقية للنظام البديهي وطبيعة البرهان والتعود على استخدام لغة الفئات (نظلة، 1985، ص.218).

2-6-أهداف تدريس الهندســـة:

نهدف من خلال تدريس الهندســـة في جميع المراحل الدراسية إلى مايلي:

1. اكتساب المعلومات المناسبة عن الأشكال الهندسيـــة في المستوى والفراغ وذلك لأهميتها في دراسات أخرى مثل التفاضل والتكامل والمثلثات...الخ، إلى جانب ارتباطها بالعالم الفيزيقي المحيط بالتلميذ . وذلك على مراحل متدرجة تبدأ بالرسم والقياس وعمل النماذج وفحص الحقائق الهندسية بطرق عملية ثم التدرج منها نحو الدراسة الاستنتاجية المبنية على المسلمات والبرهان والاستدلال.

2. تنمية فهم وتذوق الطريقة الاستدلالية كطريقة للتفكير والبرهان مع اكتساب المهارة في تطبيق هذه الطريقة في المواقف الرياضية المختلفة.

3. تشجيع الأصالة والمبادأة والتفكير المثمر عند التلامذة وإتاحة الفرصة لهم لممارسة التفكير الابتكاري من خلال دراسة الهندسة.

4. دراسة أساليب التفكير المختلفة عند معالجة المسائل الهندسية. (البنا، 1994، ص.04)

اضافة إلى ذلك اشار الكونجرس العالمي لتعليم الرياضيات (ICMI) خلال حلقة البحث الإقليمية المغلقة والتي عقدت في فبراير (1986) بالكويت، إلى وجوب أن تهدف مناهج الرياضيات في المرحلتين الإبتدائية والمتوسطة إلى تنمية الفهم للهندسة الحدسية، وتنمية القدرة على استخدام قوانين القياس للأشكال الهندسية ذات البعدين وذات ثلاثة أبعاد، فضلاًعن تنمية القدرة على فهم خواص الأشكال المعروفة المستوية والفراغية (رباب، 2008، ص.29). وقد قسمت الأهداف المنشودة من تدريس الهندسة إلى ثلاثة أهداف رئيسية موضحةً كيفية تحقيقها، وهي:

1. أن يتعرف التلامذة على أهم مفاهيم وخصائص الفراغ الإقليدي وأن يتعلموا كيفية إستعمالها.
2. يتمثل في تزويد التلامذة بالأساليب اللازمة لحل مسائل هندسية وإنجاز تمارين عملية ذات طابع هندسي.
3. تزويد التلامذة بالقدرة على إستعمال بعض الأساليب والطرق الرياضية البسيطة مثل الرسم الفني للتمارين الهندسية وحل التمارين الهندسية بالحساب أو الاستنتاج الاشتقاقي أو التصنيف.

وصنف العزب (1987) كما ورد عن رباب (2008) أهداف تدريس الهندسة إلى عدة محاور (أهداف معرفية، وجدانية، مهارية) تتمثل الأولى منها في:

1. تعرف التلامذة على الأشكال الهندسية وإدراك خصائصها.
2. فهم التلامذة لمعنى التعريفات والفروض والحقائق والنظريات الهندسية.
3. إدراك وفهم التلامذة للمفاهيم والعلاقات الهندسية التي تساعد على مواصلة دراسة الرياضيات.

ولخص (الكسندروف) الهدف من تدريس الهندسة في تنمية ثلاث صفات لدى المتعلمين هي: الخيال الفراغي، والفهم العملي، التفكير المنطقي (رباب، 2008، ص.29).

2-7-أهمية تدريس الهندسة:

تعد الهندسة من أكثر فروع الرياضيات التي تحتاج الى أساليب خاصة في حل مشكلات تعليمها وتعلمها على مختلف مراحل التعليم العام، لأهميتها في تزويد متعلميها بالمهارات الأساسية والضرورية للحياة العملية، كما أنها تتضمن جوانب التعلم المتضمنة في فروع الرياضيات الأخرى. كما أنها من أكثر فروع الرياضيات ارتباطاً ببيئة المتعلم، وحياته اليومية، فضلاًعن ارتباطها الوثيق بموضوعات رياضية

وعلمية أخرى، فهي تساعد المتعلم في تمثيل العالم الذي يعيش فيه ووصفه بطريقة منظمة، وتنتمي طريقته في التفكير وتنظيم البرهان، كما يسهم فهم النماذج الهندسيـــة في تنمية القــدرة على التحليل والاستدلال وحل المشكلات وأنماط التفكير التأملي والتفكير الناقد والتفكير البصري والتفكير الابداعي وهي جميعها أنماط هامة في تعليم المضامين الهندسية المختلفة وتعلمها (عبد القوي، 2007، ص.163). كما تعتبر الهندسة فرعاً من فروع الرياضيات المدرسية، ولها أهميتها في الحياة لما توفره من فرص كبيرة للتلامذة لكي ينظروا ويقارنوا ويقيسوا ويخمنوا وينقدوا الأفكار ويبنوا علاقات جديدة مما يساهم في توفير مجال خصب لتنمية التفكير لديهم (عباس، 2002، ص.16).

وهو ما أشار إليه ابن خلدون، بقوله: " بأن تعلم الهندسة يفيد صاحبها إضاءة في عقله واستقامة في فكره، لأن براهينها كلها بنية الانتظام، جلية الترتيب، لا يكاد الغلط يدخل أقيستها لترتيبها وانتظامها، فيبعد الفكر عن الخطأ ". (عبد القادر، 1997، ص.26)

ويرى هوفر (Hoffer,1981) أن أهمية الهندســـة تتمثل في كونها مادة تنمي البنية العقلية، فهي أيضاً مهارات متعددة يرافقها مستويات تفكير تحكم طبيعة الأداء لهذه المهارات. ولخص خمس مهارات هندسية أساسية هي مهارة نظرية ومهارة لفظية أو وصفية، ومهارة الرسم، ومهارة منطقية، ومهارة تطبيقية. ويضيف ماردیس (Marrades,2000) أن دراسة الهندســـة ترتبط بدراسة كل البنيـــات الأساسيـــة في الرياضيات، لذا فإن طبيعة الهندســـة وطرائق تدريسها ينبغي أن تكون مجالاً خصباً للتدريب على أنماط التفكير المختلفة، كما أن لغة الرياضيات عامة، واللغة والمفاهيم والمصطلحات والرموز الهندسية تتصف بالدقة والإيجاز في التعبير، إذ يؤدي ذلك إلى توجيه تفكير التلميذ في مسارات صحيحة . (سعيد ردمان، 2007، ص.168)

وللهندسة من المميزات من حيث المحتوى والطريقة ما يجعلها مجالاً ممتازاً لتدريب التلامذة على أساليب التفكير السليم وينبع ذلك من خاصتين مهمتين، هما:

1. أن لغة الهندســـة تمتاز عن اللغة العادية بدقة التعبير ووضوحه وإيجازه.

2. أن الهندســـة من حيث الموضوع لها مميزات خاصة في تنمية التفكير، وذلك بالتأكيد على الناحية المنطقية، لوضوح عناصرها وخلوها من العاطفـــة التي تؤثر في استخلاص النتائـــج. (البنا، 1994، ص.04)

ويذكر سيدهو (Sidhu) في هذا السياق، أن من أهم القيم التعليمية التي يمكن أن نحصل عليها من تدريس الهندســـة من خلال إجابتنا عن السؤال لماذا ندرس الهندســـة؟

(أ) تمكين المتعلم من إكتساب كم من الحقائق الهندسيـــة؛ (ب) تطوير ثقافة المتعلم الرياضيـــة وتنمية القدرة على رسم أشكال دقيقة؛ (ج) تعريف المتعلم فائدة الهندســـة في المجالات الأخرى لصناعات البناء وغيرها؛ (د) تنمية أساليب التفكير الموضوعي لدى المتعلم (رباب، 2008، ص.30). ويرى هتفيلد وآخرون (Hatfield et al ,2001) كما ورد عن عبد القوي في موضع آخر، بأن هناك أسباب متعددة تبرز أهمية تعلم الهندسة، ولعل أهم هذه الأسباب أنها تنمي المهارات الفراغية لدى التلامذة. وهذه المهارات تسمى بالتالي قدرات أخرى مثل القدرة على التعليل والقدرة على التخمين ويمكن بذلك اعتبار جميع المهارات اليومية الأخرى ضمن المهارات الفراغية. (2007، ص.163)

كما حدد كوكو ومارك (Cuco and Mark,1998) دور الهندســـة في التعلم العام في محورين، هما:

1. أن الهندســـة تساعد التلامذة على الارتباط بالرياضيات.
2. أن الهندســـة يمكن أن تعتبر محركة لبناء منظور العادات العقلية.

وبالتالي فإن الهندسة تمثل موضوعاً مهماً في مناهجنا الرياضية في جميع مراحل التعليم العام، حيث يواجه المتعلم في سنواته العمرية الأولى معطيات فضائية، لذا فإنه في حاجة إلى تمثيلات حسية متعددة في شتى الميادين. وتسهم التجارب الأساسيـــة في الهندســـة في دعم قـــدرات المتعلم على تصور هذه التمثيلات. ونعني بالتجارب الأساسية في السنوات الأولى من الدراسة هندسة الأجسام الطبيعية و الإدراكات المحسوسة، حيث يحصل على معرفة مناسبة بالأشكال الهندسية المختلفة وخواصها بملاحظتها ومن ثم تسميتها ووصف خواصها وتحليلها، ويدرس الحقائق الهندسية البسيطة، ويكتب مهارة في العمليات الهندسية البسيطة أيضاً، ويقوم في المرحلة التالية حتي يصل في نهاية مرحلة التعليم العام بالدراسة الاستدلالية بدراسة بعض الخواص التي سبق للمتعلم التعامل معها حدسياً أو عملياً، ومن ثم إدراك العلاقات بين خواص هذه الأشكال والمفاهيم الهندسية المختلفة، واستخدام طرق وأساليب البرهان الهندسي في اثبات صحة هذه العلاقات (NCTM , 2000). (عبد القوي، 2007، ص.164)

2-8- مهارات التفكير الهندســـي:

تشير العديد من الدراسات والأبحاث – وخاصة الدراسات المعرفية – إلى أن مهارات التفكير الهندســـي لا يمكن تعلمها لذاتها وبصورة مباشرة، ولكن يتم تعلمها من خلال مواقف التعلم المباشرة، التي يمكن أن تتوافر عند استخدام الأنشطة الإثرائية في التدريس، ومن أهم هذه المهارات، ما يلي:

2-8-1- استخدام علاقات الأشكال الهندسيـــة وأنماطها:

يعد التعرف على علاقات الأشكال الهندسية وأنماطها من أبرز مهارات التفكير التي لا تلقى إهتماماً واسع في المناهج الدراسية، وقد لا يتعرض التلامذة لخبرات تعليمية في مجال التعرف على علاقات الأشكال وأنماطها إلا في حالات إستثنائية أو نادرة لا تشكل أساساً صلباً لتنمية هذه المهارة. أن التعرف على علاقات الأشكال الهندسيـــة وأنماطها مهارة تفكيرية تنطوي على نوع من الاستدلال العقلي لاكتشاف أوجه الشبه والاختلاف أو النسق العام بين مكونات مجموعة من الأشكال الهندسيـــة أو الرسومات أو بين مجموعات من هذه الأشكال والرسومات. وتعتمد فاعلية الاستدلال العقلـــي في اكتشاف نمط العلاقة على مجموعة مهارات تفكيرية أساسية من بينها مهارات التحليل والمقارنة والتصنيف والملاحظة. وتأخذ الأسئلة والتمرينات التي تتناول العلاقات بين الأشكال أنماطاً كثيرة من بينها:

أ. الشكل المختلف: وهو إعطاء مجموعة من الأشكال أو الرسومات التي توجد بينها صفة عامة مشتركة لا نطبق على أحدها، والمطلوب هو إكتشاف الشكل المختلف أو الشاذ.

ب. الأشكال المتناوبة: تضم أسئلة الأشكال المتناوبة سلسلة من الأشكال الهندسيـــة التي رتبت بطريقة معينة، والمطلوب هو إكتشاف نمط الترتيب واختيار الشكل أو الأشكال التي تنسجم مع السياق.

ج. مصفوفات الأشكال: هي عبارة عن مجموعة من الأشكال الهندسيـــة المختلفة التي تتغير مكونات كل مجموعة منها بالحذف أو الاضافة وفق نمط معين تشترك به المجموعات.

د. الأشكال المتناظرة: تتضمن أسئلة الأشكال المتناظرة وجود مجموعتين أو طرفين متقابلين، في كل منهما شكلان يرتبطان بعلاقة تناظر أو تماثل.

هـ. الأشكال المتداخلة: تأتي هذه الأشكال متضمنة متغيراً متداخلا، وقد يكون هذا المتغير على شكل رموز أو حروف أو أرقام. (فتحي، 2007، ص.196-199)

وقد أوصت بعض الدوريات العلمية بضرورة تعليم التفكير الهندســـي من خلال المناهج التعليمية من خلال عدة إتجاهات معاً وليس إتجاهاً واحداً، وضرورة إمكانية تطبيق التلامذة لمهارات التفكير الهندسي التي يكتسبونها في المواقف والمشكلات التي تقابلهم في حياتهم وضرورة تفعيلها عمليا لتحقيق أقصى استفادة منها.(مجدي عزيز، 2005، ص.09)

3- التفكير الهندســـي و نموذج (فان هيل):

يحتل التفكير الهندســـي مكانة مهمة في العمل المدرسي، لإن المناهج المدرسية في معظم دول العالم المتقدم أصبحت تهدف إلى تعلم الفرد طرق التفكير وإكتساب مهاراته حتى يمكنه من التكيف مع المجتمع الذي يعيش فيه وحل المشكلات التي تواجهه في الحياة سواء المدرسية أو خارجها.

لهذا قد تزايد الإهتمام بالبحث في مستويات التفكير الهندســـي لدى المتعلمين من قبل الباحثين والتربويين على المستوى العالمي، وظهرت عدة محاولات في كتابات كل من بسكالو(Pyhskalo)، وفان هيل(Van hiele) وهوفر(Hoffer) :

كما طرح (بسكالو) بعض الأفكار حول مستويات نمو التفكير الهندسي لدى المتعلم تتلخص في مايلي:

1. يرى الأطفال الأشكال الهندسيـــة ككل دون تصنيف أجزائها أو إدراك العلاقات من داخلها أو بين الأشكال المختلفة.

2. بداية التحليل من خلال (الملاحظة، الرسم، التشكيل) حيث يكتشف الطفل المعرفة بالأجزاء الأساسية وخواص الأشكال الهندسيـــة، إلا أنه يستطيع رؤية العلاقات بين الخواص، كما أن فكرة تعريف الشكل بواسطة بعض خواصه لا تكون معروفة بالنسبة له.

3. بداية التفكير الاستدلالـــي، وتتمثل في إدراك بعض خواص الشكل الهندســـي واستخدامها في استنتاج خواص أخرى، وإدراك التعاريف، ويصبح التنظيم المنطقى ممكناً.

4. النظر إلى الهندسة كنظام استدلالي و إدراك المفاهيم المعرفة وغير المعرفة، وبرهنة النظريات وإدراك البناء المنطقي للبرهان.

5. إدراك نظام المسلمـــات كتعرف للبناء المجرد، واستخدام قواعد المنطق الشكلـــي، وإدراك صدق الاستقلال لمجموعة من المسلمـــات في نظام رياضي معين.(مجدي عزيز، 2005، ص.347)

3-1- نموذج (فان هيل) لتنمية التفكير في الهندســة:

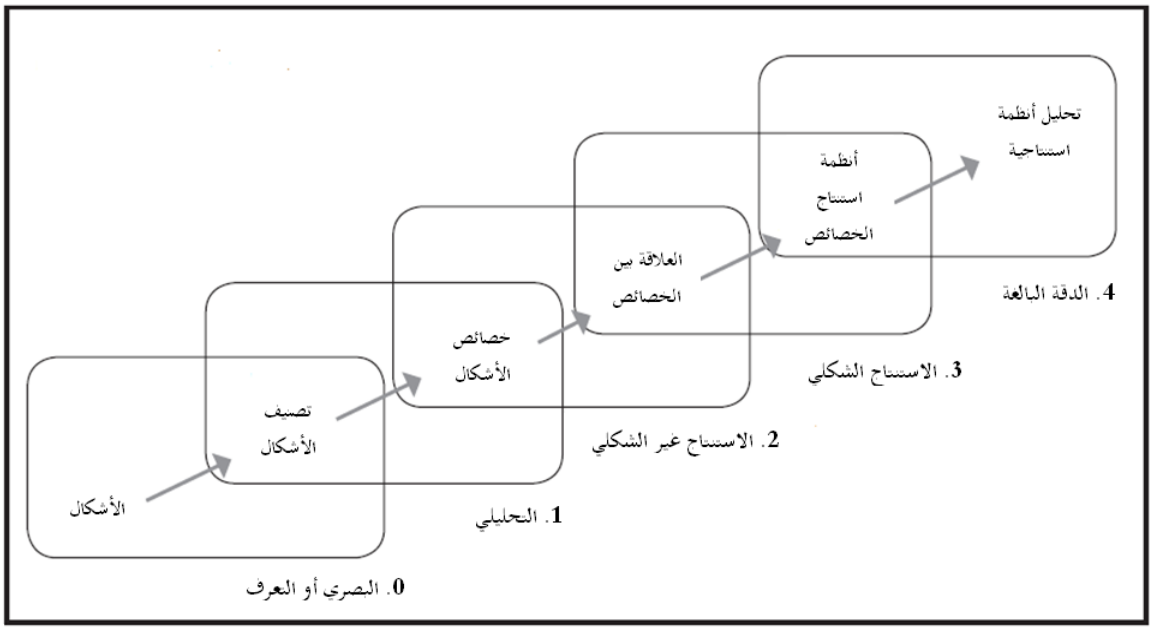
قدم بيير وزوجته دينا فان هيل (Dina Van hiele) نموذجاً للتفكير الهندســي أواخر الخمسينات (1957) أثناء دراستهم للدكتوراه في جامعة Utrecht في هولندا . كان موضوع دراسة بيير "دور الحدس في تعليم الهندسة"، أما دينا فقد كان موضوعها "تعليم الهندسة" (Didactics in geometry)، وكان الإثنين يعملان معلمان للرياضيات في المدارس بهولندا. توفيت (دينا) بعد دراستها للدكتوراه بوقت قصير، وقام (بيير) بمهمة توضيح النظرية وتطويرها لاحقاً. في عام 1959 نشر (بيير فان هيل) دراسة بعنوان "الهندسة وتفكير الطفل" (The Thought of the Child and Geometry)، ناقش فيها خمسة مستويات تصف تطور التفكير الهندســي في الهندسة.(Fuys,Geddes and Tischler, 1988, p.04)

واستناداً للدراستين السابقتين لهما حول الصعوبات التي يواجهها التلامذة في دراسة الهندسة بهولندا، والتي أشارت إلى أن التفكير الهندســي وتعلم الهندسة يسير في مستويات متتابعة تتضمن نمواً في طرق ونوعية التفكير، حيث لكل مستوى لغته ومصطلحاته التي يمكن استخدامها، وأن تعلم مستوى معين يتطلب تعلماً للمستوى السابق له وأن الانتقال من مستوى لآخر يتطلب وقتاً لنضوجــه قبل الإنتقال إلى المستوى التالي، حيث حذرت النظريــة من أنه إذا كان المتعلم ما زال في مستوى معين وكان التدريس في مستوى أعلى (أي يتطلب مستوى أعلى)، فإنه لا يحدث تعلم ولكن من الممكن فقط أن المتعلم يستظهر المعلومات وقد يتذكرها ولكن دون فهم، ذلك أنه من وجهة نظر (فان هيل) تنمو مستويــات التفكير من خلال التعليم والتعلم وليس من خلال العمليات البيولوجية. ويفترض (فان هيل) أن التفكير الهندســي لدى التلامذة يمر بعدة مستويــات من التعقيد، وعلى الرغم من أن تلك المستويات قد تم اثباتها من خلال الأبحاث، إلا أن الآليات السيكولوجية الأساسية في عملية التفكير على كل مستوى من المستويات لا تزال من الموضوعات الهامة محل الدراسة والبحث (الصادق، 2001، ص.276)، ويتكون نموذج (فان هيل) من ثلاثة محاور أساسية، وهي: (1) مستويات النموذج؛ (2) خصائص النموذج؛ (3) الإنتقال بين المستويات في نموذج.

3-2- مستويات نموذج (فان هيل) للتفكير الهندســي:

حسب نظرية (فان هيل) فان تطور التفكير الهندسي لدى التلامذة يرتقى خلال خمسة مستويــات إبتداء من المستوى الجشطالتي أو التبصر (visualisation) لمستويات أكثر فأكثر اتقاناً للتحليل والتجريد والاستنتاج والدقة .(BURTON, 1999, p05)، فهم يتقدمون بتسلسل هرمــي من خلال خمسة مستويات

للتفكير في الهندسة، حيث استخدمت الأدبيات بنيتين رقميتين مختلفتين لتحديد تلك المستويـــات، الأولى ترقيم المستويات من 0 إلى 4 وهو نظام يماثل النظام الأوربـــي لعد الطوابق في بناية بادئاً بالطابق الأرضي، ثم الأول وبعده الثاني وهكذا. ونظام آخر يرقم من 1 إلى 5 وهو ما استخدمه هوفر و ورزيوب (أحمد، 2014، ص.31). وقد صنف (فان هيل) في البداية مستويات التفكير الهندســـي إلى خمسة مستويات، كما يظهره الشكل رقم (16):



شكل (16) يوضح مستويات التفكير الهندســـي الخمسة حسب نظرية (فان هيل).

(Van de Walle, 2015, p.514)

3-2-1-المستوى (0) التفكير البصري الكلـــي (Visual –Holistic Reasoning Level):

ويتحدد هذا المستوى بملاحظة الصورة أو الشكل الهندســـي ككل دون إدراك لخواصه أو عناصره، فهو يتعرف على شكل المستطيل لأنه يشبه الباب مثلا، وهذا نفس الشكل لأنه يبدو متشابه، وهذا مربع وليس مستطيل لإن المستطيل أطول، ويكون توجه التلامذة بقوة في هذا المستوى نحو الشكل، كما يميلون إلى تحديد المربع فقط من جانبيه الأفقـــي والعمودي (Frank, 2007, p.851)، كما يمكنه أن يتعلم المصطلحات بشرط أن تكون في لغة محسوسة، وبأسلوب ملموس(رباب، 2008، ص.46)، ويمكن للتلميذ في هذا المستوى أن يقوم بالآتي: (أ) يتعرف على هيئة الشكل وهو في أوضاع مختلفة؛ (ب) تحديد بعض الأشكال الهندسيـــة في صورتها الفيزيقية أو المرسومة؛ (ج) رسم بعض الأشكال الهندسية البسيطة؛ (د) يسمي أشكالاً بأسماء عامة (مثلاً المستطيل على شكل باب)؛ (ه) يميز بين الأشكال بحسب مظهرها

ويصفها بالكلام؛ (و) يتعرف على أجزاء شكل؛ (ز) ينظر لكل شكل أضلاعه مستقيمة (مربع مثلاً) وشكل محيطه على شكل منحنيات ولكنه لا يميز بين الأشكال من نفس النوع.(عبيد، 2004، ص.95) (مجدي عزيز، 2005، ص.34)

كما حاول كل من كليمنتس و باتيستا (Clements and Battista, 1992) النظر إلى مستويات (فان هيل) من وجهة نظر (بياجيه) حيث اقترح كلاهما مستوى يسبق المستوى الأول (البصري)، كما ورد عن فرانك (Frank, 2007)، وبأن التلامذة في هذا المستوى يمرون بمستويين فرعيين، وهما:

3-2-1-1- **مستوى ما قبل الإدراك (Pre-recognition):** يكون التلامذة في هذا المستوى غير قادرين على تحديد أسماء العديد من الأشكال المشتركة (المتشابهة). أي لا يميزون الأشكال التي من نفس الفئة، مثلا، قد يميزون بين المربع والدائرة، ولكن لا يستطعون التمييز بين المربع والمثلث.

3-2-1-2- **مستوى الإدراك (Recognition):** يستطيع التلامذة في هذا المستوى تحديد أسماء العديد من الأشكال المشتركة (المتشابهة) بشكل صحيح. (2007, p.851)

3-2-2- **المستوى(1) التفكير التحليلي (Analytic - componential reasoning):**

وفي هذا المستوى يتمكن التلميذ من تحليل الأشكال الهندسية على أساس خصائصها أو مكوناتها والعلاقات المتداخلة بين تلك المكونات، ويتحدد ذلك بملاحظة خواص الأشكال ووصفها دون ربطها ببعضها بعض، سواء على مستوى خواص الشكل الواحد، أو خواص الأشكال المختلفة، ويتميز بتحليل واع لخواص الشكل الهندســي، ورسم الشكل من خلال خواصه، فمثلاً يمكن أن يعرف أن جميع أضلاع المربع متساوية وأن كلاً من قطري المعين هو المنصف العمودي للآخر ولكن يصعب عليه في نفس الوقت إدراك أن كل مربع هو معين. (عبد القوي، 2007، ص.167)، ويمكن للتلميذ في هذا المستوى أن يقوم بالآتي : (أ) يميز بين الأشكال الهندسية بحسب خواصها ومكوناتها؛ (ب) استخدام التعبيرات اللفظية الصحيحة الدالة على خصائص الأشكال الهندسية؛ (ج) يتعرف على شكل من خواصه ويختبرها بالقياس؛ (د) مقارنة الأشكال الهندسية طبقا لخصائصها؛ (هـ) يستخدم الخواص في رسم الشكل؛ (و) يعمم خواصاً على مجموعة من الأشكال (المربعات لها 4 أضلاع ، 4 زوايا قائمة)؛ (ز) يحل بعض التمارين على خواص مثل مجموع قياسات زوايا المثلث؛ (ح) لا يرى حاجة لإثبات صحة الخواص التي يدركها فيكفي القياس مثلاً. (مجدي عزيز، 2005، ص.348) (عبيد، 2004، ص.95)

ويرى باتيستا (Battista) في هذا كما ورد عن فرانك (Frank,2007) بأن التلامذة في هذا المستوى يمرون بثلاثة مستويات فرعية في المستوى الثاني (التحليلي)، وهي:

3-2-2-1- **التفكيـــر البصري غير الشكلـــي(Visual-informal componential reasoning):** وصف التلامذة لأجزاء وخصائص الأشكال غير رسمي وغير دقيق، فهم لا يملكون المفاهيم أو التصورات الشكلية التي تمكنهم من تحديد خصائص الشكل بدقة.

3-2-2-2- **التفكير الشكلـــي وغير الشكلـــي غير الكافـــي (informal and insufficient-formal tial reasoning componen):** هنا يبدأ التلامذة في اكتساب المفاهيم الرسمية (الشكلية) التي يمكن استخدامها في الرؤية أو النظر ووصف العلاقات المكانيـــة بين أجزاء الأشكال، و يتم فيه استخدام المزج بين الأوصاف الرسمية (الشكلية)، وغير الرسمية (غير الشكلية) لمعرفة الأشكال الهندسية.

3-2-2-3- **التفكير المعتمد على الخصائص الشكليـــة الكافي (formal property based reasoning sufficient):** التلامذة في هذه المرحلة يستخدمون صراحة وحصراً المفاهيم الهندسية الشكلية (الرسمية) ولغة الوصف وتصور الأشكال بطريقة تهتم بمجموعة من الخصائص لتحديد الأشكال الهندسية.(p.852)

3-2-3- **المستوى(2) الاستـــدلال غير الشكلـــي(informal Deductive level) أو الاستنتاج العلائقي (Relational-inferential Property – based reasoning):**

ويطلق عليه أيضا المستوى شبه الاستدلالي، يتحدد فيه بوعي المتعلم العلاقات بين الأشكال الهندسية و خواصها، ويتميز بالقدرة على إعطاء تعريف للشكل الهندسي، وإيجاد علاقات بين خواص الشكل الواحد والأشكال المختلفة.(عبد القوي، 2007، ص.167)، ويدرك تعاريف مجردة ويستخدم ألفاظا لها طابع منطقي مثل "بعض"، "كل" ويمكنه أن يستدل على خاصية ما بدون حاجة لبرهان منطقي (مجموع الزوايا للشكل الرباعي 360^0 يكفي الاستدلال على ذلك أنه مكون من مثلثيـــن وكل مثلث مجموع زواياه 180^0 ويمكن للطفل في هذا المستوى أن يقوم بالآتي: (أ) تحديد أقل عدد من الخصائص لتعريف شكل هندســـي معين؛ (ب) يرتب أشكالاً هندسية في ضوء خواصها، ولكن دون الاستناد إلى برهان منطقي؛ (ج) يدرك الخصائص التي تكفي لتمييز شكل عن آخر؛ (د) يستنتج بعض خواص العلاقات مثل إذا كان أ=ب، فإن ب=أ و أنه إذا كان أ=ب ، ب=ج فإن أ=ج؛ (هـــ) يصل إلى نتائج من معطيات ويدلل على صحتها بطرق غير شكلية؛ (و) الإتيان ببراهين غير شكلية لإثبات صحة بعض القواعد أو النظريات باستخدام

الرسوم والمواد والأدوات الهندسية؛ (ز) يتابع برهاناً منطقياً ولكنه لا يقيمه بنفسه؛ (ح) يدرك الفرق بين "نظرية" هندسية ومعكوسها ويشرحها بطرق غير شكلية؛ (ط) لا يستطيع الربط بين مجموعة نظريات متعلقة بموضوع واحد. (عبيد، 2004، ص.96) (مجدي عزيز، 2005، ص.348)، كما أن اللغة المستخدمة في هذا المستوى بين المعلم والتلميذ تكون ذات صيغة مجردة، لأنها ترتبط بالعلاقات المنطقية والعلاقات القائمة على المسببات، بالإضافة إلى علاقات أخرى لتراكيب تكون غير واضحة في المستوى التحليلي السابق، مثال : قد يقول التلميذ: " إذا كان الشكل لديه خاصية x ، فإن لديه إذن الخاصية y "(Frank, 2007, p.854)، لهذا فإن الحجج والبراهين عن العلاقات المنطقية بين النظريات في الهندسة تظهر بلغة المستوى الثالث، إذ يفهم التلامذة العلاقات بين النظريات والمسلمات، وتتضح لديهم بعض المفاهيم مثل الشروط الضرورية والكافية (Teppo,1991, p.211) .

ويرى كل من بورو وباتيستا(Borrow and Battista,2000) كما ورد عن فرانك (Frank,2007) بأن التلامذة في هذا المستوى (الاستدلال غير الشكلي) يمرون بأربعة مستويات فرعية، وهي:

3-2-3-1- **العلاقات التجريبية(Empirical relations):** التلامذة في هذا المستوى يستخدمون الأدلة التجريبية في الاستنتاج إذا كان للشكل خاصية واحدة أو له خصائص أخرى.

3-2-3-2- **المكون التحليلي(Componential analysis):** من خلال تحليل أنماط أو أنواع الأشكال، يمكن أن يتم بناء مكون أو عنصر واحد في وقت واحد، ويستنتج التلامذة أنه عندما تظهر خاصية واحدة، فلابد أن تظهر أو تحدث الخاصية الأخرى.

3-2-3-3- **الاستدلال المنطقـــي(logical inference):** يستخدم التلامذة استدلالات منطقية حول الخصائص، وتعمل ذهنيا (عقليا) حول بيانات الخاصية وليس صورها.

3-2-3-4- **تصنيف الأشكال هرمياً باستخدام الاستدلال المنطقي(Hierarchical shape classification based on logical inference):** يستخدم التلميذ الاستدلال المنطقي لإعادة تنظيم وتصنيف الأشكال في ترتيب هرمي منطقي. ولا يقتصر التسأل هنا على لماذا المربع هو المستطيل؟ بل يصبح جزء ضروري من المنطق، أي يعطي التلامذة الحجج المنطقية لتبرير تصنيفاتهم الهرمية. فمثلا كل مربع مستطيل ولكن ليس كل مستطيل مربع لإن للمستطيل خصائص إضافية. وأخيراً فإن استخدام التلامذة للمنطق في استخلاص النتائج يمنح لهم وسيلة جديدة لتجميع المعرفة. أي أن المعرفة الجديدة لا يمكن أن تتولد فقط

من خلال وسائل تجريبية بديهية (عملية)، ولكن من خلال الاستدلال المنطقي (العقلي). (Frank, 2007, p.852)

3-2-4- مستوى (3) برهان الاستنتاج الشكلـــي(formal deduction Proof):

ويطلق عليه أيضا مستوى الاستدلال المجرد، وفي هذا المستوى تتحدد القدرة على الاستنتاج من خلال بناء البراهين الرياضية البسيطة، والقدرة على التعليل ضمن خطوات البرهان، وفهم دور كل من التعاريف والمسلمات والنظريات (عبد القوي، 2007، ص.168)، و في هذا المستوى كذلك يستطيع التلميذ أن يفكر نظرياً ويقيم براهين منطقية، ويدرك العلاقات بين الخواص كما يدرك أهمية الاستنتاج ذهنياً واستخلاص نتائج من خواص ومعطيات معطاة، ويمكن للتلميذ في هذا المستوى أن يقوم بالآتي: (أ) يميز بين المصطلحات المعرفة والغير المعرفة، وبين العبارة التي تقبل كمسلمة وتلك اللازم برهنتها (النظرية)؛ (ب) ينتج تتابعاً من العبارة التي يستنتج فيها كل عبارة من السابقة لها وحتى يصل إلى نتيجة مطلوبة أو تساعد في الوصول إلى المطلوب إثباته بالبرهان؛ (ج) يدرك معنى الشرط اللازم والشرط الكافي (مثلاً: إذا كان "أ ب ج د" متوازي أضلاع فإن أب//ج د، كون "أ ب ج د" متوازي أضلاع فهذا الشرط كاف لأن يكون أب//ج د . ولكن إذا كان أب//ج د فهذا شرط لازم ولكنه ليس كاف لأن يكون "أ ب ج د" متوازي أضلاع)؛ (د) يقيم برهاناً يستند إلى المنطق لإثبات صحية قضية ما؛ (هـــ) لا يدرك استقلالية مجموعة من المسلمات أو النظريات؛ (و) يدرك خواص عامة تجمع بين مجموعة من الأشكال أو مجموعة من النظريات؛ (ز) مناقشة نظام المسلمـــات ومدى اتساقـــه واستقلاله وتكامله دون التعرض لكيفية استخدام تلك المفاهيم في بناء الأنظمة الرياضيـــة . (مجدي عزيز، 2005، ص.349) (عبيد، 2004، ص.97)

3-2-5- مستوى (4) الاستدلال المجرد الكامل (Rigor Dedction):

ويسمى أيضا بمستوى الدقة البالغة، وهو أرقى مستويات التفكير الهندســـي في نموذج (فان هيل)، يتميز بقدرة التلميذ على المقارنة بين أنظمة هندسيـــة مختلفة (هندسة اقليدية، هندسة غير اقليدية، هندسة محايدة لا تعتمد على مسلمة التوازي الاقليدية و لا على مسلمات التوازي اللااقليدية)، ويكون التلميذ على وعي وفهم لدور المنطق والطرق المختلفة للبرهان وأسانيده في المنطق الشكلي مثل البرهان المباشر والغير المباشر وذلك الذي يعتمد على رفض التعارض(عبد القوي، 2007، ص.168)، ويمكن للتلميذ في

هذا المستوى أن يقوم، بالآتي: (أ) استنتاج وإثبات بعض النظريات في مختلف أنظمة المسلمات الهندسية (الاقليدية، واللاقليدية)؛ (ب) يبرهن على بعض النظريات بعدة طرق (بحسب طبيعة النظرية)، مثل: البرهان المباشر من المعطيات إلى المطلوب (بما أن.....إذن)، البرهان باستنفاذ جميع الحالات (في المواقف محدودة الامكانات)، رفض النقيض (إذا لم يكن المطلوب صحيحاً فإنه يؤدي إلى تناقض)؛ (ج) يتعامل مع أنظمة هندسية محدودة العناصر (مثل هندسة الأربع فقط)؛ (د) يدرك أهمية استقلال المسلمات التي يبني عليها نظام هندسي معين؛ (هـ) استحداث طرق لحل بعض المشكلات الهندسية أو البرهنة على بعض النظريات الهندسية؛ (و) يمكنه التعامل مع أشكال ثلاثية البعد بمعالجات نظرية.(عبيد، 2004، ص.98) (مجدي عزيز، 2005، ص.349). كما قدم بورتن (BURTON, 1999) تلخيصاً لمستويات التفكير الهندسي لفان هيل، كما جاء في (حسن شوكارة وزملاؤه،2007)، يعرضها الجدول رقم (03):

جدول (03)

يوضح وصف ومثال لكل مستوى من مستويات التفكير الهندسي الخمسة لـ(فان هيل).

المستوى	الوصف	مثال
المستوى1: التبصر	يدرك التلامذة الكائنات الهندسية اعتماداً على مظهرها الفيزيائي ويستدلون بواسطة إعتبارات بصرية بدون الإستعمال الصريح لخصائص هذه الكائنات.	يعتبر التلامذة بأن مربعاً هو مربع اعتماداً على مظهرها العام وما توحي به العين المجردة وأن إرتفاعاً هو إرتفاع لأنه عمودي.
المستوى2: التحليل	يكون المتعلم قادراً على ربط الكائنات الهندسية بخصائصها في حين يستعمل لائحة الخصائص الضرورية للتحديد ووصف هذه الكائنات.	يعتبر التلامذة بأن مربعاً هو مربعاً لأنه يتوفر على 4 أضلاع متقايسة و4 زوايا قائمة وأن الأضلاع المتقابلة متوازية.
المستوى3: التجريد	يكون المتعلم قادراً على ترتيب خصائص الكائنات الهندسية وبناء تعاريف صورية وتمييز الخصائص اللازمة عن الخصائص الكافية لتحديد مفهوم وفهم الاستنتاجات البسيطة في حين أنه لا يفهم البرهان.	يعتبر التلامذة بأن مربعاً هو مربع لأنه مستطيل له 4 أضلاع متقايسة.
المستوى4: الاستنتاج	يكون المتعلم قادراً على فهم دور مختلف عناصر بنية استنتاجية وبناء براهين أو على الأقل فهمها.	يكون المتعلم قادراً على أن يبين بأن متوازي أضلاع له ضلعين متتاليين لهما نفس الطول هو معين.
المستوى5: الدقة	يكون المتعلم قادراً على الإشتغال على أنظمة أكسيوماتية مختلفة ودراسة هندسات متنوعة في غياب نموذج ملموس.	يكون التلامذة قادرين على فهم الهندسات اللاإقليدية

(ص.25)

كما أطلقت (دينا فان هيل) على المستويات من الثاني إلى الخامس مسميات: شكل الهندســة، جوهر الهندســة، التبصر في نظرية الهندســة، التبصر العلمي في الهندســة.

وقد لاقى هذا النموذج إقبالاً منقطع النظير في الولايات المتحدة الامريكية بصفة خاصة وفي أغلب دول أوربا الشرقية والغربية، حيث أجريت البحوث وصممت المناهج التي تتلاءم مع مستويات (فان هيل) للتفكير الهندســي، حيث ورد عن (Fuys,Geddes and Tischler) أن السوفييت راجعو منهاج الهندســة لديهم على أساس مستويات (فان هيل) للتفكير الهندســي.

ولا شك أن نظرية (فان هيل) تأثرت بنظرية (بياجيه) من حيث أن نمو التفكير يسير في مراحل متتالية ولكنها لم تحدد مراحل عمرية معينة مرتبطة بها. والتي تؤكد على أنه لا يمكن العمل في اطار مستوى معين ما لم يكن قد تم نضوج التفكيــر بالنسبة للمستــوى السابق له. (عبد القوي، 2007، ص.170). ويرى (فان هيل) في نظريته أن تطور النشاط المعرفي الهندسي لدى الفرد المتعلم يتميز بخمسة مستويات مختلفة في كم التفكير حيث يقوم بعملية وصف التفكير الهندســي، ويفترض سلسلة من القدرات المعرفية المميزة لسلسلة من مراحل النمو. وقد استخدمت هذه النظرية في التطور المعرفي الهندســي عن طريق العديد من الباحثين بوصفها أساساً للبحث عن تفكير التلامذة في مراحل تعليمية مختلفة. (Pardikaris, 1994, p.41)

وطبقا لنموذج (فان هيل) فإن هذه المستويات تتحقق بمرورها خلال فترات تعلم مختلفة، والإنتقال من مستوى إلى آخر ليس عملية طبيعية بل تأخذ مكانها تحت تأثير برنامج تعلم وتعليم مناسب. ففي ما يتعلق بالمرحلة الإبتدائية الدنيا من السنة الأولى إلى السنة الرابعة إبتدائي، يدعو معيار "الهندسة والقدرة المكانية" إلى تنمية قدرة التلميذ على وصف ورسم وتصنيف الأشكال وتنمية قدرته على التنبؤ بالنتائج من تركيب وتجزئة الأشكال، وكذلك تطوير الحس المكاني لديه. وفيما يتعلق بالسنوات من الخامس إبتدائي إلى الثالثة متوسط، يدعو معيار "الهندسة" التلامذة إلى تنمية قدرتهم على تمييز ووصف ومقارنة الأشكال الهندسية، وتصور وتمثيل الأشكال الهندسية، وإكتشاف تحويلاتها، وتمثيل وحل المسائل باستخدام النماذج الهندسية، وتطبيق الخواص والعلاقات الهندسية. أما على صعيد السنوات من الرابعة متوسط إلى الثالثة ثانوي، يدعو معيار "الهندسة من منظور تركيبي" التلامذة إلى تمثيل المسألة باستخدام النماذج الهندسية وتطبيق خواص الأشكال، وتصنيف الأشكال بدلالة التطابق والتشابه، واستنتاج خواص العلاقات بين الأشكال

الهندسية من افتراضات معطاة، ومن ثم تطوير فهم لنظام المسلمات من خلال تقصي ومقارنة الهندسات المختلفة. (مجدي عزيز، 2005، ص.350)

3-3- خصائص نموذج (فان هيل) لمستويات التفكير الهندســـي:

يتميز نموذج / مستويات (فان هيل) بعدة خصائص (Usiskin,1982):

الخاصية الأولى: الهرمية (hierarchical) أو التتابع الثابت(fixed sequence):

لا يمكن للتلميذ الإنتقال إلى مستوى معين (n) دون المرور بالمستوى السابق له (n-1). وأضاف فان دو وال (Van de Walle, 2015)، بأن (فان هيل) يرى بأن وصول التلامذة إلى مستوى أعلى يكون من خلال التحرك بين المستويات السابقة، فعملية التفكير في مستوى ما، هو نفس الأشياء قيد التفكير في المستوى التالي، كما يوضحه الشكل5، لذا يجب وضع الأشياء قيد التفكير في كل مستوى حيث أن العلاقات بين هذه الأشياء من التفكير يمكن أن تصبح محور المستوى التالي. (p.51)

الخاصية الثانية: التجاور(adjacency):

كل ما يكون ضمني في مستوى التفكير السابق، يصبح صريحاً في المستوى التالي، بمعنى أن التلميذ قد يتعرف على بعض الخصائص، ولكنه لا يعيها أو يدركها في مستوى ما، يدركها ويعيها ويصبح قادراً على التعبير عنها في المستوى التالي، مثال: في المستوى الأول، يتم التعرف على الأشكال من خلال خصائصها، ولكن التلميذ في هذا المستوى من التفكير لا يدرك هذه الخصائص أو يتعرف عليها إلا في المستوى الثاني (wirzup,1976).

الخاصية الثالثة: التمييز(distinetion):

لكل مستوى تفكير لغته الخاصة، ورموزه الخاصة، وشبكات علاقات خاصة تربط هذه الرموز ويرتبط الإنتقال من مستوى لآخر بمدى توسع اللغة، بمعنى ظهور مصطلحات هندسية وتعريفات ورموز جديدة، مثال: العلاقة بين المربع والمستطيل التي لا تظهر ثباتاً في المستوى الأول، وتكون داخلية/ باطنية في المستوى الثاني، حيث يتعرف التلميذ على المربع والمستطيل منفصلين ولا يربط بينهما، أما في المستوى الثالث فيصبح تعريف المربع أنه مستطيل أضلاعه متساوية.

الخاصية الرابعة: الانفصال(separation):

لا يستطيع شخصان في مستوى تفكير مختلفين أن يفهم كل منهما الآخر، وهذا ما يحدث غالباً بين المعلم والتلامذة (wirzup,1976)، فالمعلم يستخدم لغة من مستوى تفكير عالي لا يتمكن التلميذ من فهمها.(Fuys, Ged des, and Tischler, 1988)

الخاصية الخامسة: الاكتساب(attainment):

يمكن لعملية التعليم أن تمكن التلميذ من الانتقال من مستوى الآخر، ولكنها يجب أن تمر بخمس مراحل (تقريباً متسلسلة)، هي: الاستقصاء، التوجيه المباشر، التوضيح أو التفسير، التوجيه الحر، التكامل (Usiskin, 1982)، حيث تساعد هذه الخصائص على العديد من الأمور التي تحدث داخل الأقسام الدراسية في تعليم الهندسة. (الشويخ، 2005، ص.201)

و قد أضاف للخصائص السابقة فان دو وال (Van de Walle, 2015)، ثلاثة خصائص أخرى، هي:

1. **التطور أو النمو(Developmental):**

عندما يكون التدريس أو اللغة في مستوى أعلى من مستوى التلميذ، سيكون التلميذ في موقف المتحدي لفهم المفهوم الذي يجري تطويره، على سبيل المثال: جميع المربعات هي مستطيلات. ليس هناك بناء عقلـــي للعلاقة الفعلية لكيفية إرتباط خصائص المربع مع المستطيل. وبالتالي فإنه لكل مستوى لغته ورموزه الخاصة به.

2. **الاستقلالية عن العمر (Age independent):**

يمكن أن يكون تلميذاً في السنة الثالثة إبتدائي و تلميذاً في المرحلة الثانوية في المستوى الأول (0).

3. **الاعتماد على الخبرة (Experience dependent):**

يتطلب التقدم من خلال الخبارات التعليمية في الهندســـة، حيث يجب على التلامذة الاستكشاف، والنقاش، والتفاعل مع محتوى كل مستوى تالي، مع زيادة الخبرات في المستوى الحالي. (Van de Walle, 2015, p.514)

3-1-3- الانتقال بين المستويات في نموذج (فان هيل):

كان (فان هيل) أكثر تفاؤلاً من (بياجيه) في هذا الشأن، حيث اعتقد أنه يمكن تسريع النمو المعرفي/ الذهني في تعلم الهندســـة من خلال التعليم (Usiskin, 1982)، وليس العمر أو النضج البيولوجي (wirzup,

1976 ; Fuys, Teppo,1991, van Hiele,1986)، ويرى (فان هيل) أن الإنتقال من مستوى الذي يليه يعتمد بشكل كبير على التدريس فضلاًعن العمر أو النضج (Hatfield et al, 2001, p.113)، حيث يقول (van Hiele, 1999):

" أعتقد أن الانتقال يعتمد على التدريس أكثر من اعتماده على العمـــر أو النضـــج، وأن الخبرات التعليمية يمكنها أن تعزز أو تعيق هذا الانتقال أو النمو" (p.311)

وللمعلم دور جوهري وأساسي في انتقال التلامذة من مستوى لآخر، أهمها فكرة المراحل، وقد أعتبر (Usiskin, 1982) ذلك كخاصية خامسة للمستويات، وهي الاكتساب (attainment)، حيث يمكن لعملية التعليم أن تمكن التلميذ من الانتقال من مستوى لآخر من خلال خمس مراحل، هي: الاستقصاء، التوجه المباشر، التوضيح أو التفسير، التوجيه الحر، التكامل. وقدم (فان هيل، 1999) شرحاً لهذه المراحل، كالتالي: (Hiele, Pierr M, 1999, pp.310-316)

الاستقصاء(inquiry): ينبغي أن يبدأ التدريس بتزويد الطفل بمواد تساعده على استكشاف بنى معينة.

التوجيه المباشر(direct Orientation): حيث تقدم المهام بطريقة تظهر فيها خصائص البنى بالتدرج للتلامذة.

التوضيح (explicitation): يقدم المعلم المصطلحات الهندسية ويشجع التلامذة على استخدامها أثناء نقاشاتهم وكتاباتهم في الهندسة.

التوجيه الحر(free-Orientation): يقدم المعلم مهاماً يمكن إنجازها بطرق مختلفة تصقل قدرات التلامذة التي اكتسبوها في المراحل السابقة.

التكامل(integration): حيث تتوفر الفرصة للتلامذة لتجميع ما تعلموه سابقاً، كأن يصمموا أنشطتهم بأنفسهم.

وعلى الرغم من أنه يمكن ربط المستويات بفترات عمرية محددة (بمعنى أن معظم التلامذة يمرون خلال نفس الخبرات التعليمية في نفس الفترة العمرية)، إلا أن (فان هيل) لم يذكر أي جدول لتطور النمو خلال المستويات، وبالأخص فإنه استفسر عن ارتباط مفاهيم النمو بالنضج الحيوي، ويرى أن النمو ليس كافياً لاكتساب المفاهيم، وخلص (فان هيل) أن التلامذة يمرون خلال مستويات التفكير في الهندســـة بشكل مقارب جداً لانتقال التلامذة في مراحل النمو العقلي لبياجيه. (Hatfield et al, 2001, p.113)

ويرى فان هيل (Hiele, 1986)، أن النمو خلال المستويات ليس متداخلاً وإنما فجائياً إلا أن بعض الباحثين عارض ذلك وأيده آخرون. فقد أوجد بيج ودافري (Pegg and Davery,1998) تبريراً للفصل بين المستويات واعتبراه منطقياً إذا أخذ بالنظرة العريضة والمثالية للنظرية، أما برغر و شاونيسي (Burger and Shaughnessy,1986) و فيوز (Fuys et al,1988) فوجد أن التلامذة يمكن أن يتحركوا داخل المستويات جيئة وذهاباً. (طلال، 2003، ص.37)

و يرى (فان هيل)، أن أدوار المعلم خلال هذه المراحل متعددة: التخطيط للمهام، لفت انتباه التلامذة إلى الخصائص الهندسيـــة للأشكال، استخدام مصطلحات (لغة) الهندسة وتشجيع التلامذة على استخدامها، وتشجيع التفسيرات وحل المشكلات التي تتطلب من التلامذة استخدام تفكيرهم التحليلـــي حول الأشكال. كذلك لابد من استخدام مواد محسوسة مثل الأحجيات ولوحة المسامير geoboard، إذ " تبدأ الهندسة باللعب " (van hiele, 1999). بالاضافة إلى هرمية المستويات، ودور المعلم الأساسي في انتقال التلامذة من مستوى لآخر، فإن اللغة المستخدمة في التدريس لها دور جوهري في الانتقال من مستوى لآخر. ويؤكد (فان هيل) كما ورد في (Fuys, Geddes, and Tischler, 1988)، أن من أهم أسباب الفشل في تعلم الهندسة تعود إلى حواجز اللغة، إذ يستخدم المعلم لغة مستوى أعلى من المستوى الذي يتواجد به التلامذة. أو أن التواصل بين المعلم والتلميذ ضعيف بسبب اختلاف المعاني أو الأطر المرجعية لكل منهما (فان هيل) كما ورد في (Gravemeijr, 1998)، مثلاً، معنى المعين بالنسبة للتلميذ يختلف عن معناه بالنسبة للمعلم، فقد يتمكن التلميذ من التعرف على المعين من بعض الخصائص أو مظهره العام، وقد لا يعرف المربع على أنه معين، ولكن بالنسبة للمعلم، فالمعين هو مجموعة من الخصائص والعلاقات: متوازي أضلاع، متساوي الأضلاع، كما أن المربع معين. هذه الأطر المرجعية المختلفة تعيق التواصل بين المعلم والتلامذة رغم استخدامها للغة واحدة (أو تبدو أنها كذلك)، ولكنها مختلفة المعاني والوسيلة الوحيدة التي يراها (فان هيل) كما ورد في (Gravemeijr, 1998)، هي تمكيـــن التلامذة من أن يشكلوا أطراً مرجعية بواسطة العمل المحسوس.

وقد أظهرت العديد من الأبحاث على غرار دراسة كل من (يوسيسكين، 1982؛ الشويخ، 2005) ودراسة مايكل، كلرن (Michiel klaren, 2012) والتي اجريت باستخدام على مستويات التفكير الهندسي لـــ (فان هيل)، أن المستوى الرابع نادراً ما يظهر لدى تلامذة في المراحل الدراسية الإبتدائية والمتوسطة

والثانوية، وأن المستوى الخامس لا يظهر على الإطلاق في المراحل الدراسية الإبتدائية والمتوسطة والثانوية. (Pardikaris, 1994, p.41)، وبالتالي إقتصرت معظم الأبحاث والدراسات والتي أجريت لقياس التفكير الهندســـي على المستويات الأربعة الأولى لنظرية (فان هيل) على غرار دراسة كل من (ويلسون، 1990؛ عزو اسماعيل، 2002)، نظراً لصعوبة رصد المستوى الخامس لدى التلامذة قبل التعليم الجامعي.

3-4- التعديلات على مستويات التفكير الهندســـي لفان هيل:

قام (فان هيل) بتعديلين أساسيين في المستويات، الأول كان في ترقيم المستويات، والثاني كان جوهرياً حيث قدم مجموعة بديلة من المستويات يقول إنها مناسبة لتركيبة الرياضيات ومناسبة للتعامل مع الرياضيين (Pagg and Davey, 1998, p.113).

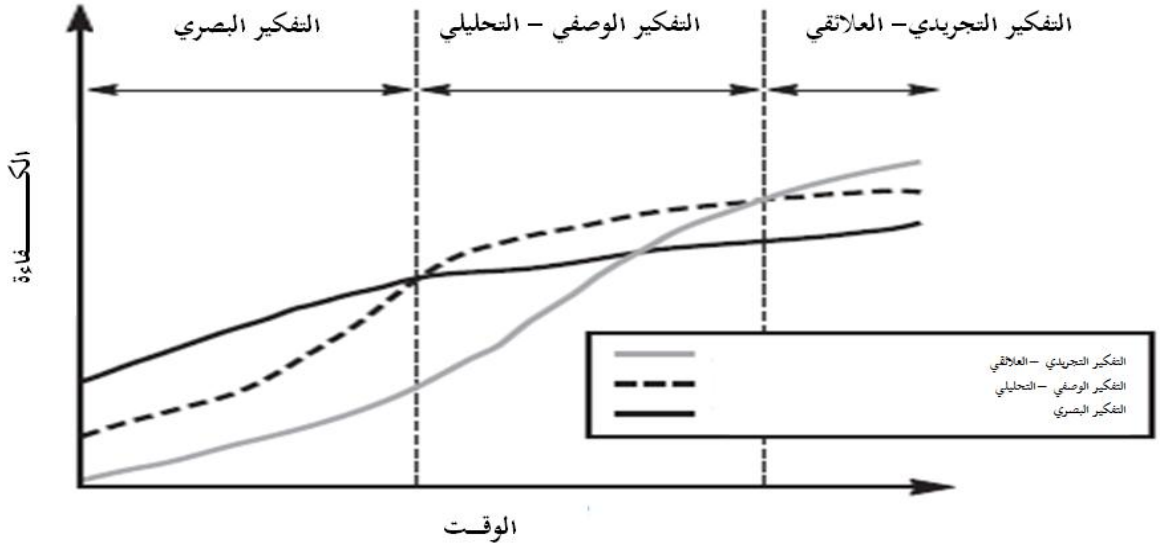
التعديل الأول : مستوى ما قبل التعرف على الشكل الهندسي:

لاحظ الباحثون أنه باستطاعة الأطفال الصغار تشكيل مفاهيم حول الأشكال الهندسية، وذلك قبل دخولهم المدرسة بفترات طويلة مما قاد إلى اجراء بحوث مستفيضة حول الأطفال وممن أشتهر في هذا المجال كلمنت وباتيستا (Clements & Battista) اللذان استنتجا من بحوثهما وجود مستوى يسبق المستوى البصري، وهو ما أسموه بمستوى قبل التعرف على الشكل الهندسي، وفيه يدرك الأطفال مجموعة جزئية من خصائص الأشكال بصرياً، ولكنهم يكونون غير قادرين على تعريف العديد من الأشكال بشكل ثابت مثل الدوائر والمربعات والمثلثات. وبامكان الأطفال في هذا المستوى التمييز بين الأشكال ذات الخطوط المنحنية والأشكال ذات الخطوط المستقيمة ولكنهم لا يميزون داخل كل مجموعة فمثلاً يميزون بين الدائرة والمربع ولكنهم لا يميزون بين المثلث والمربع حسب وجهة نظر (بياجيه) و(انهيلدر).(أحمد، 2014، ص.36)

التعديل الثاني: إعادة تصنيف المستويات:

أكد (فان هيل) إعادة تصنيف مستويات التفكير الهندســـي لاحقاً إلى ثلاثة مستويات فقط، وهي: المستوى البصري ويتميز بالقدرة على ملاحظة الأشكال الهندسية بشكل كلي، والمستوى التحليلي أو الوصفي ويتميز بالقدرة على ملاحظة الأشكال من خلال خواصها الهندسية، والمستوى التجريدي أو النظري ويتميز بالقدرة على استخدام التفكير الاستنتاجي لبرهنة العلاقات الهندسية. (Teppo, 1991, p.210)، وأكد فان هيل نفسه (van hiele, 1999, p.311) على ذلك مع تسمية المستوى الثالث بالاستنتاج

غير الشكلي informal deduction Level. وقد بقيت خصائص النموذج/النظرية كما هي، واحتفظت فكرة المراحل بأهميتها، وهو ما يوضحه الشكل رقم (17): (Teppo, 1991, p.210) (مجدي عزيز، 2005، ص.350)



شكل (17) يوضح منحنى المراحل الثلاثة في اكتساب التفكير الهندسي وفق نظرية (فان هيل) لمستويات التفكير الهندسي. (Frank, 2007, p.849)

يظهر من خلال الشكل رقم (17) أن التلامذة يمرون بثلاثة مراحل لاكتساب التفكير الهندسي وفق نظرية (فان هيل) للتفكير الهندسي، وقد افترض العديد من الباحثين أن هناك فروق في أنماط التفكير الهندسي وخصائصه وتتطوره وفق مستويات (فان هيل) في وقت واحد ومعدلات وأنماط مختلفة أو "موجات" من التفكير المسيطر أو الغالب وفق الكفاءة النسبية للتلامذة في اظهار كل نمط من أنماط التفكير الهندسي، وهذا التطور أو النمو يعتمد على النضج والتعلم معاً(Frank, 2007, p.849). وأن ما يميز كل مرحلة هو المزيج بين هذا الأنماط الثلاثة من التفكير الهندسي، إلا أنه في كل مرحلة تتميز بتفكير سائد أو غالب على بقية أنماط التفكير الهندسي الأخرى، ففي المرحلة الأولى نجد أن التفكير السائد هو التفكير البصري، وفي المرحلة الثانية نجد التفكير السائد أو الغالب هو التفكير الوصفي– التحليلي، وفي المرحلة الثالثة والأخيرة فيغلب عليها التفكير التجريدي – العلائقي أو الاستنتاج الشكلي (النظري).

وعلى الرغم من أنه يمكن ربط المستويات بتغيرات عمرية محددة (بمعنى أن معظم التلامذة يمرون خلال نفس الخبرات التعليمية في نفس الفترة العمرية)، إلا أن (فان هيل) لم يذكر أي جدول لتطور النمو

خلال المستويات، وبالأخص فإنه استفسر عن ارتباط مفاهيم النمو بالنضج الحيوي ويرى أن النمو ليس كافياً لإكتساب المفاهيم، وخلص (فان هيل) إلى أن التلامذة يمرون خلال مستويات التفكير في الهندسة بشكل مقارب جداً لإنتقال التلامذة في مراحل النمو العقلي لبياجيه (Hatfield et al, 2001, p.113).

وعلى الرغم من أن مستويات (فان هيل) تظهر للوهلة الأولى أنها خاصة بالمحتوى التدريسي، إلا أنها في الحقيقة مراحل للنمو المعرفي، حيث أفاد (Hiele,1986) أن "حالة المستويات ليست في المادة المتعلمة ولكن في التفكير الإنساني".(Hiele, 1986, p.41). ويعتمد التفكير لدى (فان هيل) بشكل أساسي على عامل اللغة، وإن لم تكن في تفكير (فان هيل) عندما وصف انتقال الطفل خلال مراحل التعلم، فالإدراك مستحيل بدون عامل اللغة، وبالتالي فإن التدريس أمر ذو حساسية بالغة في نظرية (فان هيل) وربما هذا هو الأمر الذي جعلها أكثر متعة ونجاحاً لدى التربويين، ولكن (Hiele,1999) تراجع بعض الشيء عن عامل اللغة بقوله: " عندما أثرت أن التفكير بدون كلمات ليس تفكيراً ولكن هذا لم يسعد علماء النفس الأمريكيين، إنهم على حق فالتفكير غير اللفظي ذو أهمية خاصة فكل تفكير منطقي له جذور في التفكير غير اللفظي، ويتم إتخاذ العديد من القرارات فقط مع هذا النوع من التفكير". وبين أيضاً أن اللغة تعتبر عاملاً مهماً إبتداءً من المستوى الوصفي. (Hiele, 1999, p.311). كما تضمنت نظرية (فان هيل) ثلاث أقسام رئيسة، وهي:

1. الاستبصار.
2. مستويات التفكير.
3. مراحل الإنتقال في المستويات.

وعديد الأفكار الخاصة بالاستبصار في نظرية (فان هيل) مأخوذة من نظرية الجشطالت ويقول (فان هيل) أن الاستبصار موجود عندما يقوم شخص بموقف جديد على نحو كاف، ويضيف إنه وعلماء الجشطالت يقولون الشيء نفسه لكن بكلمات مختلفة. (Van hiele, 1986, p.24)

اعتمد (فان هيل) (1986)، على علم النفس الجشطالتي لتطوير أفكاره للبنية وأعتقد أن البنيات لها أربعة خصائص مهمة، وهي: (أ) امكانية توسيع البنية؛ (ب) يمكن اعتبار البنية كجزء من بنية دقيقة/ بديعة/ ممتعة؛ (ج) يمكن اعتبار البنية جزءً من بنيات أكثر شمولاً؛ (د) قد تتماثل البنية مع بنية أخرى. حيث يرى بأن الاستبصار هو أن يعمل الشخص بشكل ملائم وعن قصد أو عي، كما أن الاستبصار هو

أن يفهم التلامذة ما يفعلون، ومتى يفعلون، ولماذا يفعلون ذلك، وبالتالي، فإن تصور نموذج التعلم وفق (فان هيل) يكون كالشكل رقم (18):

إدراك البنية ⟵ الاستبصار ⟵ الفهم .

(Pegg and Davey, 1998, p.31)

شكل (18) تصور(فان هيل) لنموذج التعلم.

وقدم كذلك ألن هوفر(Alan Hoffer) نموذجاً للتفكير يتبنى فيه أفكار (فان هيل)، وتكون اطار العمل الذي اقترحه (هوفر) من مصفوفة ذات بعدين: يتضمن البعد الأول المهارات الهندسية الخمس التالية: (أ) المهارات البصرية؛ (ب) المهارات اللفظية؛ (ج) مهارات الرسم؛ (د) المهارات المنطقية؛ (هـ) المهارات التطبيقية. في حين يتضمن البعد الثاني مستويات تطور التفكير في الهندسة لـ(فان هيل) والمتمثلة في: (التعرف، التحليل، الترتيب، الاستنتاج، التجريد و التدقيق).

3-5- النشاط والتفكير من وجهة نظر (بياجيه):

تعد نظرية (بياجيه) للنمو المعرفي القائمة على المنهج الوصفي التحليلي في تناول النمو العقلي المعرفي، مدخلا يتوسط كلٍ من المنحى السيكومتري والمنحى المعرفي في تناول النشاط العقلي المعرفي. (فتحي، 2006، ص.181)، لهذا تتكون المفاهيم الرياضية من وجهة نظر (بياجيه) نتيجة تفكير ذهني ينبثق من الأنشطة التي يقوم بها الطفل ويقصد بالفكر هنا ذلك التدفق من الأفكار الموجهة نحو غاية محددة مثل الإجابة عن سؤال يتطلب وصفاً أو إعطاء معنى، أو تفسير أو مقارنة أو استدلال، ويرى جان بياجيه (Piaget) أن الأفكار تنشأ من الأفعال، وأن المفاهيم الرياضية تنشأ من الأفعال والأنشطة التي يقوم بها الطفل بأشياء وليس من الأشياء ذاتها. ولا شك أن هناك مرحلة تسبق تكوين المفاهيم هي ما أطلق عليها (بياجيه) مرحلة ما قبل المفهوم "pre-concept" وهي التي يبدأ فيها الطفل بأن يفرق بين الشئ وخواصه.

وتشير العديد من الادبيات إلى ما قدمته نظرية (بياجيه)، من أن الأطفال وهم يتفاعلون مع بيئاتهم الإجتماعية والفيزيقية يتقدم نموهم العقلي من خلال أربعة مراحل متتالية، هي: (أ) المرحلة الحسية الحركية (من الميلاد-سنتين)؛ (ب) مرحلة ما قبل العمليات (من 2-7سنة)؛ (ج) مرحلة العمليات

المحسوسة (من 7– 12 سنة)؛ (د) مرحلة العمليات الشكلية أوالمجردة. (من 12 وما بعدها): (عبيد، 2004، ص.89–92)

وحسب النظرية النشوئية لـــ (Inhelder and Piaget) فإن الاستدلال الاستنتاجي يبرز في مرحلة العمليات الصورية. وتظهر المظاهر الأولى للاستدلال عامة عند سن 11– 12 سنة، وتبنى إلى حين استقرارها عند سن 14 – 15 سنة . كما توجد مستويات تتعلق بتطور مفاهيم التحقق والبرهنة.

المستوى الأول: في هذا المستوى تتم عملية تحديد صحة تعبير تجريبياً.

مثال: يكون المتعلم في هذا المستوى قادراً على أن يحدد بأن مجموع قياسات زوايا مثلث يساوي قياس زاوية مستوية باستعمال تقنية التقطيع والتجميع على الرغم من أنه يجهل مفهوم قياس الزوايا ودون أن يفهم لماذا تتكرر الظاهرة.

المستوى الثاني: عند نهاية هذا المستوى تصبح التعميمات الاستنتاجية منتظمة على الرغم من أن التلامذة غير قادرين على بناءً استنتاجات فتفكيرهم يبقى دائماً تجريبي بطبيعته.

مثال: يكون المتعلم قادراً على استنتاج أن مجموع قياسات زوايا مثلث تساوي 180° إنطلاقاً من سلسلة من المثلثات.

المستوى الثالث: في هذا المستوى يستعمل التلامذة استدلالات منطقية لتبرير اقتراحاتهم على الرغم من أن هذه الاستدلالات لا تركز بالضرورة على رياضيات صورية.

مثال: في هذا المستوى يكون المتعلم قادراً على أن يبين بأن مجموع قياسات زوايا مثلث يساوي 180° باستعمال خاصية الزوايا الخارجية لمثلث.(BURTON, 1999, p03-05) (حسن شوكارة وزملاؤه،2007، ص.23)

كما وضح (بياجيه) أيضا من خلال تجاربه أن المفاهيم الأساسية للهندسة الاقليدية والتي ندرسها منذ أجيال طويلة هي مفاهيم مركبة يصعب على التلميذ استيعابها قبل التمهيد له بمفاهيم أبسط في التوبولوجي مثل السطح المقفول، الجوار، الداخل، الخارج، هذا وقد أظهرت تجارب (بياجيه) كذلك أن خصائص الهندســـة الإسقاطية (المستوي، المجسمة) يسهل على التلميذ استيعابها قبل استيعاب علاقات الهندســـة الإقليدية المستوية، ومن تجارب (بياجيه) الشيقة في نمو مفاهيم الخط المستقيم الذي يعتبر عنصراً أساسياً في الهندســـة الأسقاطية تجربة استخدمت فيها منضدة (مربعة أو مستطيلة أو مستديرة). (نظلة حسن،

1985، ص.168). أما بالنسبة لمفاهيم التشابه فقد وجد (بياجيه) أن مفهوم التشابه لطفل في المرحلة الأولى لا يعني له شيء إلا من الوجهة التوبولوجية (الأشكال المتشابهة المستوية كلها منحنيات بسيطة مقفلة)، أما في المرحلة الثانية ما بين 4.5 إلى 6.5 سنة، فتبدأ فكرة التشابه في الظهور ولكن بصورة سطحية أي يستطيع الطفل أن يفرق بين المربع والمستطيل ولكن إذا سئل الطفل أن يرسم مستطيلاً يشابه نموذجاً لمستطيل فإنه يبالغ في استطالة المستطيل (يزيد من طوله) ولا يكون للطفل هنا فكرة عن محاولة مقارنة الأضلاع أو قياسها، ويظهر مفهوم التشابه بين المثلثات عن طريق الأضلاع المتوازية فقط، أما عن طريق الزوايا فيظهر في مرحلة متأخرة. ويلاحظ أن الأطفال في المرحلة الثالثة يأخذون في الحسبان عاملاً واحد مثل إرتفاع أو طول قاعدة الشكل (مثلث مثلاً) عندما يحاولون مشابهة الشكلين.

كما بين (بياجيه) عن طريق تجاربه المتضمنة الخواص القياسية للفراغ كيف يمر الطفل بمراحل تكون فيه فكرته مشوشة عن علاقات الأطوال و المساحات إلى مراحل تتبلور فيها مفاهيم القياس حتي يستطيع عمل وسائل يقارن بها الأطوال والمساحات والحجوم. أما بالنسبة لنمو مفاهيم المساحات فقد ظهر من تجارب (بياجيه) أن الطفل لا يستطيع فهم حساب المساحات المستطيلة قبل المرحلة الرابعة خاصة في عمر 11 سنة فقد يستطيع الطفل في أوائل المرحلة الثانية أن يقوم بعد وحدات من المربعات الصغيرة التي تحتويها مساحة معينة والتي تحتويها مساحة أخرى، ولكنه لا يستطيع أن يميز أي المساحتين أكبر.(نظلة حسن، 1985، ص.130)

3-6-الاستنتاج حول أفكار (بياجيه) وأفكار (فان هيل):

تقترح نظرية كلٍ من (بياجيه) و (فان هيل)، بأنه ينبغي مرور التلامذة من خلال مستويات تفكير دنيا وصولاً إلى مستويات التفكير العليا غير أن ذلك يستغرق بعض الوقت.

وترى نظرية (فان هيل) أن التعلـــم عملية ليست متصلة وأن هناك قفزات في منحنى التعلم، الأمر الذي يعني وجود مستويات تفكير منفصلة ومختلفة .(Fuys, et al, 1988) ، فيما تعزو نظرية (بياجيه) ذلك إلى النضج أو النمو، وترى كلا النظريتين إنه بإمكان التلامذة الفهم والتوسع في نظام المسلمات فقط عندما يصلون إلى مستويات التفكير العليا في الهرمية مع الأخذ بعين الاعتبار أن نظرية (بياجيه) هي نظرية في النمو العقلـــي، ولا تختص بالتعليم أو التعلم بشكل عام، أو بالرياضيات بشكل خاص ولكن تم

توجيه الكثير من الكتابات نحو الافتراضات التي تضمنتها نظريـــة (بياجيه) لتعلم الرياضيات. (أحمد، 2014، ص.37)

لهذا تركز اهتمام (فان هيل) بوصف تطور التفكير الهندســـي لدى التلامذة من خلال المستويات أثناء تعلم الهندسة في سياق المنهاج التعليمي، أما (بياجيه) فقد اهتم بوصف تطور التفكير بشكل عام من تفكير غير ممنهج وغير انعكاسي إلى تفكير تطبيقي وصولاً إلى التفكير المنطقي الاستنتاجي. (Battista and Clements,1995)

ملخص الفصل:

وتأسيساً على ما تقدم يتضح أنه في السنوات الأخيرة و في ميدان التعلم وتعليم الهندسة أصبح هناك اهتماماً متزايداً بدراسة مستويات التفكير الهندســـي عند المتعلمين، ومحاولات تنمية هذا التفكير باعتباره أحد أهم أهداف تدريس الرياضيات بجميع مراحل التعليم العام في معظم الدول، وقد كانت دراسات بياجيه (Biaget) أحد الطلائع الأولى التي اهتمت بدراسة نمو التفكير لدى الطفل. كما يعد نموذج فان هيل (Van hiele Model) من أشهر النماذج الحديثة التي اهتمت بتنمية التفكير الهندســـي لدى المتعلمين، والذي نشأ نتيجة أبحاث كل من (بيير فان هيل) وزوجته (دينا فان هيل) سنة1957 بهولندا، اللذان شرح فيها مستويات نمو التفكير الهندســـي الخمسة، والمكونات المنهجية المناسبة لكل مستوى من مستويات التفكير الهندســـي.

وعليه تم في هذا الفصل، التطرق إلى مفهوم التفكير الهندســـي، وذلك من خلال تجزئته، حيث عالجنا في الجزء الأول من الفصل مفهوم التفكير وماهيته وتعاريفه وأهم نظرياته وتصنيفاته وأدواته، وكذلك إلى دور المفاهيم في التفكير؛ أما الجزء الثاني تطرقنا فيه إلى موضوع الهندســـة، وماهيتها وتعريفها و بنيتها وطبيعة المفاهيم الهندسيـــة والنظرة الحديثة للهندســـة، و منه إلى الهندســـة المدرسية وأهداف وأهمية تدريسها، و أخيراً تم عرض مختصر لأهم مهارت التفكير الهندســـي؛ أما الجزء الثالث من الفصل تناولنا فيه مستويات التفكير الهندســـي ونموذج (فان هيل)، من خلال التعريف بهذا النموذج وتطوره التاريخي و التعريف كذلك بمستوياته الخمسة وخصائصه والتعديلات التي اجريت لاحقاً عليه، وأخيراً المقارنة بين أفكار (جان بياجيه) وأفكار (فان هيل) حول موضوع التفكير الهندســـي.

الفصـــل الخامس

الدراســـــات والبحـــوث السابقـــة

تمهيد:

يزخر ميدان البحوث في مجال القيـــاس والتقويـــم التربوي، بعديد الدراسات التي تناولت موضوع المقارنة باستخدام النظريـــة الكلاسيكيـــة والنظرية الحديثـــة في القياس وبالأخص نموذج راش، في بناء وتدريج وتطوير الاختبارات النفسيـــة والتربويـــة بمختلف أنماطها وتحقيق التفسير الموضوعي لنتائجها، والذي يعتبر من بين أهم المواضيع السائدة حالياً في أوساط علماء النفس القياســـي في تبني مناهج أكثر موضوعية في عملية القياس.

وبعد اطلاع الباحث على ما تيسر منها، تم تصنيفها من حيث موضوعاتها إلى أربعة محاور، ثم استعراض بعض الدراسات والبحوث التي تمت الاستفادة منها مع الإشارة إلى أبرز ملامحها، حيث شمل المحور الأول دراسات تناولت المقارنة بين نموذج راش اللوغاريتمـــي والنظرية الكلاسيكية في القياس، ثم التعليق عليها، وشمل المحور الثاني دراسات تناولت استخدام النظرية الكلاسيكيـــة للقياس في تطوير اختبارات لمستويات التفكير الهندســـي وفق نموذج (فان هيل)، ثم التعليق عليها، و شمل المحور الثالث دراسات تناولت استخدام نظرية الاستجابـــة للمفردة الاختبارية ونموذج راش في بنـــاء وتدريج اختبارات لمستويات التفكير الهندســـي مبنية وفق نموذج (فان هيل)، ثم التعليق عليها، أما المحور الرابع تناول دراسات حول الأداء التفاضلـــي للمفردة الاختبارية باستخدام نظرية الاستجابة للمفردة الاختبارية، ثم التعليق عليها. وبعدما تم عرض المحاور الأربعة، قام الباحث بالتعليق العام عليها، والتي تضمنت جوانب الاتفاق والاختلاف، وبيان الفجوة العلمية التي يعالجها البحث الحالي.

أولاً: دراسات تناولت المقارنة بين نموذج راش اللوغاريتمـــي والنظرية الكلاسيكيـــة في القياس(CCT) ، والتعليق عليها:

1- دراسة كاظم، أمينة محمد (1988ب):

هدفت الدراسة إلى استخدام نموذج راش في بناء اختبار تحصيلي مدخل في علم النفس وتحقيق التفسير الموضوعي لنتائجه، تكون الاختبار من 100 مفردة من نوع الاختبار من متعدد، تغطي ثمان موضوعات في مقرر علم النفس، و ثلاث مستويات من مستويات النشاط العقلي، وهي الحفظ، الفهم، التطبيق. بلغت عينة التقنين 418 طالباً وطالبة من جامعة الكويت. استخدمت برنامج بيكال (BICAL) في

تحليل النتائج وفق نموذج راش. توصلت الدراسة إلى تقديرات لكل من صعوبة المفردات، ولقدرة الأفراد، حيث بلغ معامل ثبات قدرات الأفراد 0.9 لوجيت، بينما بلغ معامل ثبات صعوبة المفردات 0.82 لوجيت، أسفر التحليل على حذف 5 مفردات غير ملائمة للنموذج، وحذف 23 فرداً، والبقاء على 95 مفردة تلائم النموذج وتحقق الاستقلال الموضوعي. توصلت الباحثة إلى الأهمية العلمية لتقارب مستوى الاختبار مع مستوى عينة التحليل للوصول إلى دقة التقديرات المؤدية إلى الموضوعية في عملية التعادل الرأسي في تدريج مفردات الاختبار، بحيث تصبح خصائصه مشتركة وصالحة لجميع خصائص الأفراد، أي تحرر القياس من صعوبة المفردات وقدرة الأفراد.

2- **دراسة عوض الله** (2000):

هدفت الدراسة المقارنة بين النظرية الكلاسيكية ونموذج راش في بناء اختبار للذكاء باستخدام محك التنبؤ بالتحصيل الدراسي من خلال اختبار ذكاء جمعي لطلبة الصف الثالث المتوسط في العراق، يشتمل على أربعة مجالات: القدرة اللفظية 32 مفردة، و48 مفردة للقدرة الاستدلالية، و51 مفردة للقدرة المكانية و40 مفردة للقدرة العددية. طبق على عينة اختيرت بالطريقة العشوائية الطبقية من 400 تلميذاً وتلميذة. استعمل برنامج راسكال (RASCAL) لتحليل البيانـــات وفقاً لأنموذج راش. وقد أشارت نتائـــج التحليل بالطريقة الكلاسيكية إلى حذف 35 مفردة كونها ضعيفة وغير ملاءمة، في حين أشار التحليل وفقاً لأنموذج راش إلى وجود 76 مفردة غير مطابقة للأنموذج، مما يشير إلى اختلاف كبير بين الأسلوبين في حذف المفردات، كما خلص الباحث إلى تفضيل استعمال نموذج راش في انتقاء مفردات الاختبار، بسبب سرعة إجراءات البناء وقلتها فضلاً عن الخصائص الأخرى التي يتمتع بها.

3- **دراسة كريسليب، وتشين تشانص (Crislip and Chin-Chance, 2001):**

استخدمت الدراسة منهجية القيـــاس الكلاسيكي ونموذج راش، على الذين تقدموا لاختبارات ولاية هاواي للقدرات الأساسية (HSTEC)، وهو نوع من الاختبارات المحكيـــة، مكون من اختبارين فرعيين هما: اختبار للمهارات الأساسية، واختبار للمهارات الحياتية. يقيس الاختبار الكلي 16 قدرة أساسية، ويشتمل على 60 مفردة. طبق على عينة حجمها 836 تلميذاً وتلميذة من تلامذة الصف العاشر. أظهرت نتائج التحليل وفق القياس الكلاسيكي باستخدام برنامج ايتمان (ITEMAN)، أن هناك أربعة مفردات 3، 6، 23، 26 سجلت قيم تمييز أقل من 0.29. ونسبة الإجابة الصحيحة على هذه المفردات هي 90% ، مما

يدل على أن القوة التميزية لهذه المفردات كانت سهلة، ويبدو أنها لا تساهم في الهدف من الاختبار (تمييز الكفاءة)، وإن كان هناك حاجة لمفردات سهلـــة في المراجعة النهائية للاختبار. استخدمت القيمة 0.40 كمحك لقيم الارتباط البيسريال النقطي وبه تم حذف المفردات التالية 2، 27، 40، 44 . تراوحت قيم مؤشر التمييز ما بين 0.31 و 0.50. أما في ما يخص صعوبة المفردات التالية: 2، 3، 6، 23، 26، 27، 40، 44 فإنها تحتاج لتنقيح أو الحذف من الاختبار. وباستخدام نموذج راش ببرنامج رسكال، والمؤشر الإحصائي كاي مربع (χ^2)، أظهرت نتائجه أن ما نسبته 37% من مفردات الاختبار غير ملائمة لنموذج راش. كما أن ما نسبته 55% من التلامذة أجابوا على 73% من مفردات الاختبار المعدل باستخدام النظريتين.

4- **دراسة بريتو وآخرون** 2003 **(Prieto et al):**

هدفت الدراسة للمقارنة التجريبية بين نتائج النظرية الكلاسيكية للاختبـــارات و نموذج راش في التخفيض المتوازي لمقياس نوعية الحياة الصحية، وهو بروفيل الصحة لنوتنجهام (NHP). استخدمت الدراسة بيانات 45 دراسة من بين 119 دراسة اجريت على المقياس خلال الفترة مابين سنة 1987 إلى 1995 بمساعدة المركز الطبي الإسباني ومركز البحث بمادلين، تكونت العينة من (9.419) شخص، منهم 4.478 ذكور، و4.908 إناث. استخدمت الدراسة مقياس بروفيل الصحة لنوتنجهام (NHP) المطور في المملكة المتحدة سنة 1987، والمكون من 38 مفردة وستة أبعاد وهي : (الطاقة، الألم، التفاعل العاطفي، النوم، العزلة الاجتماعية، الحركة البدنية). أسفرت النتائج عن مقياس مكون من (NHP20) مفردة، وأربعة أبعاد، باستخدام النظرية الكلاسيكية للاختبـــارات، ومقياس مكون من (NHP22) مفردة، وبعدين فقط، وباستخدام نموذج راش و برنامج بيغ ستابس (BIGSTEPS)، أظهرت النسختان خصائص متشابهة فالمقياس المطور حسب النظرية الكلاسيكية للاختبارات تراوحت قيم العلاقة الارتباطية الكلية بين مفرداته ما بين 0.45 و0.75، بينما حسب نموذج راش فتراوحت ما بين 0.46 و0.68، في حين تراوحت قيم معامل الثبات ما بين 0.82 و0.93، و 0.87 و0.94 على التوالي. وعلى الرغم من الاختلاف في محتوى النسختين (NHP20)، و (NHP22)، إلا أنهما أظهرتا درجة متقاربة ومرتفعة لقيم معاملات الارتباط مابين 0.78 و0.95. وعلى الرغم كذالك من تحقق افتراض أحادية البعد لكلا النسختين، إلا أن أبعاد المقياس

المطور وفق النظرية الكلاسيكية أخفق في تحقيق الملاءمة التي يتطلبها نموذج راش، أي عدم وجود وحدة قياس ثابتة لدرجات المقياس.

5- **دراسة الهاشمي** (2005):

هدفت الدراسة إلى المقارنة بين أسلوبي القيـــاس الكلاسيكي وأنموذج راش في بناء اختبار الذكاء اللغوي لتلامذة المرحلة الابتدائية ببغداد في قدرته على التنبؤ بتحصيلهم في مادة اللغة العربية، مكون من 65 مفردة اشتمل على تسعة مجالات، وهي الإملاء 9 مفردات، و12 مفردة للفهم اللفظي، و14 مفردة للنحو، و10 مفردات لإدراك العلاقات اللفظية، و10 مفردات للطلاقة اللفظية، و5 مفردات للذاكرة اللفظية، و5 مفردات للاستدلال اللفظـــي، طبق على عينة اختيرت بالطريقة العشوائية الطبقيـــة من 504 تلميذاً وتلميذة. استعمل في ذلك برنامج راسكال لتحليل البيانات وفقاً لأنموذج راش، وقد أشارت النتائج إلى أن التحليل بالطريقة الكلاسيكية إلى وجود 7 مفردات تطلب حذفها لانطباق أكثر من محك عليها، في حين أشار التحليل وفقاً لأنموذج راش إلى وجود 26 مفردة غير مطابقة للأنموذج، مما يشير إلى اختلاف كبير بين الأسلوبين ودقة نموذج راش في انتقاء مفردات الاختبار.

6- **دراسة ميمي السيد، إسماعيل** (2007):

هدفت الدراسة إلى التعرف على الخصائص السيكومترية لاختبـــار القدرة العقليـــة باستخدام النظرية الكلاسيكية للقياس ونموذج راش. تكونت العينة من 202 تلميذاً وتلميذة من تلامذة السنة الأولى ثانوي بمحافظة الدقهلية بمصر. استخدمت الدراسة اختبار القدرة العقلية المستوى 15 -17 . إعداد فاروق عبد الفتاح موسى(1984) والمكون من 90 مفردة. ولتحليل المفردات وفق المقاربتيـــن استخدمت البرنامج الإحصائي (Spss)، وبرنامج اكويست (Quest). أظهرت النتائج أن معاملات الارتباط بين المفردات والدرجة الكلية للاختبار قد تراوحت ما بين 0.20 و0.45 باستخدام النظرية الكلاسيكية. حذفت 20 مفردة كانت غير دالـــة إحصائياً. سجل معامل الثبات القيمة 0.85 وفقاً للنظرية الكلاسيكية. وباستخدام نموذج راش تم حذف 6 أشخاص غير مناسبين للنموذج. امتدت صعوبة المفردات ما بين -2.85 و1.95 لوجيت، كما تراوحت قيم قدرات الأفراد ما بين - 5.02 و4.88 لوجيت باستخدام نموذج راش. ارتفع معامل ثبات صعوبة المفردات إلى 0.98، ومعامل ثبات قدرات الأفراد إلى 0.86 باستخدام نموذج راش.

وباستخدام اختبار (ت) بين التطبيق القبلي (قبل التدريج) والبعدي (بعد التدريج)، تم تسجيل القيمة 5.272 لاختبار (ت)، وهي قيمة دالـــة إحصائياً ولصالح التطبيق البعدي في لاختبار القدرة العقلية.

7- **دراسة بوسالم، عبد العزيز** (2008):

استخدمت الدراسة نموذج راش والنظرية الكلاسيكية في القياس في بناء اختبار تحصيلي محكـــي المرجع في القياس النفسي. تكون الاختبار من 46 مفردة، وهو اختبار من متعدد بأربعة بدائل. طبق على عينة استطلاعية 100 طالب وطالبة بمجموع 20 طالباً وطالبة من كل جامعة (قسنطينة، والبليدة، وسطيف، والجزائر، ووهران). للتحقق من خصائصه السيكومترية وفقا للنظرية الكلاسيكية في القياس، أشارت نتائج معامل ثبات سبيرمان وبراون إلى القيمة 0.88، و معامل كيودر- ريتشاردسون (kr_{20}) القيمة 0.86، ومعامل ألفا كرونباخ القيمة 0.81، وإلى صدق محتوى للاختبار في قياس النطاق السلوكي، وكذلك صدق تمييزي بأسلوب المقارنـــة الطرفية . طبق الاختبار في الدراسة الأساسية على عينة مكونة من 916 طالباً و طالبة. وللتحليل وفقاً لنموذج راش استخدمت طريقة بروكس(PROX)، حيث حذفت بهذه الطريقة 45 فرداً و14 بنداً. أظهرت نتائج البحث أن مفردات اختبار القياس النفســـي تقيس متغير أحادي البعد، و تحقق استقلالية عملية القياس من ناحيتين استقلالية أو تحرر عملية القياس في حد ذاتها من قدرات الأفراد، حيث لم يتغير تدريج المفردات برغم تغير مستويات قدرة الطلبة المستخدمين في التدريج، وعدم تأثر كذلك خصائص المفردات بقدرات الأفراد في حالة نموذج راش. كذلك تحرر عملية قيـــاس قدرات الطلبة من أثر صعوبـــة المفردات . كما توصلت الدراسة إلى أن استخدام المدخل الكلاسيكي في القيـــاس لا يؤدي إلى تحقيق قياس موضوعي، حيث اختلفت درجات الأفراد من جامعة إلى أخرى، وتأثر خصائص المفردات بقدرات الأفراد. وأن استخدام نموذج راش اللوغاريتمي في بناء اختبار موضوعي في وحدة القيـــاس النفسي تتدرج مفرداته بوحدة قياس مطلقة قدم لنا تفسيراً موضوعياً لأداء الطلبة على هذا الاختبار.

8- **دراسة علي بن محمد عبد الله زكري** (2009):

هدفت الدراسة للتحقق من الخصائص السيكومترية لاختبار (أوتيس- لينون) للقدرة العقلية وفقراته وفقاً للقياس الكلاسيكي ونموذج راش. بلغ حجم العينة 1515 طالباً من طلاب المرحلة المتوسطة. أظهرت النتائج أن حوالي 70% من فقرات الاختبار تتمتع بمؤشرات صعوبة متوسطة، و حوالي 30% من فقرات

الاختبار يرتفع بها مؤشر الصعوبة، كما أن مستوى الصعوبة يقل عند الانتقال من الصف الأدنى إلى الصف الأعلى منه. وأن حوالي 53% من فقرات الاختبار تتمتع بمؤشرات تمييز جيدة، بينما حوالي 47% من فقرات الاختبار ينخفض بها مؤشر التمييز تدريجياً. وأن قيمة مؤشر الثبات بمعادلة كيودر-ريتشارسون(kr_{20}) بلغ 0.84، في حين بلغت 0.74 بطريقة التجزئة النصفية. وارتباط الفقرات بالدرجة الكلية للاختبار ارتباطاً دالاً إحصائياً عند مستوى 0.01 و 0.05 . كما حققت بيانات اختبار (أوتيس-لينون) للقدرة العقلية ملاءمة مقبولـــة وذلك بتحقيقها لافتراضات نموذج راش. امتدت تقديرات صعوبة الفقرات ما بين -2.28 و1.22 لوجيت. كما امتدت تقديرات القدرة ما بين -4.64 و2.23 لوجيت. والعلاقة بين الدرجة الكلية الخـــام المحتملة على الاختبار والقدرة المقابلة علاقة طردية موجبة. تحقيق الاختبار للصدق من خلال ما يوفره نموذج راش من أحادية البعد في القياس. كما تم إيجاد معايير الرتب المئينية ودرجته التائية، ونسبة الذكاء الانحرافية. توصل الباحث إلى عدم وجود اختلاف في نتائج استخدام المدخلين الكلاسيكـــي ونموذج راش في القياس في تحليل وتدريج اختبار القدرة العقلية.

9- **دراسة ماجد محمد، الخياط** (2012):

هدفت الدراسة إلى التحقق من فاعلية النموذج اللوغاريتمـــي ذي المعلمة الواحدة ومطابقة البيانات لنموذج راش، باستخدام اختبار تيمس (TIMSS) والذي يقيس المهارات الرياضية لمستوى الصف الثامن الأساسي. تم تطبيق مفردات الاختبار على عينة من طلبة الصف الثامن الأساسي بلغ عددها 599 تلميذاً وتلميذة؛ بعد إيجاد الخصائص السيكومترية (الصدق والثبات) لاختبار تيمس، وللإجابة على أسئلة الدراسة تم استخدام البرنامج الإحصائي رام (Rumm,2010)؛ أشارت نتائج الدراسة إلى تمتع الاختبار بالصدق والثبات؛ حيث بلغت قيمة معامل الثبات للاختبار باستخدام معامل كرونباخ ألفا 0.90، أما معامل الصدق بدلالة المحك بين درجات التلامذة على الاختبار ودرجاتهم في مادة الرياضيات فقد بلغت قيمته 0.87. وأشارت النتائج أيضاً إلى مناسبة 25 مفردة من مفردات الاختبار والبالغ عددها 35 مفردة للبيئة المحلية بحيث تمتعت بخصائص إحصائية مناسبة، من حيث معامل صعوبة المفردة.

10- دراسة أيمن سليمان، سلمان القهوجي (2014):

هدفت الدراسة إلى تطوير اختبار في مهارات البحث العلمي لدى طلبة الجامعة، وإستقصاء خصائصه السيكومترية وفقاً للنظرية الكلاسيكية في القياس ونموذج راش. تكون الاختبار من 104 مفردة من اختيار

من متعدد، ضم أربعة مجالات، وهي (المقدمة، والاطار النظري، والاجراءات، والتوثيق). تكونت عينة الدراسة من 997 طالباً وطالبة جامعية، تم اختيارهم بالطريقة المتيسرة. بينت النتائج وفقاً للنظرية الكلاسيكية، بأن معاملات صعوبـــة المفردات تراوحت ما بين 0.38 و0.68، ومعاملات التمييز ما بين 0.18 و0.51، وبأن جميع المموهات امتازت بنسبة جذب مقبولة، حيث لم تسجل أي نسبة لأي مموه أقل من 10%. كما بينت النتائج أن قيم معامل الثبات باستخدام معادلة ألفا كرونباخ للاختبار ككل سجل القيمة 0.90، وأبعاده بالقيم التالية: 0.71؛ 0.60؛ 0.75؛ 0.47 على التوالي، وبطريقة التجزئة النصفية القيمة 0.80 للاختبار ككل و أبعاده بالقيم التالية: 0.65؛ 0.51؛ 0.65؛ 0.33 على التوالي. وللتحقق من صدق البناء تم استخدام المقارنة الطرفية، حيث أشارت نتائج اختبار (ت) لوجود فروق ذات دلالة إحصائية بين الفئة العلياً والفئة الدنياً. بينما بينت نتائج التدريـــج باستخدام نموذج راش بعد التحقق من افتراضات النموذج، باستخدام برنامج (Bilog-mg)، وبرنامج (Winsteps) إلى أن 16 مفردة غير مطابقة للنموذج، وأن أقصى كمية لدالـــة المعلومات بلغت 25.62 عند الدرجة 52. في حين بلغ ثبات المفردات 0.92 و معامل الفصل 4.52، وطبقات قيم الصعوبة 6.36، بينما بلغ ثبات الأفراد 0.91، ومعامل الفصل 3.21، وطبقات قيم القدرة 4.61.

❖ التعليق على الدراسات السابقة التي تناولت المقارنة بين نموذج راش أحادي البارامتر ونظرية القياس الكلاسيكية:

- **بالنسبة للأهداف:**

لقد تباينت أهداف الدراسات السابقة والتي تناولت موضوع المقارنة بين نظريتي القياس الكلاسيكية و الاستجابة للمفردة الاختبارية باستخدام نموذج راش، حيث تناولت دراسة (كاظم، 1988ب) وهي من بين الدراسات الرائدة في الوطن العربي في هذا الموضوع، استخدام نموذج راش في بنـــاء اختبار تحصيلي وتحقيق التفسير الموضوعي لنتائجه، على غرار دراسة (أيمن سليمان، 2014)، في حين أن دراسة (بوسالم، 2008) هدفت إلى بناء اختبار تحصيلي محكي المرجع، ودراسة (كريسليب وتشين تشانص، 2001) التي هدفت إلى استخدام نموذج راش والنظرية الكلاسيكية للقياس في التحقق من فاعلية اختبار محكي المرجع للقدرات الأساسية، في حين أن دراسة (ماجد، 2012) هدفت إلى التحقق من فاعلية كذلك

اختبار تحصيلي للمهارات الرياضية تيمس، على غرار بحثنا الحالي. إلا أن دراسة كل من (ميمي السيد، 2007؛ عوض الله، 2000؛ الهاشمي، 2005؛ علي بن محمد زكري، 2009) تناولت بناء وتدريج اختبارات للذكاء والقدرات العقلية باستخدام نموذج راش والنظرية الكلاسيكية للقياس. كما تناولت دراسة (بريتو وآخرون، 2003) المقارنة التجريبية بين النظريتين الكلاسيكيـــة والحديثة في التخفيض المتوازي لمقياس نوعية الحياة الصحية لنوتنجهام(NHP).

- **بالنسبة للعينات:**

أما في ما يتعلق بالعينة تنوعت أحجام وأنواع العينات المستخدمة في دراسات هذا المحور، من حيث المرحلة العمرية لأفرادها، حيث تناولت دراسة كل من (كاظم، 1988ب؛ بوسالم، 2008؛ أيمن سليمان، 2014) طلبة المرحلة الجامعية، بينما تناولت دراسة كل من (عوض الله، 2000؛ على بن محمد زكري، 2009؛ ماجد، 2012) تلامذة المرحلة المتوسطة أو الاعدادية، في حين تناولت دراسة (الهاشمي، 2005) تلامذة المرحلة الابتدائية، أما الدراسات التي تناولت تلامذة المرحلة الثانوية على غرار بحثنا الحالي نجد دراسة كل من (كريسليب، 2001؛ ميمي السيد، 2007)، في حين نجد أن دراسة (بريتو وآخرون، 2003) الدراسة الوحيدة التي تناولت شرائح عمريـــة مختلفة. أما من ناحية أحجام العينات المستخدمة فنجدها متباينة وذلك تبعا لنوع العينة وطبيعتها، حيث انحصر عدد أفرادها ما بين 202 و1515، حيث كانت أصغر عينة في دراسة (ميمي السيد، 2007) وأكبر عينة نجدها في دراسة (علي بن محمد زكري، 2009)، بمتوسط حسابي يساوي 710 فرداً، وهو عدد قريب من أعداد أفراد عينة بحثنا الحالي، في حين نجد الدراسة الوحيدة التي استخدمت حجم عينة كبيراً جداً 9.419 شخص، هي دراسة (بريتو وآخرون، 2003) والتي اعتمدت على بيانات جاهزة لـــ 45 دراسة سابقة.

- **بالنسبة لأدوات جمع البيانات المستخدمة:**

استخدمت معظم هذه الدراسات اختبارات تحصيلية معيارية المرجع، عدا دراسة كل من (بوسالم، 2008؛ كريسليب، 2001؛ ماجد، 2012) والتي استخدمت اختبارات تحصيلية محكية المرجع، على غرار بحثنا الحالي، في حين استخدمت دراسة كل من (ميمي السيد، 2007؛ عوض الله، 2000؛ الهاشمي،

2005؛ علي بن محمد زكري، 2009) اختبارات للذكاء والقدرات العقلية، كما استخدمت دراسة (بريتو وآخرين، 2003) مقياس لنوعية الحياة الصحية لنوتنجهام.

- **بالنسبة الأساليب والبرامج الإحصائية:**

تنوعت الأساليب والبرامج الإحصائية المستخدمة في هذه الدراسات، حيث استخدمت دراسة (كاظم، 1988ب) في التحليل وفقاً لنموذج راش برنامج بيكال، في حين استخدمت دراسة (عوض الله، 2000؛ كريسليب، 2001؛ الهاشمي، 2005) برنامج راسكال، أما دراسة (علي زكري، 2009) فاستخدم برنامج بيلوغ ماغ، أما دراسة (ماجد، 2012) استخدمت برنامج رام، أما دراسة(علي زكري، 2009) استخدم كل من برنامج رام وبرنامج بيلوغ ماغ، أما دراسة (أيمن سليمان، 2014) والذي استخدم برنامج بيلوغ ماغ، وبرنامج وينستبس على غرار بحثنا الحالي، في حين استخدم بوسالـــم في دراسته (2008) طريقة بروكس(PROX)، أما دراسة (بريتو وآخرون، 2003) فاستخدم برنامج بيغستبس، ودراسة ميمي السيد(2007) والتي استخدمت برنامج أكويست. أما في ما يخص التحليل وفقاً للنظرية الكلاسيكية للقياس نجد أن أغلب الدراســات استخدمت الحزمة الإحصائيــة للعلوم الاجتماعية (SPSS)، ما عدا دراسة (كريسليب، 2001) والتي استخدمت برنامج ايتمان في التحليل.

- **بالنسبة للنتائج:**

اختلفت هذه الدراسات من حيث النتائج المتوصل إليها، حيث توصلت نتائج دراسة (أمينة كاظم، 1988ب) إلى تقديرات لكل من صعوبة المفردات ولقدرات الأفراد وإلى حذف 5 مفردات، و23 فرد غير ملائمين لنموذج راش، أما دراسة (عوض الله، 2000) تم فيها حذف 35 مفردة باستخدام النظرية الكلاسيكية و76 مفردة غير مطابقة لنموذج راش، في حين أن من نتائج دراسة (كريسليب، 2001) حذف 8 مفردات باستخدام النظرية الكلاسيكية، و22 مفرة باستخدام نموذج راش، أما دراسة (الهاشمي، 2005) فتم فيها حذف 7 مفردات باستخدام النظرية الكلاسيكية، و26 مفردة باستخدام نموذج راش، كما أظهرت نتائج دراسة (ميمي السيد، 2007) حذف 20 مفردة باستخدام النظرية الكلاسيكية، و6 مفردات باستخدام نموذج راش، في حين أن دراسة (بوسالم، 2008) تم فيها حذف 45 فرداً و 14 مفردة باستخدام نموذج راش، ولم يتم حذف أي مفردة باستخدام النظرية الكلاسيكية، أما دراسة (ماجد، 2005) تم فيها حذف 10

مفردات باستخدام نموذج راش، وعدم حذف أي مفردة باستخدام النظرية الكلاسيكية، كما بينت نتائج دراسة (أيمن سليمان، 2014) حذف 16 مفردة باستخدام نموذج راش فقط. والدراسة الوحيدة التي لم تحذف فيها أي مفردة خلال مراحل التحليل باستخدام النظريتين الكلاسيكية ونموذج راش هي دراسة (علي محمد زكري، 2009)، في حين أظهرت دراسة (بريتو وآخرون، 2003) أن أبعاد المقيـــاس المبني وفق النظرية الكلاسيكية أخفق في تحقيق الملاءمة التي يتطلبها نموذج راش.

ثانياً: دراسات تناولت استخدام النظرية الكلاسيكيـــة في القياس في تطوير اختبـــارات لمستويات التفكير الهندســـي وفق نموذج فان هيل، والتعليق عليها:

1- **دراسة يوسيسكين زلمان (Usiskin Zalman) مدير مشروع جامعة شيكاغو (1982):**

هدف مشروع جامعة شيكاغو الامريكية، والذي يعتبر من أول الدراسات التي اجريت حول نظرية (فان هيل) لمستويات التفكير الهندسي. شمل عينة مكونة من 2699 تلميذاً وتلميذة من تلامذة السنة السابعة إلى السنة النهائية من التعليم الثانـــوي، ومن 13 مدرسة ثانويـــة تمثل إلى حد واسع الحالة الاجتماعية والاقتصادية للتلامذة، معظم التلامذة من السنة الأولى ثانوي 56% والثانية ثانوي 26%. طبق في بداية المشروع اختبار مستويات التفكير الهندســـي لفان هيل (VHS)، واختبار الدخول للهندســـة (EG) ، أما في نهاية المشروع تم إعادة تطبيق اختبار (VHS)، واختبار التقييم الشامل لبرنامج الهندسة (CAP)، واختبار البرهان الهندســـي (Prf). وقد حاولت الدراسة الإجابة على الأسئلة التالية:

(أ). ما هي مستويات (فان هيل) التي يحققها التلامذة الذين يتعلمون الهندسة؟

(ب). ما هي التغيرات التي تحدث على مستويات (فان هيل) لدى التلامذة بعد دراستهم الهندسة لمدة عام؟

تم بناء اختبار لمستويات التفكير الهندسي، لتمييز مستويات التفكير الهندسي للتلامذة وفقاً لنموذج (فان هيل)، وللتحقق من صدقـــه أجريت مقابلات فردية مع التلامذة بطرح الأسئلة الشفوية عليهم في ثلاثة ولايات مختلفة في دراسات بحثية منفصلة، تم استبعاد أو تعديل المفردات التي لا تتوافق إجابات التلامذة عليها مع مستويات (فان هيل)، وبناءً على هذا الإجراء تم الاحتفاظ بـــ25 مفردة مكونة لاختبار (فان هيل) من الاختيار من المتعدد، حيث خصصت لكل مستوى من المستويات المكونة للنظرية خمسة مفردات. تم تجريب الاختبار على تلامذة أربع مدارس، للتأكد من مناسبة الوقت المحدد له بـــ 35 دقيقة،

والتأكد من توافقها مع مستويـــات (فان هيل). تم التحقق من ثبـــات الاختبـــار باستخدام معادلة كودر- ريتشارسون (Kr_{20})، ومعادلة هورست، والتي سجلت النتائج التالية: 0.31(0.36)؛ 0.44(0.48)؛ 0.49 (0.60)؛ 0.13(0.13)؛ 0.10(0.11) على التوالي (في بداية المشروع)، أما القيم المسجلة في (نهاية المشروع) فكانت كالتالي: 0.39(0.43)؛ 0.55(0.59)؛ 0.56(0.59)؛ 0.30(0.31)؛ 0.26 (0.27) على التوالي. ويرى يوسيسكين أن معاملات الثبات لهذه المستويات، فيما لو كان كل مستوى يحتوي على 25 مفردة، يكون كالتالي: 0.74؛ 0.82؛ 0.88؛ 0.43؛ 0.38 على التوالي (في بداية المشروع)، والقيم التالية: 0.79؛ 0.88؛ 0.88؛ 0.69؛ 0.65 على التوالي (في نهاية المشروع). أظهرت نتائج المشروع، التالي: (أ). تساوي الذكور والإناث في القدرة على تعلم الهندســـة من الحقائق حتى البرهان؛ (ب). أن ما نسبته 88% من التلامذة أمكن تصنيفهم على مستويات (فان هيل) وفق معيار 5/3، وما نسبته 70% وفق معيار 5/4، بمعنى أن النظرية نجحت في تصنيفهم؛ (ج). سهولة تصنيف غالبية تلامذة العينة وهذه نقطة قوية لصالح النظرية. أما نقطة الضعف في النظرية فهي أن تصنيف التلامذة على المستويات يعتمد على المعيار المستخدم حيث يختلف توزيعهم حسب معيار التصحيح 5/3؛ أو 5/4؛ (د). نظرية (فان هيل) إطار نظري يمكن استخدامه لتفسير كيفية تعلم التلامذة للهندســـة.

2- دراسة مخلوف (1994) :

هدفت الدراسة إلى معرفة واقع التفكير الهندسي وفق نموذج (فان هيل) لدى طلاب شعبة الرياضيات في كلية المعلمين بالمدينة المنورة وطلاب السنتين الثالثة والرابعة بكلية التربية بالمدينة المنورة، ومدى تأثير برامج إعدادهم في تنمية هذه المستويات. تكونت عينة الدراسة من105 طالباً وطالبة، صمم الباحث اختبار وفق نموذج (فان هيل)، تكون من 35 سؤالاً، منها 25 سؤال من نوع الاختيار المتعدد، السؤال الخامس من نوع التكملة. وعشرة أسئلة أخرى مسائل، يقوم الطالب برسمها وحلها. بينت نتائج ثبات الاختبار باستخدام كيودر- ريتشارسون(kr_{21}) لكل مستوى من مستويات التفكير الهندسي لفان هيل، كالتالي: 0.11؛ 0.26؛ 0.12؛ 0.66؛ 0.76 على التوالي، والاختبار ككل 0.62. كما استخدم معامل ليفنجستون، وكانت النتائج كالتالي: 0.97؛ 0.95؛ 0.89؛ 0.91؛ 0.85 على التوالي، والاختبار ككل 0.97. أما نتائج الصدق فقد تأكد من مدى اتساق أجزاء الاختبار من خلال حساب معاملات الارتباط بين

درجات الطلاب على كل مستوى من مستويات الاختبار، وبين درجاتهم على الاختبار ككل، وكانت النتائج حسب مستويات (فان هيل)، كالتالي: 0.32؛ 0.48؛ 0.63؛ 0.96؛ 0.58 على التوالي.

3- **دراسة ماكبريد (McBride, 1995):**

هدفت الدراسة إلى تحليل وتقييم أداء التلامذة في البرهان الهندســـي، اعتماداً على نظريات السلوك المعرفي وصنافات بلوم (Bloom,1956)، وجرينو(Greeno,1978)، و فـــان هيل (van hiele,1957). وإجراء مقارنة بين ثلاثة طرق لتصحيح اختبارات البرهان الهندسي، وهي: التنقيط وفق طريقة الاستجابة الثنائية (صح/خطأ)؛ والطريقة الشاملة (متوسطات الدرجات الخام)؛ والطريقة التحليلية (متوسط مجموع تقديرات المحكمين). على عينة مكونة من 241 من تلامذة المرحلة الثانوية بالولايات المتحدة الأمريكية وثلاثة خبراء محكمين لتصحيح اختبارات البرهان الهندســـي. استخدمت الدراسة خمسة اختبارات حول البرهان الهندسي، وهي : (أ) اختبار (فان هيل) لمستويات التفكير الهندسي المعدل (RVHGT) المكون من 24 مفردة ثمنية مفردات لكل مستوى (البصري؛ التحليلي؛ النظري). والذي سجل قيم للثبات بمعامل (جتمان) ما بين 0.80 و0.97، و قيمة ثبات الإعادة بـــ 0.83؛ (ب) اختبار حل المشكلات الهندسيـــة (PST) المكون من 5 مفردات؛ (ج) اختبار البرهان الهندســـي الجديد (NPT) المكون من 6 مفردات؛ (د) اختبار الإتزان والتعديل (EBT) لـــ(Adi,1976) المكون من 15 مفردة خمسة مفردات تمثل كل مستوى أو مرحلة من مراحل النمو المعرفي لبياجي (الملموس (المادي)، الانتقالي، الشكلي أو الرسمي)؛ (هـــ) اختبار البرهان الهندسي (CPT) المطور في دراسة مشروع جامعة شيكاغو (CDASSG,1982). أظهرت النتائج أن نظرية (فان هيل) للتفكير الهندســـي لا تتماشى مع صنافت (بلوم) المدمجة بصنافت (جرينو) في مجالات المعرفة الهندسية وتقييمها.

جدول (04)

يوضح قيم الارتباطات بين المحكـــات المستخدمة في طريقة التصحيح التحليليـــة المطبق على اختبار (CPT) و اختبار (NPT) مع الاختبارات الفرعية لاختبار فان هيل لمستويات التفكير الهندســـي المعدل.

المحكـــات	اختبار البرهان الهندسي (CPT)			اختبار البرهان الهندسي الجديد (NPT)		
أو المعايير	البصري	الوصفي	النظري	البصري	الوصفي	النظري
المحكI: معارف، مصطلحات، علامات، رموز	0.36	0.40	0.31	0.42	0.42	0.29

المحكII : فهم المشكلة.	0.36	0.37	0.31	0.40	0.37	0.27
المحكIII : اقتراح أو افتراض المعارف ذات الصلة.	0.33	0.38	0.29	0.38	0.44	0.32
المحكIV : لديه الخطة لرسم الأشياء اللازمة.	0.37	0.39	0.33	0.44	0.46	0.34
المحكV : النجاح في البرهان أو حل مشكلة.	0.34	0.35	0.31	0.37	0.39	0.35

* كل القيم لمعاملات الارتباط في الجدول دالة عند مستوى (0.01)

تظهر قيم معامل الارتباط في الجدول رقم (04) لاختبار مستويات التفكير الهندســـي المعدل، والذي يصنف الأداء على البرهان الهندسي بأنها قيم ونتيجة غير متوقعة، وبالتالي فإن الاختبار الأخير لا يعتبر مؤشراً قوياً على الأداء وإتقان مهارة البرهان الهندســـي أو مستويات التفكير الهندسي لدى تلامذة المرحلة الثانوية، أما نتائج الثبات فيوضحها الجدول رقم (05):

جدول (05)

يوضح معاملات الثبات للاختبارات المستخدمة بكيودر- ريتشاردسون(KR_{20}) أو ألفا كرونباخ.

الاختبارات	العينة	المفردات	كودر- ريتشاردسونKR_{20} أو ألفا كرونباخ
اختبار الاتزان والتعديل (EBT) ثلاث مراحل لكل اختبار فرعي	239	15 5،5،5	KR_{20}=0.54 (-،0.70،0.48)
اختبار فان هيل للتفكير الهندسي المعدل (RVHGT) وثلاثة اختبارات فرعية.	235	24 8،8،8	KR_{20} = 0.68 (0.47،0.55،0.26)
اختبار البرهان الهندسي (CPT) - محكات الاختبارات الخمس الفرعية - التقييم الشامل - مقياس ثنائي الاستجابة	122	60 12، 12، 12 ، 9 ، 15 6 6	α =0.99 0.95، 0.94 ، 0.96، 0.94، 0.97 0.89 0.86
- اختبار البرهان الهندسي الجديد (NPT) - محكات الاختبارات الخمس الفرعية - التقييم الشامل - مقياس ثنائي الاستجابة	119	60 12، 12، 12 ، 9 ، 15 6 6	α =0.98 0.92، 0.94 ، 0.93، 0.91، 0.95 0.84 0.77

تشير نتائج معامل الثبات ألفا كرونباخ بالجدول رقم (05)، بأنها قيم مرتفعة بالنسبة لاختبارات البرهان الهندســـي التي صححت بالطريقة التحليلية (طريقة تصحيح المحكمين أي الخبراء). وبالتالي

أوصت الدراسة في ضوء نتائجها، بإعادة النظر في اختبار التفكير الهندســـي لفان هيل المعدل، وإلى استخدام طريقة التصحيح التحليلية، إضافة إلى الطريقة الشاملة، و الثنائية من قبل الباحثين والمعلمين في التقييم، وتطوير المناهج الدراسية في نظام البراهين الهندسيـــة.

4- **دراسة رلى يوسف، فهد** (2001):

هدفت الدراسة للتعرف على صعوبات تعلم الهندسة لدى تلامذة الصف الثالث الاعدادي في البحرين وتفسيرها في ضوء مستويات (فان هيل) للتفكير الهندســـي. استخدمت في هذه الدراسة أداتان، وهما: اختبار تشخيصي في الهندسة، واختبار فان هيل للتفكير الهندسي المقتصر على أربع مستويات الأولى من مستويات التفكير الهندسي لفان هيل. تكونت العينة على 491 تلميذاً وتلميذة . توصلت الدراسة إلى: (أ) تمركز أكثر الصعوبات شيوعاً بين التلامذة حول مهارات التحويلات الهندسيـــة، وتطبيق النظريات، وأداء البراهين وكتابتها. وهي مهارات تتطلب مستويات مرتفعة من التفكير الهندســـي لا يمتلكها التلامذة لاعتمادهم المطلق على حفظ القواعد والنظريات الهندسيـــة واستظهارها دون استخدام أي من عمليات التفكير الهندســـي؛ (ب) تمركزت أقل الصعوبات شيوعاً بين التلامذة حول مهارات تذكر التعاريف والمفاهيم، وتطبيق النظريات في حل التمارين، والرسم. وهي مهارات تتطلب مستويات متدنية للتفكير الهندســـي مقارنة بغيرها؛ (ج) مستويات تفكير غالبية التلامذة متدنية جداً في ضوء نظرية (فان هيل)، حيث بلغت نسبة التلامذة في المستويين (صفراً) و(1) من مستويات (فان هيل) حوالي 90%؛ (د) وجود صعوبات في 82% من المهارات التي تم تدريسها في مادة الهندســـة، لأن معظمها يتطلب انتقال التلامذة إلى المستويين الثاني و الثالث من التفكير، بينما طرق تدريسهم عشوائية لا تساعدهم على رفع مستواهم التفكيري والارتقاء به؛ (هـــ) تمركزت الصعوبات التي يعاني منها تلامذة المستويين التفكيريين (صفر) و الأول حول التعاريف والمفاهيم، والخصائص والعلاقات، والتحويلات الهندسيـــة، وإدراك النظريات وتطبيقاتها، والرسم، واستخراج المعطيات والمطلوب، والبرهنة؛ (و) تمركزت الصعوبات التي يعاني منها تلامذة المستوى التفكيري الثاني حول إدراك النظريـــات وتطبيقاتها، والرسم، واستخراج المعطيات والمطلوب، وكتابة البراهين؛ (ز) تمركزت الصعوبات التي يعاني منها تلامذة المستوى التفكيري (صفر) حول إدراك النظريات وتطبيقاتها، البرهنة، وكتابة البراهين.

5- **عفانة، عزو اسماعيل** (2002):

هدفت الدراسة إلى التعرف على مستويات التفكير الهندســي لفان هيل، المتضمنة في موضوعات الهندســة الجزأين الأول والثاني من مناهج الرياضيات المطور للصف السادس الأساسي في فلسطين، والكشف عن مستويات الأداء الفعلي لتلامذة السنة السادسة الأساسي في اختبار التفكير الهندسي لفان هيل. تكونت العينة من 1555 تلميذاً وتلميذة. طبق اختبار (فان هيل) لمستويات التفكير الهندسي عليهم، بعد القيام بترجمته والاستعانة بالاختبارات المتوفرة الأخرى. تم التحقق من صدق اختبار (فان هيل) للتفكير الهندســي من خلال صدق المحكمين، والذين أشارو إلى حذف المستوى الخامس، تم حساب صدق الاتساق الداخلي من خلال معاملات ارتباط كل مستوى من مستويات (فان هيل) والاختبار ككل، والنتائج كالتالي: 0.64؛ 0.72؛ 0.59؛ 0.36 على التوالي. أما الثبات استخدم معامل كودر- ريتشاردسون21 والذي يعتمد على تباين درجات الاختبار، والذي سجل القيم التالية لكل مستوى: 0.76؛ 0.86؛ 0.64؛ 0.51 على التوالي، و القيمة 0.83 للاختبار الكلي.

6- **دراسة الشويخ، جهاد** (2005):

هدفت الدراسة إلى استكشاف أنماط التفكير الهندســي لدى الطلبة الفلسطينيين، وقياس مستويات تفكيرهم الهندســي وفق نظرية (فان هيل)، ومقارنة أدائهم بأداء أقرانهم في الدول الأخرى. تكونت عينة الدراسة من 1240 من تلامذة صفوف السادس والثامن والعاشر الأساسية. ترجم خلال هذه الدراسة اختبار فان هيل للهندســة والمطور من خلال مشروع جامعة شيكاغو من قبل (Usiskin,1982). استخدم المقابلات الفردية للتعرف بعمق على التفكير الهندسي للطلبة، تطلبت المقابلة أداء مهام هندسية، مثل رسم الأشكال وتصنيفها والتعرف عليها بالإضافة إلى لعبة الاستدلال حول الأشكال الهندسية. تم التحقق من صدق الاختبار بعد ترجمته والتواصل مع واضعه (يويسيسكين) الذي أفاده أن الاختبار صادق طالما أن أسئلته تتناسب مع تفكير التلامذة والرياضيات التي تعلموها. وللتحقق من الثبــات استخدمت معاملات ألفا كرونباخ لكل مستوى تفكير حيث بلغت: 0.40؛ 0.09؛ 0.31؛ 0.23 على التوالي. وبنفس الأسلوب الذي اتبعه يويسيكين(1982) بلغ معاملات ثبات هذه المستويات، فيما لو كان كل مستوى يحتوي على 25 سؤالاً باستخدام معادلة سبيرمان-براون، كالتالي: 0.77؛ 0.33؛ 0.69؛ 0.60 على التوالي. أظهرت النتائج أن ما نسبته 30.9% من الطلبة لم يصنفوا بحسب مستويات (فان هيل)، وأن ما نسبته 45.7% فقط

حقق المستوى الأول (الادراك البصري)، و16.8% حققوا المستوى الثاني، و5.6% حققوا المستوى الثالث، و0.7% حققوا المستوى الرابع، و 0.2% حققوا المستوى الخامس من مستويات (فان هيل) للتفكير الهندســـي.

7- **دراسة رباب ابراهيم، الطنة** (2008):

هدفت الدراسة إلى تحليل محتوى منهاج الرياضيات للصف الثامن الأساسي في ضوء مستويات التفكير الهندســـي لفان هايل بمدينة غزة، بلغ حجم العينة 420 تلميذاً وتلميذة. تم فيها بناء اختبار لقياس مستويات التفكير الهندســـي وفقاً لنموذج (فان هيل) مكون من 28 مفردة، أسئلة من اختيار من متعدد وأسئلة مفتوحة، والاكتفاء بالمستويات الأربعة لفان هيل فقط. تم التأكد من صدقه بطريقة صدق المحكمين وصدق الاتساق الداخلي من خلال حساب معامل الارتباط بين كل مفردة من مفردات مستويات الاختبار والدرجة الكلية لمفردات كل مستوى، وكذلك حساب معاملات ارتباط كل درجة كلية لكل مستوى من مستويات الاختبار والمجموع الكلي لدرجات الاختبار، كما تم التأكد من ثبات الاختبار بطريقتين، وهما طريقة التجزئة النصفية، حيث بلغ القيمة 0.88، وطريقة كودر- ريتشارسون21 القيمة 0.89. أظهرت النتائج أن مستوى مهارات التفكير الهندسي لدى تلامذة لم تصل إلى حد الكفاية وهو 60%. ووجود فروق ذات دلالة إحصائية عند مستوى دلالة 0.01 في مستويات التفكير الهندســـي تعزى لمتغير الجنس (ذكور، إناث) ولصالح الإناث.

8- **دراسة لينا فؤاد، جواد** (2011):

هدفت الدراسة إلى الكشف عن مستويات التفكير الهندسي لدى طلبة قسم الرياضيات في كلية التربية بالجامعة المستنصرية، من خلال اختبار تم بناؤه مكون من 50 مفردة من نوع الاختيار من متعدد تقيس مستويات التفكير الهندســـي الأربعة الأولى لفان هيل . طبق الاختبار على عينة مكونة من 180 طالباً وطالبة. تراوحت قيم معاملات تمييز المفردات ما بين 0.33 و0.69، وقيم معاملات صعوبة مفردات الاختبار ما بين 0.31 و0.79. وللدلالة على صدق البناء تم حساب العلاقة الارتباطية بين درجة كل مفردة والدرجة الكلية للاختبار باستخدام معادلة معامل الارتباط الثنائي، تراوحت القيم ما بين 0.32 و0.79، وباستخدام صدق المحك الداخلي، وهو معامل الارتباط بين درجة كل مجال والدرجة الكلية

للاختبار تراوحت القيم ما بين 0.75 و0.90 بمتوسط حسابي 0.83. أما نتائج الثبات وباستخدام معادلة ألفا كرونباخ بلغ القيمة 0.91 للاختبار الكلي، و القيم 0.90، 0.87، 0.84، 0.89 لمستويات (فان هيل) الأربعة على التوالي. وباستخدام درجات القطع 70% فأكثر على المستوى الإدراكي؛ و65% فأكثر على المستوى التحليلي؛ و60% فأكثر على المستوى الترتيبي؛ و55% على المستوى الاستنتاجي. حيث أظهرت النتائج تصنيف 13% من طلبة المرحلة الأولى، و 13% من طلبة المرحلة الثانية، و 23% من طلبة المرحلة الثالثة، إلى أحد المستويات الأربعة (الإدراكي، التحليلي، الترتيبي، الاستنتاجي). وأن 73.3% من طلبة المرحلة الأولى، و 83.3% من طلبة المرحلة الثانية، و 45% من طلبة المرحلة الثالثة كانوا دون المستوى الأول (الإدراكي).

9- **دراسة العطاس، أحمد بن عبد الله** (2014):

هدفت الدراسة إلى إيجاد دلالات الصدق والثبات لاختبار مستويات التفكير الهندسي لفان هيل لدى تلامذة الصف الثاني ثانوي في مدينة مكة المكرمة، وهو الاختبار الذي أعدته جامعة شيكاغو(1982). استخدم الباحث المنهج التحليلي المقارن. طبق النسخة المترجمة للاختبار في دراسة الحربي (1424هـ)، ودراسة الشويخ (2005)، على عينة طبقية عشوائية مكونة من 520 تلميذاً. أشارت نتائج البحث إلى تمتع الاختبار بتدرج واسع من معاملات السهولة والصعوبة بمدى يتراوح بين 0.12 و 0.88، وبمتوسط حسابي بلغ 0.65، ومدى معامل التمييز تراوح بين 0.45 و 1، وبمتوسط حسابي بلغ 0.83. ومدى تباين المفردات تراوح بين 0.1 و 0.25، كما توفرت دلالات صدق جيدة، حيث تراوحت قيم معاملات ارتباط مفردات الاختبار مع الدرجة الكلية للبعد الذي تنتمي إليه المفردة بين 0.44 و0.67، وارتفاع مؤشرات ثبات الإعادة حيث تراوحت قيم معامل الارتباط بيرسون بين التطبيقين من 0.69 و 0.72. وكذلك توفير معايير الأداء لأفراد العينة تمثلت في الرتب المئينية، والدرجات المعيارية، والدرجات التائية.

10- **دراسة مدحت نوري، جليل** (2015):

هدفت الدراسة إلى التعرف على مستويات التفكير الهندسي لدى طلبة المعهد التقني بعقوبة، من خلال تطبيق اختبار تم بناؤه يقيس المستويات الأربعة الأولى لمستويات التفكير الهندسي لفان هيل، والمتكون من 48 مفردة من اختيار من متعدد . تكونت العينة من 140 طالباً وطالبة. أظهرت نتائج صدق البناء من

خلال حساب العلاقة الارتباطية بين درجة كل مفردة والدرجة الكلية للاختبار، والتي تراوحت قيمه ما بين 0.30 و 0.42، و قيم معامل الارتباط بين درجة كل مفردة في كل مجال والدرجة الكلية للمجال تراوحت ما بين 0.30 و 0.40 للمستوى الإدراكي؛ ومابين 0.30 و 0.42 للمستوى التحليلي؛ و ما بين 0.30 و 0.42 للمستوى الترتيبي؛ وما بين 0.30 و 0.42 للمستوى الاستنتاجي. أما نتائج الثبات وباستخدام التجزئة النصفية بلغ معامل الثبات للاختبار الكلي 0.87، و القيم 0.90؛ 0.86؛ 0.88؛ 0.82 لمستويات فان هيل الأربعة على التوالي. وباستخدام محكات النجاح، حقق 70 % فأكثر المستوى الإدراكي؛ و65% فأكثر المستوى التحليلي؛ و60% فأكثر المستوى الترتيبي؛ و55 % على المستوى الاستنتاجي . كما أظهرت النتائج أن 31.4% دون المستوى الإدراكي، و أن 10 % من الطلبة في المستوى الادراكي؛ وأن 5% في المستوى التحليلي؛ و3.5 % في المستوى الترتيبي؛ و3.5 % في المستوى الاستنتاجي.

❖ التعليق على الدراســات السابقة التي تناولت استخدام النظرية الكلاسيكية للقيــاس في تطوير اختبارات لمستويات التفكير الهندســي وفقاً لنموذج (فان هيل):

- **بالنسبة للأهداف:**

تباينت أهداف الدراسات السابقة التي تناولت موضوع استخدام نظرية القياس الكلاسيكية في بناء وتطوير اختبارات لمستويات التفكير الهندسي وفقاً لنموذج (فان هيل)، فمنها من كان هدفها الأول هو التحقق السيكومترية لاختبار مستويات التفكير الهندسي لفان هيل، كدراسة (العطاس، 2014)، أما دراسة كل من (يوسيسكين، 1982؛ نوري، 2015؛ مخلوف، 1994؛ الطنة، 2008؛ لينا فؤاد، 2011) كان هدفها الرئيس التعرف على مستويات وأنماط وصعوبات تعلم التفكير الهندسي وفق نموذج (فان هيل)، ومن خلالها تم بناء اختبارات لقياس مستويات التفكير الهندســي وفقاً لنموذج (فان هيل)، ومنها من استعانت باختبار (فان هيل) لمستويات التفكير الهندسي ذي الشهرة العالمية بعد ترجمته على غرار دراسة كل من (شويخ، 2005؛ ماكبريد، 1995؛ رلى يوسف، 2001؛ عفانة، 2002) لاستكشاف أنماط ومستويات التفكير الهندســي، في حين هدفت دراسة (ماكبريد، 1995) للمقارنة بين طرق تصحيح اختبارات البرهان ومستويات التفكير الهندســي.

- **بالنسبة للعينات:**

أما في ما يتعلق بالعينات تنوعت أحجامها وأنواعها العينات، من حيث المرحلة العمرية لأفرادها، حيث تناولت دراسة كل من (ماكبريد، 1995؛ العطاس، 2014) تلامذة المرحلة الثانوية على غرار بحثنا الحالي، في حين تناولت دراسة (يوسيسكين، 1982؛ رلى يوسف، 2001؛ الشويخ، 2005؛ الطنة، 2008) تلامذة المرحلتين المتوسطة والثانوية، أما دراسة (لينا فؤاد، 2011؛ نوري، 2015؛ مخلوف، 1994) فتناولت طلبة كليات التربية تخصص الرياضيات. أما ناحية أحجام العينات فتراوحت ما بين 105 و 2699، حيث كانت أقل عينة في دراسة (مخلوف، 1994)، وكانت أكبر عينة في دراسة (يوسيسكين، 1982) في مشروع جامعة شيكاغو الأمريكية.

- **بالنسبة لأدوات جمع البيانات المستخدمة:**

تنوعت دراسات هذا المحور في استخدامها لأدوات جمع البيانات، وهي اختبارات لمستويات التفكير الهندسي مبنية ومطورة وفقاً لنموذج (فان هيل)، حيث نجد دراسة كل من (يوسيسكين، 1982؛ العطاس، 2014؛ ماكبريد، 1995؛ رلى يوسف، 2001؛ عفانة، 2002؛ الشويخ، 2005) تم فيها الاستعانة باختبار مستويات التفكير الهندسي لفان هيل المطور بمشروع جامعة شيكاغو، والمكون من 25 مفردة من اختيار من متعدد. في حين نجد دراسة كل من (نوري، 2015؛ مخلوف، 1994؛ الطنة، 2008؛ لينا فؤاد، 2011) تم فيها بناء اختبارات لقياس مستويات التفكير الهندسي وفقاً لنموذج (فان هيل).

- **بالنسبة الأساليب الإحصائية المستخدمة:**

تنوعت الأساليب الإحصائية المستخدمة، حيث استخدمت دراسة كل من (يوسيسكين، 1982؛ كامبريد، 1995؛ عفانة، 2002؛ الشويخ، 2005؛ الطنة، 2008؛ نوري، 2015) معادلة كيودر- ريتشارسون(kr_{20}) و(kr_{21})، وألفا كرونباخ والتجزئة النصفية في التحقق من ثبات الاختبارات المستخدمة، في حين أن دراسة (مخلوف، 1994) استخدمت بالاضافة للأساليب السابقة أسلوب معادلة لفينجستون للتحقق من الثبات و معادلة كيودر- ريتشارسون (kr_{21})، في حين أضافت دراسة (يوسيسكين، 1982) أسلوب معادلة هورست. أما من ناحية الصدق استخدمت دراسة كل من (الغطاس، 2014؛ نوري، 2015؛ لينا فؤاد، 2011؛ الطنة، 2008؛ عفانة، 2002؛ مخلوف، 1994) معامل الارتباط بيرسون وكذلك معامل الارتباط

الثنائي في التحقق من الاتساق الداخلي لمفردات اختبارات مستويات التفكير الهندسي، اضافة إلى صدق المحكمين، كما تفردت دراسة (العطاس، 2014) بإيجاد الرتب الميئنية، والدرجات المعيارية، والتائية لاختبار مستويات التفكير الهندسي لفان هيل الأصلي.

بالنسبة للنتائج:

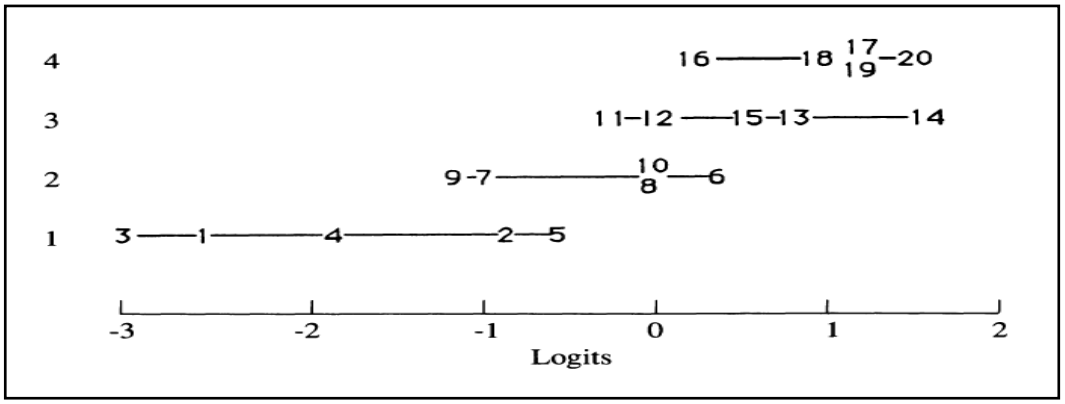
اختلفت نتائج دراسات هذا المحور، حيث توصلت نتائج التحليل وفقاً للنظرية الكلاسيكية للقياس لاختبارات مستويات التفكير الهندسي المبنية والمطورة وفقاً لنموذج (فان هيل)، في دراسة (يوسيسكين، 1982) إلى الاحتفاظ بـ25 مفردة، باستخدام أسلوب صدق المحكمين في التحقق من الصدق، و معادلة كيودر- ريتشارسون(kr_{20}) وهورست، في التحقق من الثبات والتي كانت نتائجه متواضعة نوعاً ما بحسب المستويات الخمسة، ونفس النتائج تقريباً توصلت إليها دراسة (مخلوف، 1994) والتي أضاف فيها ثبات الاختبار الكلي باستخدام معادلة كيودر- ريتشارسون(kr_{21}) والذي سجل القيمة 0.62، وكذلك استخدام معادلة ليفنجستون، أما صدق الاتساق الداخلي كانت قيمه متواضعة نوعاً ما، في حين استخدمت دراسة (كامبريد، 1995) أسلوب معامل (جتمان) وثبات الإعادة للتحقق من الثبات، اضافة لصدق المحكمين، كما أظهرت نتائجها أن نظرية (فان هيل) لمستويات التفكير الهندسي لا تتماشى مع صنافت (بلوم) والمدمجة بصنافت(جرينو) في مجال المعرفة الهندسية وتقييمها، أما دراسة كل من (عفانة، 2002؛ الطنة، 2008؛ نوري، 2015)، توصلت إلى نتائج مقبولة للثبات باستخدام أسلوب كيودر- ريتشارسون (kr_{21})، وكذلك التجزئة النصفية، أما الصدق وبأسلوب الاتساق الداخلي باستخدام معامل الارتباط بيرسون توصلوا كذلك إلى ارتباطات مقبولة بين درجة كل مفردة والدرجة الكلية للاختبار وما بين درجة كل مفردة في كل مجال والدرجة الكلية للمجال. كما توصلت نتائج دراسة (الشويخ، 2005) إلى قيم للثبات ضعيفة على حسب مستويات (فان هيل) الخمسة على غرار الدراسة الأصلية لـ (يوسيسكين، 1982) بمشروع جامعة شيكاغو، وبنفس الأسلوب الذي استخدمه يوسيسكين في دراسته باستخدام طريقة (سبيرمان- براون) لو كان كل مستوى يحتوى على 25 سؤالاً، والتي جاءت قيمه متوسطة نوعاً ما، في حين اضافت دراسة كل من (لينا فؤاد، 2011؛ العطاس، 2014) نتائج لمؤشرات السهولة والصعوبة والتمييز لمفردات اختبارات (فان هيل) لمستويات التفكير الهندسي، حيث تراوح معامل الصعوبة ما بين 0.12 و 0,88، ومعامل للتمييز تراوح ما بين 0.33 و1، كما توفرت دلالات صدق الاتساق الداخلي،

وارتفاع لمؤشر ثبات الإعادة وكذلك معامل ألفاكرونباخ، اضافة إلى توفير معايير الرتب الميئنية والدرجات المعيارية والتائية في دراسة (العطاس، 2014)، وما يؤخذ على دراسة العطاس هو استخدامه للمنهج التحليلي المقارن في دراسة سيكومترية بحتة.

ثالثاً: دراسات تناولت استخدام نظرية الاستجابـــة للمفردة الاختبارية (IRT) ونموذج راش في بناء وتدريج اختبارات مستويات التفكير الهندسي وفقاً لنموذج (فان هيل)، والتعليق عليها:

1- دراسة ويلسون (Wilson, 1990):

والتي موضوعها " قياس تسلسل هندســـة (فان هيل): إعادة التحليل". هدفت الدراسة إلى إعادة تحليل بيانات اختبار مستويات التفكير الهندسي لفان هيل المطور في مشروع جامعة شيكاغو (1982)، وبالتالي فهي امتداد لدراسة (يوسيسكين) ولكن باستخدام النظرية الحديثة للقياس و نموذج راش اللوغاريتمي. استخدمت الدراسة بيانات عينة فرعية مكونة من 655 تلميذاً وتلميذة. بعد تحليل البيانات باستخدام برامج (FORTRAN) وفقاً للنظرية الحديثة للقياس باستخدام نموذج راش. انحصرت القيمة المعيارية المقبولة ما بين -2.0 و 2.0 والتي تتناسب مع مستوى الدلالة 0.05. أظهرت نتائج متوسط الصعوبة لمستويات (فان هيل) الخمسة للتفكير الهندســـي، كالتالي: 1.81؛ 0.32؛ 0.59؛ 1.07 و 0.46 على التوالي، والتي تشير إلى أن المستويات متسلسلة بحسب توقعات (فان هيل)، ما عدا المستوى الخامس، وهذا يتفق مع ملاحظات مصممي الاختبار حول قدرة الاختبـــار على قياس المستوى الخامس (التجريد الصارم)، إلا أنها أظهرت نوعاً من التداخل في المستويات بعد استبعاد المستوى الخامس، كما يبينه الشكل رقم (19):



شكل (19) يوضح نتائج تدريج اختبار (فان هيل) للتفكير الهندســـي لأربعة مستويات. (ويلسون،1990)

كما أشارت النتائج إلى أن ما نسبته 3% من العينة والتي تحددت بحسب معيار 5/4 بأنها غير متسقة ومتسلسلة هرمياً وفق النظرية الكلاسيكية للقياس، فيما حددت ما نسبته 4% من العينة بأنها غير متسقة بنفس المعيار السابق وفق النظرية الحديثة للقياس ونموذج راش، وهكذا يتضح أن التحليل وفق النظريتين الكلاسيكية و نموذج راش يعطي تقريباً نفس النتائج بحسب معيار 5/4. وأوصت الدراسة بإعطاء أهمية كبيرة لنظرية (فان هيل) و اختبار (فان هيل) لمستويات التفكير الهندســي المطور.

2- دراسة سيرو (Serow, 2007):

موضوعها استخدام نموذج راش في اكتساب الاستبصــار في إدراك التلامذة و تضمين المفاهيم الهندسية في فئات. هدفت الدراسة إلى البحث والتعمق في نظرية (فان هيل) والتركيز على تحليل الفهم وتطور المفاهيم الهندسية وتضمينها في فئات. دمجت منهجين كيفي وكمي، فالكيفي تمثل في التحليل و الاستفادة و استخدام نموذج مخرجات التعلم الملاحظ (SOLO) لــ (Biggs and Collis,1982) والذي يشير إلى تطور نمو فهم العلاقات بين الأشكال والعلاقات بين خصائصها. والكمي المتمثل في استخدام نموذج راش للمقارنة بين سبع مهام مختلفة و معقدة في سياق المثلثات والأشكال الرباعية. تم اعتبار أن مستوى ترابط المفاهيم لدى التلامذة باستخدام المعلومات وتضمينها في فئات وهو ما يناسب المستوى الثالث من مستويات (فان هيل) و شرط أساسي لأجل اكتساب التفكير الاستنتاجي الشكلي، والتي يمكن وصفها بأنها تلك التي تتعامل مع العلاقات بين الأشكال والعلاقات بين خصائصها، حيث أن نموذج راش يسمح بتحديد ودراسة المسارات التطورية للنظريات المتأصلة في تطوير المفاهيم الرياضية (Callingham and Bond, 2006). تكونت عينة الدراسة من 24 تلميذاً وتلميذة 12 ذكور و12 إناث أعمارهم تراوحت ما بين 13 و18 سنة، تم اختيارهم من الصفوف الدراسية من الصف 8 إلى الصف 12من مدرستين ثانويتين بأستراليا. استخدمت الدراسة المقابلة لجمع البيانات من 36 استجابة مدتها 1 ساعة، تضم المقابلة برتوكول به سبع مهام أو بنود تركز على العلاقات المعروفة بين الأشكال وبين خصائصها ضمن سياق المثلثات والأشكال الرباعية. استخدمت الدراسة برنامج التحليل الإحصائي بنموذج راش (QUEST) لــ (Adams and Khoo,1993) . أظهرت خريطة بند- فرد للمهام السبعة أن ستة بنود أو مهام وقعت ضمن الحدود المقبولة، في حين أن البند رقم 4 والذي يتعلق بفهم التلامذة للعلاقات بين خصائص المثلث القائم المتساوي الساقين أظهر ميلاً قليلاً إلى اليمين من الحدود المقبولة، والذي يشير إلى أنه هناك عشوائية في الترميز أو

التشفير لدى التلامذة على هذا البند، حيث تم تدريج المهام وفقاً لصعوبتها على النحو التالي: (أ) العلاقات بين أشكال المثلث؛ (ب) العلاقات بين أشكال الرباعيات؛ (ج) العلاقات بين خصائص المثلث المتساوى الأضلاع؛ (د) العلاقات بين خصائص المثلث القائم المتساوي الساقين؛ (هـ) العلاقات بين خصائص المربع؛ (و) العلاقات بين خصائص متوازي الأضلاع؛ (ز) العلاقات بين خصائص المعين. أظهرت النتائج كذلك بأن المهام السبع تتدرج هرمياً واضحاً ومتسلسلاً وفقاً لمستويات نظرية مخرجات التعلم الملاحظ (SOLO) (مستوى (1) وحيد البنية (U)؛ و مستوى (2) المتعدد البنية(M)؛ والمستوى (3) العلائقي (R)) باستخدام نموذج راش، حيث سمحت هذه النتائج بالمقارنة بين صعوبة البنود من خلال فئة الاستجابات بشأن العلاقات بين الأشكال (المستوى الرمزي الملموس أو المادي (CS)) والاستجابات حول البنود المتعلقة بالعلاقات بين خصائص الأشكال (المستوى الاستنتاجي الشكلي (F)).

3- **دراسة مايكل، كلرن (Michiel klaren, 2012):**

موضوعها "خلل التماسك المركزي لدى مضطربي طيف التوحد ومستويات (فان هيل) للتفكير الهندسي". هدفت الدراسة إلى الكشف عن التماسك لدى التلامذة الذين يعانون من اضطرابات طيف التوحد (ASD)، فعادة التلامذة الذين يعانون من التوحد يظهرون خللً في التماسك المركزي (Weal contral coherence)، والذي يخص أسلوب تحليل المعلومات في سياق ملائم أو القدرة على ربط معلومات مع بعض ليصدر معنى وقيمة للشيء، وأن مضطربي طيف التوحد عندهم تركيز زائد على التفاصيل بدل النظرة الشمولية، ويتلخص ذلك في الملاحظة أو الاهتمام أكثر بالتفاصيل الدقيقة بدل التعميم. ولهذا فإن تأثير بما يسمى خلل التماسكيـــة المركزية على عملية التعلم واضح. لهذا انطلق الباحث من افتراض أن التلامذة الذين يعانون من طيف التوحد (ASD) سيجدون صعوبة في الكشف ومعرفة الأشكال الهندسية والسؤال الذي طرح، هو هل أن مستويات صعوبة بنود اختبار (فان هيل) للتفكير الهندســـي لدى الذين يعانون من اضطرابات طيف التوحد هي نفس الصعوبة لدى التلاميذ العاديين؟ وبما أن المستوى الأول لمستويات (فان هيل) للتفكير الهندسي هو مستوى البصري (التعرف أو الإدراك) للشكل الهندسي ككل، لذا فمن المتوقع أن التلامذة الذين يعانون من اضطرابات طيف التوحد سيجدون صعوبة نسبية مع بنود هذا المستوى أكثر من المستوى الثاني (التحليلي) عكس التلامذة العاديين، وبما أن المستوى الثاني التحليلي يهتم بالخصائص والعلاقات إذن فهو النمط المناسب والمفضل لدى التلامذة الذين يعانون اضطرابات طيف

التوحد لذا نجدهم نسبياً ليس لديهم صعوبة في التعامل مع بنود هذا المستوى كغيرهم من العاديين. تكونت العينة من 100 تلميذاً وتلميذة من مدرسة ثانوية بمدينة (إيندهوفن) الهولندية للتلامذة الذين يعانون من اضطرابات طيف التوحد ودرجة ذكاءهم عالي IQ110 ومافوق، تم تحديدهم حسب معايير التشخيص DSM-IV، طبق اختبار (فان هيل) على 81 تلميذاً وتلميذة ،61 تلميذاً وتلميذة بالمرحلة الابتدائية، و 20 تلميذاً وتلميذة بالمرحلة الثانوية. استخدام اختبار مستويات التفكير الهندســـي لفان هيل بعد ترجمته إلى اللغة الهولندية المكونة من 25 سؤال، تم الاستغناء في هذه الدراسة على المستوى الخامس لصعوبة رصده. وباستخدام نموذج راش، أظهرت النتائج أن جميع البنود مطابقة لنموذج راش من المستوى الأول إلى المستوى الرابع، وأن قيم صعوبتها مناسبة تراوحت ما بين −2 و 2 لوغاريتم، باستثناء البند رقم 18 من المستوى الرابع ظهر بأنه أقل صعوبة من البند رقم 13 و14 من المستوى الثالث، والبند رقم 6 من المستوى الثاني، كما تم حذف البند رقم 3 من المستوى الأول والذي أجاب عليه جميع التلامذة والذي قيمته −3.0 لوغريتم، كما أن البند رقم 8 كان أسهل بالنسبة لمجموعة مضطربي طيف التوحد، وأن البنود رقم 16 و19 كانت أكثر صعوبة لهم . أكدت النتائج كذلك بأن البنود في المستويات الأربعة متسلسلة وهرمية سواء عند التلامذة مضطربي طيف التوحد (ASD) أو العاديين، كما أثبتت أن بنود المستوى الأول كانت صعبة نسبياً من المستوى الثاني وهي النتيجة التي كانت متوقعة من البحث.

4- دراسة ستولس (Stols, 2015):

هدفت الدراسة إلى استخدام نموذج راش للتحقق والتقييم التفاعلي لكل من نظرية (فان هيل) للتفكير الهندسي واختبار الاتساق والارتباط لدى اختبار فان هيل للتفكير الهندســـي. والتأكد من أن الاختبار يحقق شروط ومتطلبات القياس لتحقيق الأهداف التعليمية في سياق جديد. وكذلك التحقق من الأداء التفاضلي للبنود لتحديد الفروق في الأداء بين المجموعة قبل الخدمة وبعد الخدمة. تكونت العينة من 244 طالباً وطالبة من طلاب جامعة (بريتوريا) بجنوب إفريقيا من معلمي الرياضيات قبل الخدمة. استخدمت الدراسة اختبار فان هيل للتفكير الهندســـي المبني في مشروع جامعة شيكاغو، وبرنامج (RUMM2030) للتحليل وفقاً نموذج راش . لم تظهر إحصاءات البنود (chi- square) اختلافات بين الدرجات الملاحظة والمتوقعة حسب نموذج راش اللوغاريتمـــي، والتي أكدت إلى حد ما قوة نظرية (فان هيل) في تنمية التفكير الهندســـي. كما سجلت الدراسة إنخفاضاً في مؤشر معامل الفصل لدى الأفراد (PSI) بـــ0.49، ومتوسط

حسابي لموقع الأفراد –0.15، وهي قيمة قريبة من الصفر، أما الانحراف المعياري فسجل 0.70 وهي أقل من المتوقع أي 1 على الرغم من أنها ضمن الحدود المقبولة، كما أظهرت خريطة الفرد – البند رايت (Wright) تكافؤ في توزيع العتبات بين الفرد– البند، إلا أنها أشارت إلى نقطتين وهما: الأولى وهي أن بعض البنود كانت خارج مكانها المفترض حسب نظرية (فان هيل) الهرمية كالبند رقم 19 في (المستوى الرابع) حيث تم معايرته بأن بند صعب للغاية بنسبة لهذه الفئة وهي نتيجة ضد توقعات (فان هيل)، والبند رقم9 و7 (من المستوى الثاني) والبند رقم 14 من (المستوى الثالث) والبند رقم 25 من (المستوى الخامس)، هي أيضاً بنود إلى حد ما خارج مكانها المفترض، و بعد تعديل البند 19 من أثر التخمين ظلت الإشكالية قائمة في البنود 7، 9، 14 حتى بعد تصحيح أثر التخمين. كما كشفت الدراسة أنه هناك مشكلة في بعض بنود الاختبار لأنها تتطلب سوى جانب معرفي محدد من الهندسة بدلاً من اختبار التفكير الهندسي وهي القضية التي اثيرت في دراسات أخرى. و أن هناك نمط تنازلي واضح في الإجابات الصحيحة ضمن المستويات من المستوى الأول من البند 1 الى5 وصولاً إلى المستوى الرابع من البند 16 إلى 20، كما تنبأت بها نظرية (فان هيل). هذا النمط الذي لم يحافظ عليه المستوى الخامس من البند 21 إلى 25.

5– دراسة شونغ ونور(Chong and Nor, 2015):

وموضوعها تقييم معلمي الرياضيات للمرحلة الثانوية قبل الخدمة للقدرة على التفكير الهندسي للأشكال ثنائية الأبعاد. من خلال تدريج اختبار (فان هيل) لمستويات التفكير الهندسي وفقاً لنموذج راش، وكذلك التحقق من مستويات التفكير الهندسي وفقاً لنظرية (فان هيل) لدى معلمي الرياضيات للمرحلة الثانوية بماليزيا. تكونت عينة الدراسة من 140 من معلمي الرياضيات بالمرحلة الثانوية، منهم 23 معلم و 117معلمة، الذين هم ببرنامج البكالوريوس ومن المتفوقين في الرياضيات، للسنة الدراسية 2010–2011 الفصل الثاني. تم حذف 21 معلم ومعلمة بنسبة 17.1% لتصبح العينة 116 بنسبة 82.9%، تم هذا الحذف بناء على معيار الإجابة على بنود الاختبار في كل مستوى 5/3 و 5/4 . استخدمت الدراسة اختبار فان هيل للتفكير الهندسي المبني في مشروع جامعة شيكاغو، وفقاً لنظرية (فان هيل) لمستويات التفكير الهندسي، اقتصرت الدراسة على 20 بنداً الأولى والتي تمثل الأربعة المستويات الأولى لنظرية

(فان هيل). أظهر تدريج الاختبار باستخدام نموذج راش، بأن بنود المستوى الأول سجلت قيم الصعوبة ما بين 3.81 و0.36، بمتوسط صعوبة –1.91، أما في المستوى الثاني كانت قيم الصعوبة ما بين 1.68 و 0.38، بمتوسط صعوبة –1.02، أما المستوى الثالث تراوحت قيم الصعوبة مابين 0.45 و 3.12 بمتوسط صعوبة 1.21، وأخيراً المستوى الرابع فتراوحت قيم صعوبة بنوده مابين 1.37 و 2.06، بمتوسط صعوبة 1.71 . نلاحظ من خلال هذه النتائج أن البند رقم 13 من المستوى الثالث تساوى في الصعوبة مع البند رقم 17 من المستوى الرابع . وأن البند رقم 14 في المستوى الثالث كان أكثر صعوبة بـــ3.12 لوغاريتم. كما توزعت مستويات التفكير الهندســـي لدى معلمي الرياضيات كالتالي: المستوى الأول 10 معلمين بنسبة 7.1%، والمستوى الثاني 28 معلم بنسبة 20%، ثم المستوى الثالث 42 معلم بنسبة 30%، وأخيراً المستوى الرابع 3 معلمين بنسبة 2.1%.

6- **دراسة سيف خليل، إسماعيل** (2017):

هدفت الدراسة إلى تدريج اختبار لمستويات التفكير الهندســـي لتلامذة المرحلة المتوسطة باستخدام نظرية الاستجابة للمفردة. استخدم البحث اختبار (فان هيل) لمستويات التفكير الهندســـي الذي طوره يوسيسكين (1982) بعد ترجمته، طبق الاختبار على عينة استطلاعية مكونة من 60 تلميذاً وتلميذة من غير عينة التحليل الإحصائي، طبق الاختبار على عينة مكونة من 500 تلميذاً وتلميذة من تلامذة المرحلة المتوسطة، تم اختيارهم بأسلوب الطبقي العشوائي، استخدم نموذج راش أحادي المعلم و البرنامج الإحصائي راسكال، تم التحقق من افتراض أحاديـــة البعد باستخدام التحليل العاملي بطريقة المكونات الأساسية، إذ تم الحصول على عامل واحد ذي معنى مفسر للاختبار، واعتمد العامل نفسه على الحدود الدنيا لـــ(جتمان) الذي يعد العامل دالاً إحصائياً حينما يكون الجذر الكامن الذي يمكن تفسيره يساوي أو يزيد على 1، واعتماد نسبة 0.30 فما فوق على أنها نسبة تشبع مفردات الاختبار بالعامل العام على وفق معيار جيلفورد، ولم يحذف البرنامج أية مفردة من المفردات؛ وللتحقق من مطابقة المفردات للأنموذج استناداً إلى قيمة مربع كاي لحسن المطابقة، تبين أن قيم مربع كاي المحسوبة لجميع المفردات غير دالة إحصائياً، عند مستوى دلالة 0.05، وهذا يعني أنه لم تستبعد أي مفردة؛ كما أن قيمة معامل التمييز للاختبار اقتربت من 1 فقد بلغت قيمته 0.78. تم التحقق كذلك من استقلالية القياس بما يحقق موضوعية

القياس كما تمثل بنموذج راش، وتبين أن هناك مفردتين لم تكن متكافئتين للتقديرات الإحصائية المناظرة لها، وعليه استبعدت هذين المفردتين من الاختبار، وللتخلص من الكسور والإشارات السالبة حولت وحدة اللوجيت لتقديرات صعوبة المفردات ولتقديرات قدرة الأفراد وأخطائها المعيارية إلى وحدة الواط. تم تحديد درجة قطع للاختبار، بطريقة أنجوف التي تعتمد على آراء المحكمين، وقد كانت درجة القطع النهائية 14 وهي نسبة 64% من مفردات الاختبار. طبق الاختبار بصيغته النهائية على عينة التطبيق النهائي التي بلغ عددها 200 تلميذاً وتلميذة تم اختيارهم عشوائياً. وقد بينت النتائج أن أفراد العينة يتمتعون بالتفكير الهندسي، وأنه لا توجد فروق ذو دلالة إحصائية بين الذكور والإناث في مستوى التفكير الهندسي، إلا أنه ظهر هناك فروقاً بين الصفوف الثلاثة ولصالح الصف الثالث على حساب الصفين الأول والثاني .

❖ التعليق على الدراسات السابقة التي تناولت استخدام نظرية الاستجابة للمفردة الاختبارية (IRT) ونموذج راش في بناء وتدريج اختبارات مستويات التفكير الهندسي وفقاً لنموذج (فان هيل):

- **بالنسبة للأهداف:**

لقد تباينت أهداف الدراسات السابقة والتي تناولت موضوع استخدام نظرية الاستجابة للمفردة ونموذج راش، حيث هدفت دراسة (ويلسون، 1990) إلى إعادة تحليل بيانات سابقة مستخدمة في مشروع جامعة شيكاغو، لتطوير اختبار مستويات التفكير الهندسي لفان هيل، وقياس تسلسل عملية التعلم اعتماداً على نظرية (فان هيل) بجامعة شيكاغو (Usiskin,1982)، وبالتالي فهي امتداداً لدراسة يوسيسكين (1982) ولكن باستخدام نموذج راش اللوغاريتمي، كما هدفت دراسة كل من (سيرو، 2007؛ مايكل، 2012) للبحث عن كيفية اكتساب الاستبصار في الإدراك في تضمين المفاهيم الهندسية، وكذلك معرفة العلاقة بين مضطربي طيف التوحد والذين لديهم خلل في التماسك المركزي ومستويات (فان هيل) للتفكير الهندسي، أما دراسة (ستولس، 2015) هدفت إلى استخدام نموذج راش في التقييم التفاعلي بين نظرية (فان هيل) للتفكير الهندسي واختبار (فان هيل) للتفكير الهندسي، كما هدفت دراسة كل من (شونغ و نور، 2015؛ سيف خليل، 2017) إلى تدريج اختبار مستويات التفكير الهندسي لفان هيل الأصلي وفقاً لنموذج راش.

- **بالنسبة للعينات:**

أما في ما يتعلق بالعينة تنوعت أحجام وأنواع العينات المستخدمة في هذه الدراسات من حيث المرحلة العمرية لأفرادها، حيث تناولت دراسة كل من (ويلسون، 1990؛ سيرو، 2007؛ مايكل، 2012؛ سيف خليل، 2017) عينات من تلامذة المرحلة الإبتدائية إلى المرحلة الثانوية وأغلبها تلامذة المرحلة الثانوية على غرار دراستنا الحالية، في حين تناولت دراسة كل من (ستولس، 2015؛ شونغ ونور، 2015) الطلبة معلمي الرياضيات قبل الخدمة. أما من ناحية أحجامها فقد تراوحت ما بين 24 و 655 فرداً، حيث كان أقل حجم للعينة في دراسة (سيرو، 2007) وأكبر حجم للعينة كان في دراسة (ويلسون، 1990) وهي عينة فرعية من بيانات دراسة مشروع جامعة شيكاغو، وهو نفس حجم العينة المستخدم في بحثنا الحالي تقريباً.

- **بالنسبة لأدوات جمع البيانات المستخدمة:**

استخدمت معظم هذه الدراسات اختبار (فان هيل) لمستويات التفكير الهندسي الأصلي، فمنهم من استخدمه بمستوياته الخمسة، و منهم من إكتفى بالمستويات الأربعة الأولى فقط كدراسة (شونغ ونور، 2015؛ مايكل، 2012). في حين نجد الدراسة الوحيدة لـــ (سيرو، 2007) التي استخدم فيها المقابلة من خلال برتوكل به سبع مهام أو بنود تركز على العلاقات بين الأشكال وبين خصائصها ضمن سياق المثلثات والأشكال الرباعية.

- **بالنسبة الأساليب والبرامج الإحصائية:**

تنوعت الأساليب والبرامج الإحصائية المستخدمة في هذه الدراسات والتي اعتمدت على أسلوب نموذج راش اللوغاريتمي أحادي البارامتر، وباستخدام برامج إحصائية مناسبة للتحليل وفقاً للأسلوب الإحصائي للنظرية الحديثة في القياس، استخدمت دراسة (سيف، 2017) برنامج رسكال، وكذلك الفروق بين المتوسطات وتحليل التباين، كما اعتمدت على تحديد درجة القطع وفق طريقة أنجوف(المحكمين)، في حين استخدمت دراسة (ستولس، 2015) برنامج رام، ودراسة (سيرو، 2007) برنامج أكويست، أما دراسة كل من (مايكل، 2012؛ شونغ ونور، 2015) لم يتم توضيح البرنامج الإحصائي المستخدم في تحليل بياناتها.

- **بالنسبة للنتائج:**

تنوعت واختلفت هذه الدراسات من حيث النتائج المتوصل إليها، حيث توصلت نتائج دراسة (ويلسون، 1990) إلى أن متوسطات الصعوبة لمستويات (فان هيل) كانت متسلسلة بحسب توقعات (فان هيل)، ما عدا المستوى الخامس، كما ظهر هناك نوعاً من التداخل بين المستويات بعد استبعاد المستوى الخامس، وأن ما نسبته 3% من العينة بحسب معيار 5/4 بأنها غير متسقة ومتسلسلة هرمياً وفقاً للنظرية الكلاسيكية للقياس، و ما نسبته 4% من العينة وبنفس المعيار بأنها غير متسقة وفقاً لنموذج راش، في حين أظهرت دراسة (سيرو، 2007) من خلال خريطة بند- فرد للمهام أو البنود السبعة، أن ستة بنود وقعت ضمن الحدود المقبولة في حين أظهر البند4 ميلاً قليلاً إلى اليمين من الحدود المقبولة، إلا أنها أظهرت تدرج هرمياً واضحاً ومتسلسلاً وفقاً لمستويات نظرية مخرجات التعلم الملاحظ، أما دراسة (مايكل، 2012) أشارت نتائجها بأن جميع مفردات اختبار مستويات التفكير الهندسي من المستوى الأول إلى المستوى الرابع مطابقتها لنموذج راش، وأن صعوبتها كانت مناسبة، إلا أنها كانت متداخلة بين المستويات، كما أثبتت أن بنود المستوى الأول كانت صعبة نسبياً من المستوى الثاني عند مضطربي طيف التوحد كما كان متوقع من البحث، وأشارت نتائج دراسة (شونغ ونور، 2015) من خلال تدريج اختبار (فان هيل) لمستويات التفكير الهندسي، بأن متوسطات صعوبة المفردات كان بشكل متسلسل وفقاً لنظرية (فان هيل)، إلا أن بعض المفردات أظهرت بأنها ليس بمكانها المتوقع، كما لم تظهر إحصاءات كاي مربع إختلاف بين الدرجات الملاحظـــة والمتوقعـــة وفق نموذج راش في دراسة (ستولس، 2015) لجميع المفردات، إلا أنه تم تسجيل إنخفاض في مؤشر معامل الفصل، وبمتوسط حسابي قريب من الصفر لموقع الأفراد، وانحراف معياري أقل من الواحد وهو أقل من المتوقع، كما أشارت خريطة رايت تداخل بين المفردات حتى بعد تصحيح أثر التخمين، بالرغم من وجود نمط تنازلي واضح في الإجابات الصحيحة ضمن المستويات، إلا أنها كشفت عن مشكلة في بعض مفردات الاختبار بأنها تتطلب سوى جانب معرفي محدد من الهندســـة بدلاً من اختبار التفكير الهندســـي، في حين أشارت دراسة (سيف، 2017) كذلك إلى أن قيم إحصاءات كاي مربع دلت على تطابق جميع مفردات الاختبار مع نموذج راش، وإقتراب معامل التمييز من الواحد، إلا أنه تم حذف مفردتين خلال عملية التحقق من استقلالية القياس، ومن أجل التخلص

من الكسور والإشارات السالبة تم تحويل وحدة اللوجيت إلى وحدة الواط، كما لم يجد فروقاً ذات دلالة إحصائية بين الذكور والإناث في مستويات التفكير الهندســـي.

رابعاً: دراسات تناولت الأداء التفاضلي للمفردة الاختبارية (DIF) باستخدام نظرية الاستجابة للمفردة الاختبارية، والتعليق عليها:

1- **دراسة هاريس وكارلتون (Harris and Carlton, 1993):**

هدفت الدراسة إلى الكشف عن المفردات ذات الأداء التفاضلي لمتغير الجنس، في مفردات الجزء الرياضـــي من اختبار الاستعداد المدرسي (SAT) المكون من مفردات في الهندسة والحساب، ومفردات تحتوى على أشكال ورسوم بيانية. تكونت العينة من 1074 طالباً وطالبة. استخدمت طريقة (مانتل- هانزل) للكشف عن الأداء التفاضلي. بينت النتائج أن متوسط علامات الذكور أعلى من متوسط علامات الإناث على الاختبار وبدلالة إحصائية، كما كانت مفردات الاختبار ككل أسهل على الذكور من الإناث.

2- دراسة غامر وانغلهارد (Gamer and Englhard, 1999):

هدفت إلى الكشف عن المفردات التي تظهر أداءً تفاضلياً حسب متغير الجنس في اختبار القدرة الرياضية، باستخدام البرنامج الإحصائـــي (FACETS) في الكشف عن الأداء التفاضلي لاختبار مكون من 60 مفردة من نوع الاختيار من متعدد، توزعت المفردات كالتالي: 14 مفردة في موضوع تحليل البيانات، و23 مفردة في موضوع الهندسة والقياس، و11 مفردة في موضوع الحساب، و12 مفردة في موضوع الجبر، كما تضمن الاختبار 8 مفردات ذات استجابة محددة البناء، أعدت من طرف خبراء في المناهج والقياس والتقويـــم، وذلك ضمن خطة لتطوير مناهج الرياضيات في ولاية جورجيا الأمريكية. تكونت العينة من 3970 مفحوصاً من طلبة الصف الحادي عشر. أظهرت نتائج الدراسة أن تفوق الذكور على الإناث في موضوعي الهندسة والقياس، وتحليل البيانات، بينما أظهرت أن الإناث تفوقن على الذكور في موضوع الجبر.

3- **دراسة كاثرين وشوان (katherine and Shaun, 2001):**

هدفت الدراسة إلى الكشف عن تجمع المفردات التي تظهر الأداء التفاضلي حسب متغير الجنس وتأثير تغير مواقعها في نموذج اختبار مستوى الرياضيات (MMPE). تكون الاختبار من 40 مفردة من نوع

الاختيار من متعدد، يقيس الاختبار مفاهيم رياضية درسها التلامذة في السنوات الثلاثة الأخيرة من المراحل الدراسية في موضوعات الهندسة، والجبر، والمثلثات، توزعت المفردات على النحو التالي: 10 مفردات في مجال الهندســة، 18 مفردة في مجال الجبر، و12 مفردة في مجال المثلثات، والذي اعد من طرف لجان متخصصة في تطوير اختبارات الرياضيات. وضع الاختبار في نموذجين، النموذج الأول توزعت فيه مفردات الاختبار عشوائياً، والنموذج الثاني توزعت فيه مفردات الاختبار بشكل متدرج من الأسهل إلى الأصعب. طبق الاختبار على 2731 منهم 1066 تلميذاً، و1665 تلميذة، مدة الاختبار ساعة ونصف. استخدمت الدراسة طريقة تجمع المفردات ذات الأداء التفاضلي (Differntial Bundle Analysis Method, DBAM)، من أجل الكشف عن المفردات التي تظهر أداءً تفاضليــاً بحسب متغير الجنس. أظهرت النتائج أن مفردات موضوع الجبــر، كانت أسهل على الجنسين في النموذج الثانــي، منها في النموذج الأول، في حين أن المفردات في موضوعي الهندسة والمثلثات، كانت أسهل لدى الذكور منها عند الإناث في كلا النموذجين.

4- **دراسة عنابي ودودين (Innabi and Dodeen, 2006):**

هدفت الدراسة إلى تقصى المفردات التي تظهر أداءً تفاضليــاً لمتغير الجنس في اختبار الرياضيات للدراسة الدوليــة توجهات في الرياضيات والعلوم (TIMSS ,1999) عينة الأردن. تكونت العينة من 5299 مفحوص 2470 تلميذاً و 2829 تلميذة من الصف الثامن الأساسي. استخدمت طريقة (مانتل- هانزل) في الكشف عن الأداء التفاضلي لــ 124 مفردة من نوع الاختيار من متعدد. أسفرت النتائج على أن هناك 37 مفردة قد أظهرت أداءً تفاضليا، أي ما نسبته 30% من مفردات الاختبار، منها 17 مفردة لصالح الذكور، و20 مفردة لصالح الإناث، وأن أغلب المفردات التي أظهرت أداءً تفاضلياً لصالح الإناث تنتمي لموضوع الجبر والإحصاء، أما المفردات التي كانت سلبية بإتجاه الإناث فقد كانت موضوعاتها تتطلب المخاطرة في الحل، مثل موضوعات التقدير، التوقع، التقريب. كما أن أغلب المفردات التي أظهرت أداءً تفاضلياً لصالح الذكور تنتمي لموضوع الهندســة والقيــاس.

5- **دراسة حمادنة، اياد** (2007):

هدفت الدراسة إلى مقارنة الأداء التفاضلــي لمفردات اختبارات تحديد الكفــاءة اللغوية في اللغة الإنجليزية بالجامعات الأردنية تبعاً لمتغيري الجنس وفرع الثانوية العامة، وفق طريقتي فرق المساحة بين

منحنيات خصائص المفردة في النموذج اللوجيستي ثلاثي المعلم (ICC)، ومانتل- هانزل (MH)، ومعرفة ما إذا كان الأداء التفاضلي يختلف بإختلاف مستوى القدرة، والتعرف على درجة الاتفاق بين الطريقتين، وعلى طبيعة المهارات التي تظهر مفرداتها أداء تفاضلياً. استخدمت ثلاثة نماذج من اختبارات تحديد الكفاءة اللغوية في اللغة الإنجليزية التي تستخدمها ثلاث جامعات، هي آل البيت 100 مفردة، والعلوم والتكنولوجيا 50 مفردة، واليرموك 50 مفردة، وجميعها اختبارات اختيار من متعدد بأربعة بدائل. تكونت عينة الدراسة من 1935 مفحوص ممن تقدموا لهذه الاختبارات في بداية العام الجامعي 2005. من أبرز نتائجها أن أكثر عدد للمفردات التي أظهرت أداء تفاضلياً لمتغير الجنس كان في اختبار جامعة اليرموك، وبنسبة 24% تقريباً من مفردات الاختبار، ثم تبعه اختبار جامعة العلوم والتكنولوجيا بنسبة 21%، فيما بلغت النسبة في اختبار آل البيت 3% تقريباً. وأن أكثر عدد المفردات التي أظهرت أداء تفاضلياً لمتغير فرع الثانوية العامة كان في اختبار جامعة العلوم والتكنولوجيا، وذلك باستخدام طريقة (ICC) حيث أظهرت 9 مفردات أداء تفاضلياً منتظما لصالح الفرع العلمي. تبين أن طريقة (MH) استطاعت أن تكشف عن أكبر عدد من المفردات التي تظهر أداءً تفاضلياً منتظماً، في ما أظهرت طريقة (ICC) فاعليتها في الكشف عن المفردات ذات الأداء التفاضلي من النوع غير المنتظم. وأن أكثر المهارات التي أظهرت مفرداتها أداء تفاضلياً لمتغير الجنس، هما مهارتي القواعد والمفردات، وكانت في معظمها لصالح الإناث، وأما أكثر مهارة أظهرت مفرداتها أداء تفاضلياً لمتغير فرع الثانوية العامة، فهي مهارة استيعاب المقروء، وكانت في معظمها لصالح الفرع العلمي.

6- **دراسة وائل محمد، أحمد مبارك** (2010):

هدفت الدراسة إلى الكشف عن الأداء التفاضلي في اختبار الرياضيات في الدراسة الدولية بيزا 2006، تبعا لمتغير اللغة، باستخدام طريقة (مانتل-هانزل) . بلغ عدد مفردات اختبار الرياضيات 48 مفردة. تقدم للاختبار 400000 مفحوص من الصف العاشر، تم اختيارهم عشوائياً من مجتمع الدراسة. بينت النتائج وجود 27 مفردة أظهرت أداءً تفاضلياً لمتغير اللغة العربية مقابل اللغة الإنجليزية، حسب طريقة (مانتل-هانزل)، أي ما نسبته 56%، والمفردات التي أظهرت أداءً تفاضلياً كان 74% منها لصالح مجموعة اللغة الإنجليزية. كما بينت النتائج أيضا أن متوسط الأداء في الأردن وترتيبه قد تحسن بعد حذف المفردات التي أظهرت أداءً تفاضلياً، والسبب قد يعود إلى أن مفردات الاختبار الأصلية كانت

مكتوبة باللغــة الإنجليزية، وترجمــة تلك المفردات ربما يغير معناها، أو يؤدي إلى حذف معلومــات ضرورية تجعل الاستجابــة عليها صعبة.

7- دراسة عبد العزيز (Abdelaziz, 2010):

هدفت الدراسة إلى استكشاف الأداء التفاضلــي لمتغير الجنس في اختبار الرياضيات بماليزيا مكون من 40 مفردة. استخدمت الدراســة ثلاث طرق، هي: مانتل-هانزل، والصعوبــة المحولة، وفرق معلم الصعوبة (منحنى خصائص المفردة). على عينة مكونة من 3390 مفحوص، منهم 1600 تلميذاً و 1790 تلميذة من تلامذة الصف الحادي عشر. وقد كان التوافق الأكثر بطريقة (مانتل-هانزل)، بينما كان التوافق الكلي بين طريقتي الصعوبــة المحولة وفرق معلم الصعوبة، و بلغت نسبة المفــردات التي أظهرت أداء تفاضلياً 42.5% باستخدام طريقة (مانتل-هانزل)، بواقع 7 مفردات لصالح الإناث، و10 مفردات لصالح الذكور. وبشكل عام أظهرت مفردات الهندســة أداء تفاضليــاً لصالح الذكور، يضاف إلى ذلك بعض المفردات التي كانت تقيس المفاهيم المجردة الجبرية والتعريفات والحساب العددي، بينما أظهرت مفردات الجبر وبعض التطبيقات الحياتية أداءً تفاضليــاً لصالح الإنــاث.

8-دراسة حسان العمري وعبد الله شطناوي (2016):

هدفت الدراسة إلى الكشف عن الأداء التفاضلي لمفردات الاختبار الوطني الأردني لمادة الرياضيات للصف العاشر تبعا لمتغير الجنس. تكونت العينة من 2400 تلميذاً وتلميذة، بواقع 1200 تلميذ و 1200 تلميذة ممن تقدموا للاختبار للعام الدراسي 2013/2014. ولتحقيق هذا الهدف تم الحصول على استجابات التلامذة على الاختبار المكوّن من 39 مفردة. استخدم (مؤشر الأداء التفاضلي اللاتعويضي NCDIF) في الكشف عن الأداء التفاضلي للمفردات، حيث بلغ عدد المفردات في صورتها النهائية بعد مطابقة المفردات للنماذج المستخدمة 25 مفردة مطابقة للنموذج الأحادي المعلمة، و 34 مفردة مطابقة للنموذج الثنائي المعلمة، وكشفت النتائج عن وجود أداء تفاضلــي منتظم لــ 9 مفردات من أصل16 مفردة هي عدد المفردات المشتركة بين الذكور والإناث، وعن وجود أداء تفاضلي غير منتظم لــ 15 مفردة من أصل 26 مفردة هي عدد المفردات المشتركة بين الذكور والإناث. وخلصت الدراسة إلى جملة من التوصيات تدعو إلى إجراء دراســات مختلفة تتناول الأداء التفاضلــي لمفردات الاختبار الوطني لضبط نوعية التعليم

وخصوصا في مباحث أخرى غير الرياضيـــات، والكشف عن الأداء التفاضلـــي للاختبـــارات الوطنية الأخرى التي تم تطبيقها إلكترونياً على التلامذة.

❖ التعليق على الدراسات السابقة التي تناولت الأداء التفاضلـــي للمفردة الاختبارية (DIF) باستخدام نظرية الاستجابة للمفردة الاختبارية:

- **بالنسبة للأهداف:**

تباينت أهداف الدراسات السابقة التي تناولت الأداء التفاضلي للمفردة الاختبارية، حيث نجد الهدف من دراسة كل من (هاريس وكرلتون، 1993؛ غامر وانغلهارد، 1999؛ كاثرين وشوان، 2001؛ عنابي ودودين، 2006؛ عبد العزيز، 2010؛ حسان وعبد الله، 2016)، الكشف عن المفـــردات ذات الأداء التفاضلـــي حسب متغير الجنس (ذكور وإناث)، في حين أن دراسة (حمادنة، 2007) هدفت للكشف عن الأداء التفاضلـــي بحسب متغير الجنس و فرع الثانويـــة العامة، أما دراسة (وائل محمد، 2010) هدفت للكشف عن الأداء التفاضلـــي بحسب متغير اللغة.

- **بالنسبة للعينات:**

أما في ما يتعلق بالعينة تنوعت أحجام وأنواع العينات المستخدمة في دراسات هذا المحور من حيث المرحلة العمرية لأفرادها، حيث تناولت معظم هذه الدراسات (هاريس وكرلتون، 1993؛ غامر وانغلهارد، 1999؛ كاثرين وشوان، 2001؛ وائل محمد، 2010؛ حمادنة، 2007؛ عبد العزيز، 2010؛ حسان وعبد الله، 2016) تلامذة المرحلة الثانويـــة، على غرار بحثنا الحالي، ما عدا دراسة (عنابي ودودين، 2006) والتي تناولت تلامذة المرحلة المتوسطة. أما في ما يخص أحجام العينات فتراوحت ما بين 1074 كأقل عينة في دراسة (هاريس وكرلتون، 1993)، و أكبر عينة (400000) في دراسة (وائل محمد، 2010) والتي استخدم فيها بيانات الدراسة الدوليـــة بيزا 2006 (PISA).

- **بالنسبة لأدوات جمع البيانات المستخدمة:**

استخدمت معظم دراسات هذا المحور اختبارات تحصيلية للقدرة الرياضية، كدراسة (غامر وانغلهارد، 1999؛ كاثرين وشوان، 2001؛ عنابي ودودين، 2006؛ وائل محمد، 2010؛ عبد العزيز، 2010؛ حسان

وعبد الله، 2016)، ما عدا دراسة (هاريس وكرلتون، 1993)، التي استخدمت اختبار الاستعداد المدرسي (SAT) للهندســـة والحساب، ودراسة (حمادنة، 2007)، التي استخدمت اختبار الكفاءة اللغوية الإنجليزية.

- **بالنسبة الأساليب والبرامج الإحصائية:**

تنوعت الأساليب والبرامج الإحصائية المستخدمة في دراسات هذا المحور، حيث استخدمت معظم الدراسات (هاريس وكرلتون، 1993؛ غامر وانغلهارد، 1999؛ كاثرين وشوان، 2001؛ عنابي ودودين، 2006؛ وائل محمد، 2010؛ حسان وعبد الله، 2016)، طريقة (مانتل–هانزل) في الكشف عن الأداء التفاضلي، ما عدا دراسة (عبد العزيز، 2010؛ حمادنة، 2007)، التي استخدمت طريقة فرق المساحة بين منحنيات خصائص المفردة أو فرق معلم الصعوبة، والصعوبة المحولة، إضافة إلى طريقة (مانتل–هانزل) في الكشف عن الأداء التفاضلـــي، على غرار بحثنا الحالي، ودراسة (حسان وعبد الله، 2016)، والتي استخدمت فيها مؤشر الأداء التفاضلي اللاتعويضي (NCDIF).

- **بالنسبة للنتائج:**

اختلفت دراسات الأداء التفاضلـــي للمفردات الاختبارية، من حيث النتائج المتوصل إليها، حيث توصلت نتائج دراسة (هاريس وكرلتون، 1993)، أن مفردات الاختبار كانت سهلة بالنسبة للذكور مقارنة بالإناث، أما دراسة (غامر وانغلهارد، 1999) فكشفت عن تفوق للذكور على الإناث في موضوعي الهندســـة والقياس وتحليل البيانات، بينما تفوقت الإناث على الذكور في موضوع الجبر، وفي نفس الاتجاه تقريباً كشفت دراسة (عنابي ودودين، 2006)، أن أغلب المفردات التي أظهرت أداءً تفاضليـــاً لصالح الإناث تنتمي لموضوع الجبر والإحصاء، وأن أغلب المفردات التي أظهرت أداءً تفاضلياً لصالح الذكور تنتمي لموضوع الهندســـة والقياس، وكذلك دراسة (عبد العزيز، 2010) التي كشفت عن أداء تفاضلي لصالح الذكور في مفردات الهندســـة، يضاف إليها بعض المفردات التي تقيس المفاهيم المجردة الجبرية والتعريفات والحساب العددي، بينما أظهرت مفردات الجبر وبعض التطبيقات الحياتية أداءً تفاضلياً لصالح الإناث، وتقريباً وبنفس النتائج كشفت دراسة (كاثرين وشوان، 2001)، بأن مفردات موضوع الجبر كانت أسهل على الجنسين في النموذج الثاني منها في النموذج الأول، و أن مفردات موضوعي الهندسيـــة والمثلثات كانت أسهل لدى الذكور منها عند الإناث في كلا النموذجين، إلا أن دراسة (حمادنة، 2007) كشفت أن أكثر المهـــارات التي أظهرت مفرداتها أداءً تفاضليـــاً لمتغير الجنس هي مهارتي القواعد

والمفردات و معظمها كان لصالح الإناث، وفي دراسة (وائل محمد، 2010) كشفت أن النسبة الكبيرة من المفردات التي أظهرت أداءً تفاضلياً كانت لصالح مجموعة اللغة الإنجليزية، وأن متوسط الأداء في الأردن وترتيبه قد تحسن بعد حذف المفردات التي تظهر الأداء التفاضلي، بينما كشفت دراسة (حسان وعبد الله، 2016) عن وجود أداء تفاضلي منتظم لـ9 مفردات من أصل 16 مفردة، وعن وجود أداء تفاضلي غير منتظم لـ15مفردة من أصل 26 مفردة بين الطلبة والطالبات.

خامساً: التعليق العام على الدراسات السابقة و موقع الدراسة الحالية منها:

يتضح من تناولنا للبحوث والدراسات التي اجريت حول تطوير وبناء اختبارات لمستويات التفكير الهندسي وفق نموذج (فان هيل)، وباستخدام المقاربتين القياسيتين النظرية الكلاسيكية للقياس ونظرية الاستجابة للمفردة و نموذج راش خاصةً، لدراسة تطور التفكير الهندسي لدى التلامذة أو المعلمين الطلبة وفقا لنموذج (فان هيل) كنظرية أو نموذج للتدريس، للتحقق من مدى نجاعة النظرية في وصف التفكير الهندسي أو تعلم الهندسة لدى التلامذة. والتي أجمعت على أن نظرية أو نموذج (فان هيل) تشكل إطاراً نظرياً يمكن استخدامه لتفسير كيفية تعلم التلامذة الهندسة، وتشخيص المشكلات والصعوبات التي يعاني منها المتعلمين في تعلم الهندسة، وبأنها تشكل أساساً بنائياً لتعليم الهندسة، وأن مستوياتها الأربعة الأولى مفيدة وناجعة في وصف تفكير التلامذة في الهندسة. كما حاولت بعض الدراسات البحث في خصائص النظرية ومستويات التفكير فيها، من حيث كونها هرمية أم لا؟ وطبيعة هذه المستويات كونها منفصلة أو متصلة؟ ودور اللغة فيها. وأوصت معظمها بضرورة الاهتمام بتطوير أدوات لقياس مستويات التفكير الهندسي وفق نموذج (فان هيل). وقد أدى ذلك إلى محاولة بعض الباحثين لتطوير أدوات لفحص خصائص هذه النظرية، خاصة بعدما أوضحت نتائج بعض الدراسات أن التلامذة يتأرجحون في تفكيرهم الهندسي بين مستويين متتاليين أو أكثر، مما يعني أن مستويات (فان هيل) ليست منفصلة بل متصلة.

وبالرغم من الاستخدام الواسع لاختبارات مستويات التفكير الهندسي المبنية وفق نظرية (فان هيل) في البحوث والدراسات، إلا بعض الملاحظات ظهرت حولها، منها(jaime and Gutiérrez,1994):

1. هناك بعض الشكوك حول قدرة اختبارات (فان هيل) بمفرداتها التي من نوع اختيار من متعدد على قياس التفكير الهندسي للتلامذة (Crowly,1990 ;wilson,1990)، كما ورد في دراسة (jaime and

1994 Gutiérrez,) . إلا أن أهم ما يتميز به هذا النوع من الاختبارات هو سهولة تطبيقها مع أعداد كبيرة من التلامذة، وسهولة تصحيحها وتصنيفها للتلامذة على مستويات (فان هيل) للتفكير الهندســـي.

2. استخدام طريقة المقابلـــة يحتاج إلى وقت طويل، ولكنها تمتاز بالمعرفـــة العميقة التي تزود الباحثين بمعلومات حول تطور التفكير الهندســـي لدى التلامذة. (Burger and Shaugnessy,1986)

أ- العلاقة بين الدراسات السابقـــة والبحث الحالي:

-يتشابه بحثنا الحالي مع دراسة بوسالم في بناء اختبار تحصيلي محكي المرجع والتحقق من موضوعية نتائجه باستخدام نظريتي القياس الكلاسيكية والحديثة (نموذج راش).

-اتفق هذا البحث مع دراسة (بوسالم، وكريسليب، والخياط) في استخدام اختبار تحصيلي محكي المرجع، بالرغم من أن هذه الدراسات لم تعتمد على نظرية معينة في بناء اختباراتها. في حين اختلفت مع بقية الدراسات والتي اعتمدت على اختبارات معيارية المرجع، واختبارات للذكاء والقدرات العقلية.

-يتفق بحثنا الحالي مع دراسة العطاس في التحقق السيكومتري من اختبار مستويات التفكير الهندســـي وفق نموذج (فان هيل) باستخدام النظرية الكلاسيكية في القياس، في حين أن بقية الدراسات لم يكن هدفها الرئيس البناء والتحقق السيكومتري وفق النظرية الكلاسيكية للقياس.

-يتفق بحثنا الحالي مع دراسة (ويلسون، ستولس، وشونغ ونور، و سيف خليل اسماعيل) في التحقق السيكومتري من اختبار مستويات التفكير الهندســـي وفق نموذج (فان هيل) باستخدام نظرية الاستجابة للمفردة (نموذج راش)، في حين أن بقية الدراسات لم يكن هدفها الرئيس البناء والتحقق السيكومتري وفق نموذج راش.

-اتفق هذا البحث مع دراسة كل من: (يوسيسكين، كريسليب، ميمي السيد، ماكبريد، العطاس، ويلسون، سيرو، مايكل، سيف، وهاريس وكرلتون، و غامر وانغلهارد، وكاثرين وشوان، و مأمون، ووائل وحمادنة، وعبد العزيز، وعبد الله وحسان) في الاعتماد على عينة من تلامذة المرحلة الثانوية، إلا أنها اختلفت مع دراسة ميمي السيد في حجم العينـــة الذي كان صغيراً نوعاً ما 202 تلميذاً وتلميذة، و الذي يعتبر أقل حجم عينة مرغوب للتحليل وفي نموذج راش.

-يتفق بحثنا الحالي مع دراسة القهوجي في استخدامه لبرنامجي وينستبس وبيلوغ، وكذلك برنامج أس بي أس أس، في تحليل البيانات وفق نظريتي القياس الكلاسيكية والحديثة.

-يتفق بحثنا الحالي مع دراسة (أمينة كاظم، وبوسالم، والخياط، وأيمن القهوجي) في حذف المفردات وفق نموذج راش فقط ودون حذف أي مفردة وفق النظرية الكلاسيكية للقياس، كما اختلفت مع دراسة (عوض الله، والهاشمي، وميمي السيد) والتي تم فيها حذف المفردات وفق النظريتين. إلا أن حذف المفردات وفق النظرية الكلاسيكية في دراسة كريسليب يعتبره الباحث تجاوزاً بما أن الاختبار محكي المرجع، والذي لا يهتم بتحديد المكانة بالنسبة للفرد بين أقرانـــه، وبالتالي في هذا النمط من الاختبارات لا تستبعد أيا من المفردات فيه طالما أنها تقيس هدفا سلوكيا مهما.

-يتفق بحثنا الحالي مع دراسة (نوري، و مخلوف، والطنة، ولينا فؤاد) في عملية بناء اختبار لمستويات التفكير الهندسي وفق نظرية (فان هيل). إلا أنه اختلف مع بقية الدراســات، والتي استخدمت اختبار مستويات التفكير الهندسي لفان هيل الأصلي المنشور.

-اختلف بحثنا الحالي مع دراسة (الشويخ، وعفانة، والطنة، و لينا فؤاد، ونوري، وكريسليب وتشين) الذين استخدموا معادلة كيودر- ريتشاردسون(kr_{21}) للتحقق من الثبات، والتي من المفروض أنها لا تصلح للاختبارات المحكية، اضافة إلى أن دراسة كريسليب وتشين تم فيها حذف مفردات سهلة جداً وصعبة جداً وهو خلافاً لما تتطلبه الاختبارات المحكية المرجع.

-يتفق بحثنا الحالي مع دراسة مخلوف الوحيدة في استخدامها لمعادلـــة ليفنجستون للتحقق من ثبات الاختبارات المحكية المرجع.

- اختلف بحثنا الحالي مع دراسة ويلسون الذي استخدم عينة صغيرة جداً حجمها 24 تلميذاً في التحليل وفق نموذج راش، مما يطرح تساؤلا حول مصدقية نتائج هذا التدريج وفق نموذج راش.

-يتفق بحثنا الحالي مع دراسة (ويلسون، مايكل، رلى يوسف، عفانة، الطنة، لينا فؤاد، نوري جليل، شونغ، ونور) في اقتصارها على مستويـــات (فان هيل) الأربعة الأولى، في عملية بناء وتطوير اختبارات لمستويات التفكير الهندســـي وفق نظرية (فان هيل).

-اختلف بحثنا الحالي مع دراسة سيرو التي استخدم فيها المقابلـــة من خلال طرح سبع مهام أو أسئلة هندسيـــة على شكل برتوكل.

-اختلف بحثنا الحالـــي مع دراسة سيف الذي اعتمد في تحديد درجة القطـــع على طريقة أنجوف (المحكمين).

-يتفق بحثنا الحالي مع دراسة (يوسيسكين، الشويخ، مخلوف، عفانة، رباب، لينا فؤاد ونوري) في الكشف وتحديد مستويات التفكير الهندســـي وفق نموذج (فان هيل) لدى أفراد العينة.

-تشابه بحثنا الحالي مع دراسة شويخ ونوري في نتائج توزع درجات التلامذة على مستويات (فان هيل).

-يتفق بحثنا الحالي مع دراسة (هاريس وكرلتون، غامر انغلهارد، كاثرين وشوان، عنابي ودودين، عبد العزيز، عبد الله وحسان) في الكشف عن الأداء التفاضلي للمفردة حسب متغير الجنس (ذكور وإناث)، في حين اختلف عن دراسة حمادنـــة الذي اضاف فرع الثانوية العامة، ودراسة مأمون التي اقتصرت على متغير اللغة الانجليزية.

-اختلف بحثنا الحالـــي مع دراسات الأداء التفاضلـــي للمفردة، في حجم العينة المستخدمة فيها، والتي استخدمت أحجام عينات كبيرة نوعا ماً مقارنة بعينة بحثنا.

-يتفق بحثنا الحالي مع دراسة (غامر وانغلهار، كاثرين و شوان، عنابي و دودين، مأمون، وائل، عبد العزيز، عبد الله وحسان) في استخدام اختبارات تحصيلية للقدرة الرياضية في الكشف عن الأداء التفاضلي للمفردة، ويتشابه مع دراسة هاريس وكرلتون اللذان استخدما اختبار للاستعداد المدرســـي في الهندســـة والحساب.

-يتفق بحثنا الحالي مع كل من دراسات الأداء التفاضلي للمفردة المعروضة في هذا البحث في استخدام طريقتي (مانتل–هانزل)، وفرق المساحة بين منحنيات خصائص المفردة أو فرق معلم الصعوبة، في حين اختلف مع دراسة عبد الله وحسان التي استخدم فيها مؤشر الأداء التفاضلي اللاتعويضي للكشف عن الأداء التفاضلي للمفردة.

ب- مكانة البحث الحالي بين الدراسات الـــسابقة:

-يعتبر البحث الحالي من البحوث الأولى المقارنـــة – على حد علم الباحث – على المستوى الدراسات العربية والأجنبية، بين نظريتي القياس الكلاسيكية ونموذج راش، في بنـــاء وتدريـــج اختبار لمستويات التفكير الهندســـي وفقا لنموذج (فان هيل).

-يعتبر البحث الحالـــي من البحـــوث الأولى – على حد علم الباحث – التي بحثت في الكشف عن الأداء التفاضلي للمفردة بحسب متغير الجنس، لاختبار مستويات التفكير الهندســـي مبني وفق نموذج (فان هيل)، والذي يعتبر مؤشراً مهم للصدق البنائي للاختبار، وكذلك مؤشراً للثبات (اللاتغاير).

ج- استفادة الباحث من الدراسات والبحوث السابقـــة:

-استفاد الباحث من الدراسات والبحوث السابقـــة في تحديد مشكلة البحث، وصياغة فرضياته، وفي إعداد اختبار لمستويات التفكير الهندسي وفقاً لنموذج (فان هيل)، بالاعتماد على بعض مفردات اختبـــار (فان هيل) الأصلي المطور بمشروع جامعة شيكاغو (1982)، و إجراء عليها بعض التعديلات، وبموافقة مؤلفه (يوسيسكين زلمان)، والذي أبدى موافقته للاستعانة بالاختبار من خلال اتصال شخصي معه، انظر الملحق رقم (09)، وحثه لنا على استخدام أو الاستعانة بالنسخة المترجمة من طرف (جهاد الشويخ، 2005).

-التعرف إلى نموذج (فان هيل)، من حيث خصائصه، ومستوياته، و مراحل تعليمـــه.

-التعرف على الأساليب و البرامج الإحصائية المستخدمة في هذه الدراسات والاستفادة منها.

-استفادة الباحث من هذه الدراســـات في تفسير النتائـــج التي توصل إليها البحث الحالي تفسيراً علمياً وموضوعياً.

الفصـــل السادس

الإجـــراءات الميدانيـــة للبحث

تمهيد:

يتضمن هذا الفصل إجراءات البحث الاستطلاعـــي من حيث تحديد منهج البحث، ومجتمع البحث، واختيار عينـــة ممثلة للمجتمع، كما يتضمن تعريف مختصر لأداة البحث، و كيفية وإجراءات التطبيق ميدانياً لاختبار مستويات التفكير االهندســـي المبني وفقاً لنموذج (فان هيل) المستخدم في جمع بيانات عينة البحث، وعملية تفريغ بيانات الاختبار وتصحيحها، كما يهتم الفصل بتحديد البرامج، والأساليب الإحصائية المستخدمة لمعالجة بيانات البحث وتحليلها.

1- المنهج المتبع في البحث:

ولتحقيق أهداف البحث اتبع الباحث المنهج الوصفـــي الإحصائـــي، والذي يهتم بوصف الظواهر وتنظيمها وتبويبها بيانياً، لإلقاء الضوء على ما تنطوي عليه من معلومـــات، ولذا يتناول هذا المنهج بالدراسة جميع المقاييس التي تمكن الباحث من وصف بيانات بحثه، وتلخيصها بصورة كمية، وهو عبارة عن استخدام الطرق الرقمية والرياضية في معالجة وتحليل البيانات، وإعطاء التفسيرات المنطقية المناسبة لها ويتم ذلك من خلال عدة مراحل، وهي:

1. جمع البيانات الإحصائية عن الموضوع.
2. عرض هذه البيانات بشكل منظم وتمثيلها بالطرق الممكنة.
3. تحليل البيانات.
4. تفسيرات البيانات من خلال ما تعنيه الأرقام المجمعة من نتائج. (محمد على، 2011، ص.398).

وبالتالي فإن المنهج الوصفي، هو عبارة عن:

" وصف دقيق وتفصيلي لظاهرة أو موضوع محدد على صورة نوعية أو كميـــة رقمية، فالتعبير الكيفي يصف الظاهرة ويوضح خصائصها، أما التعبير الكمـــي فيعطينا وصفاً رقمياً يوضح مقدار هذه الظاهرة أو حجمها، ودرجة ارتباطها مع الظاهرات المختلفة الأخرى، وقد يقتصر هذا المنهج على وضع قائم في فترة زمنية محددة أو تطويراً يشمل عدة فترات زمنية " . (دويدري، 2002، ص.183)

مع مراعاة الباحث لخصوصية بحثنا الحالـــي، من خلال محاولة تكييف هذا المنهج لهدف ولطبيعة المشكلة موضوع البحث، وبما يلبي الشروط التي يتطلبها البحث السيكومتري المقارن لأداة البحث المتمثلة في اختبار لمستويات التفكير الهندســـي، وعملية التحقق من كفاءتها القياسية.

2- المجتمع وعينة البحث:

2-1- تحديد المجتمع الأصلي للبحث:

المجتمع هو الهدف الأساســـي من البحث، حيث أن الباحث يستطيع في النهاية تعميم النتائج عليه، فنحن لا ندرس عينـــات، وإنما ندرس مجتمعات، وما العينة التي نختارها إلا وسيلة لدراسة خصائص المجتمع . ولذلك فإن الخطوة الأولى في اختيار العينـــة، هي تعريف مجتمع البحث، والمكون من جميع تلامذة الشعب العلمية (علوم تجريبية، رياضيات، تقني رياضي)، الذين يزاولون دراستهم بالمرحلة الثانوية المسجلين في السنة الدراسيـــة 2016-2017 . وعددهم 18125 تلميذاً وتلميذة، منهم 7892 تلميذاً و 10233 تلميذة.

تحصل الباحث على الإحصاءات المتعلقة بالتلامذة في الشعب العلمية في كل مستوى من المستويات الدراسية بالمرحلة الثانوية والتي سحبت منها عينة البحث، من مصلحة التنظيم التربوي (مكتب التعليم الثانوية) بمديرية التربية لولاية الوادي، كما يوضحها الجدول رقم (06):

جدول (06)

يوضح توزيع أفراد المجتمع الأصلي وفقا لمتغيري الجنس والمستوى الدراسي.

المستوى / الجنس	المستـــوى والشعب الدراسيـــة							المجموع الكلي
	أولى ثانوي	ثانية ثانوي			الثالثة ثانوي			
	ج.م.علوم	ت.رياضي	رياضي	علوم	ت. رياضي	رياضي	علوم	
ذكور	2775	503	129	1618	560	208	2099	7892
إناث	3986	360	103	1987	511	163	3123	10233
المجموع	6761	863	232	3605	1071	371	5222	18125

المصدر: مصلحة التنظيم التربوي (مكتب التعليم الثانوي) مديرية التربية لولاية الوادي.

2-2- طريقة اختيار عينة الثانويات:

بعد حصول الباحث على تعداد تلامذة الشعب العلميـــة في المجتمع الأصلي، بحسب متغيري الجنس والمستوى الدراسي، وكذلك على تعداد ثانويـــات الولاية والتي كان عددها 61 ثانوية، من مكتب التعليم الثانوي مصلحة التنظيم التربوي بمديرية التربية لولاية الوادي، قسم الباحث الولايـــة إلى منطقتين جغرافيتين متباينتين، وهما: منطقة وادي سوف، ومنطقة وادي ريغ.

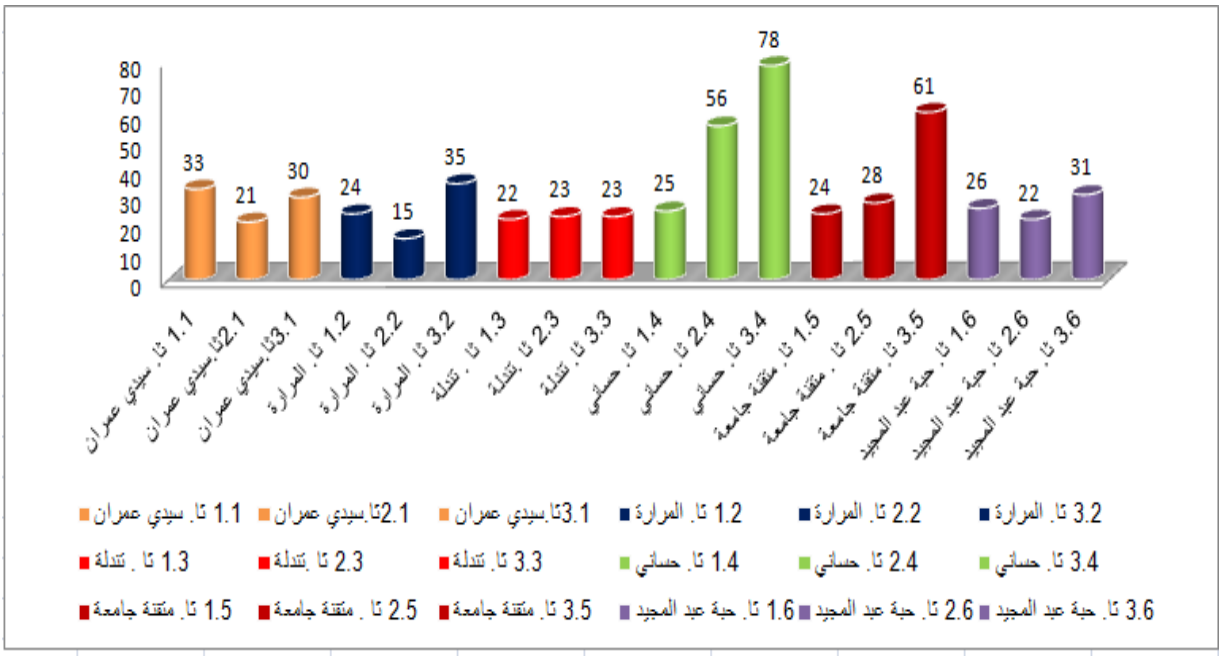
وقع اختيار الباحث على منطقة وادي ريغ، وهذا لاعتبارات عديدة، منها باعتبارها منطقة سكن وعمل الباحث، مما سهل عملية التنقل بين البلديات التي تقع فيها الثانويات المختارة أثناء عملية تطبيق الاختبار، كذلك تجانس عينة التلامذة فيها نوعاً ما، من ناحية المستوى الحضري والإجتماعي والإقتصادي والتعليمي والثقافي، والتي ضمت 15 ثانوية، حيث اختيرت منها 6 ثانويات عشوائياً لسحب عينة البحث الاستطلاعية والاساسية، وبذلك أصبحت عينة الثانويات تمثل ما نسبته تقريباً 10% من مجموع ثانويات الولاية، وما نسبته 40% من مجموع ثانويات منطقة وادي ريغ. والجدول رقم (07)، والشكل رقم (20) يوضحان توزيع أفراد عينة البحث على الثانويات الستة المعنية.

جدول (07)

يوضح توزيع عينة البحث على الثانويات الستة والمستويات الدراسية الثلاثة.

الثانويات	المستويات الدراسية						المجموع	
	أولى ثانوي		ثانية ثانوي		ثالثة ثانوي			
	العدد	النسبة	العدد	النسبة	العدد	النسبة	العدد	النسبة
ثا. سيدي عمران	33	%21.42	21	%12.72	30	%11.62	84	%14.55
ثا. المرارة	24	%15.58	15	%9.09	35	%13.56	74	%12.82
ثا. تندلة	22	%14.28	23	%13.93	23	%8.91	68	%11.78
ثا. حساني لخضر	25	%16.23	56	%33.93	78	%30.23	159	%27.55
ثا. متقنة جامعة	24	%15.58	28	%16.96	61	%23.64	113	%19.58
ثا. حبة المغير	26	%16.88	22	%13.33	31	%12.01	79	%13.69
المجموع	154	%26.69	165	%28.59	258	%44.71	577	%100

يتضح من الجدول رقم (07)، أن توزيع عينة البحث في الثانويات الستة، كان بنسب متقاربة والتي تراوحت ما بين 11.78% و 19.58%، ما عدا النسبة 27.55% والتي تشير إلى ثانوية الشهيد حساني لخضر بجامعة، والتي سجلنا فيها أعلى نسبة من بين نسب الثانويات الأخرى، ويرجع ارتفاع عدد العينة المسحوبة من هذه الثانوية، إلى التسهيلات التي وجدها الباحث في تطبيق الاختبار بحكم عمل الباحث فيها.



شكل (20). يوضح توزيع عينة البحث على الثانويات الستة والمستويات الدراسية الثلاثة.

2-3- طريقة اختيـــار عينة التلامذة:

2-3-1- اختيار عينة البحث الاستطلاعـــي الأولى:

اختار الباحث عينـــة متوفرة (Convenience Sample)، من تلامذة السنة الأولى ثانوي جذع مشترك علوم من ثانوية حساني لخضر بجامعة، مكونة من 34 تلميذاً وتلميذة، بمتوسط عمري 17 سنة، وانحراف معياري 1.12. وذلك لغرض التحقق من وضوح التعليمات، وصياغة المفردات، وملاءمة البدائل، وتحديد الزمن المناسب للإجابة عن مفردات الاختبار، وكذلك التعرف على أي مشكلات قد تواجه الباحث أثناء التطبيق، ليمكن تداركها قبل التطبيق على العينة النهائية.

2-3-2- اختيار عينة البحث الاستطلاعي الثانية:

تم اختيار عينة متوفرة، من الثانويـــات الستة للدراسة الاستطلاعية مكونة من ثلاثة أقسام دراسية (أولى ثانوي، ثانية ثانوي، ثالثة ثانوي)، ومن 104 تلميذاً وتلميذة، منهم 54 تلميذاً و50 تلميذة، بمتوسط عمري 18 سنة، وانحراف معياري1.6. والجدول رقم (08) يوضح توزيع عينة الدراسة الاستطلاعية.

جدول (08)

يوضح عدد أفراد عينة البحث الاستطلاعية ن=104 **ونسبتهم المئوية موزعين بحسب المستوى الدراسي والجنس.**

الجنس / المستوى	عدد الذكور	النسبة المئوية	عدد الإناث	النسبة المئوية	تعداد التلامذة لكل مستوى	النسبة المئوية
أولى ثانوي	23	%42.59	8	%16	31	%29.8
ثانية ثانوي	15	%27.77	18	%36	33	%31.7
ثالثة ثانوي	16	%29.62	24	%48	40	%38.5
المجموع	54	%51.92	50	%48.07	104	%100

يتضح من الجدول رقم (08) أن اختيار أفراد عينة البحث الاستطلاعي كان متناسباً ومتقارباً، بحسب متغيري الجنس (الذكور و الإناث)، والمستوى الدراسي (السنة الأولى والثانية والثالثة ثانوي)، وهو مؤشر جيد في سبيل تحقيق أهداف الدراسة الاستطلاعية المرجوة منها في التحقق الأولـــي من الخصائص السيكومترية لاختبار مستويات التفكير الهندســـي.

2-3-3- اختيار عينة البحث الأساسية:

وبنفس الطريقة التي اختيرت بها عينة الدراسة الاستطلاعية، تم اختيار العينة الأساسية، حيث اختار الباحث الأفواج التربوية (الأقسام) الشعب العلمية التي طبق عليها الاختبار جماعياً، خلال الحصة الدراسية الواحدة، طبق الاختبار في المرة الأولـــى على عينة مكونة من 682 تلميذاً وتلميذة، والتي نتج عنها إلغاء 105 استمارة إجابة، لعدم استجابة التلامذة لتعليمات الاختبار، وبالتالي تكونت العينة الأساسية النهائية من 577 تلميذاً وتلميذة، منهم 247 تلميذاً و 330 تلميذة، بمتوسط عمري 18 سنة، وانحراف معياري 1.61. وقد تم مراعاة أهم معايير اختيار العينة المتوفرة، وهي ضمان توزيع معقول للعينة، حسب جنس وعمر التلامذة، وكذلك حسب متغير المستوى الدراســـي، كما يعتبر هذا النوع من العينـــات، أقل تكلفة وجهداً ووقتاً وتمتاز بسهولة إدارتها، كما يضمن الباحث فيها مشاركة أكبر للمبحوثين، والأهم من ذلك هو إمكانية

تعميم نتائجها على أشخاص متشابهين في الخصائص (James and Mcmillan, 1996, p.96). كما حرص الباحث تقصـــي الرغبة والتعـــاون من قبل المستجيبين، لضمان دقة الاستجابة نظراً لطبيعة البحث الذي يهدف لاعداد اختبار لمستويات التفكير الهندسي وفقاً لنموذج (فان هيل).

2-4- حجم عينة البحث:

يرى نانلي (1967) Nunnally كما ورد عن (ليندا و جيمس)، إنه لا توجد قاعدة مطلقة لأقل عدد من المفحوصين يلزم في تحليل المفردات، ويؤكد بأن تحليل مفردات اختبارات تستخدم بكم واسع، كاختبارات الاستعدادات أو التحصيل المنشورة تجارياً، يجب أن تعتمد على عينة ممثلـــة و كبيرة من المفحوصين. بالمقابل طالب الدكتوراه الذي يطور أداة لأطروحته، يجب أن يعتمد على عينة أصغر من هذه بكثير، وكقاعدة عامة يمكن حساب معظم معالم نظرية الاستجابة للمفردة بعينات مستقرة نسبياً، ومؤلفة من 200 مفحوص، ويعد هذا أقل عدد مرغوب فيه. وقانون آخر أثبت علمياً، هو أن يكون عدد المفحوصين من 5 إلى 10 أضعاف عدد المفردات. وحجوم العينات اللازم لحساب معالم المفردة في نظرية الاستجابة للمفردة عادة يتراوح ما بين 200 إلى 1000 مفحوص، وهذا يعتمد على النموذج المختار في التحليـــل. (ليندا وجيمس، 2009، ص.426)

وهو ما ينطبق على بحثنا الحالـــي، حيث تكونت العينة الاساسية من 577 تلميذاً وتلميذة، وهي عينة كافية للاستخدام في التحقق السيكـــومتري لاختبار مستويـــات التفكير الهندســـي، وفقاً للمقاربتين السيكومتريتين النظرية الكلاسيكية في القياس، ونظرية الاستجابة للمفردة الاختبارية (نموذج راش).

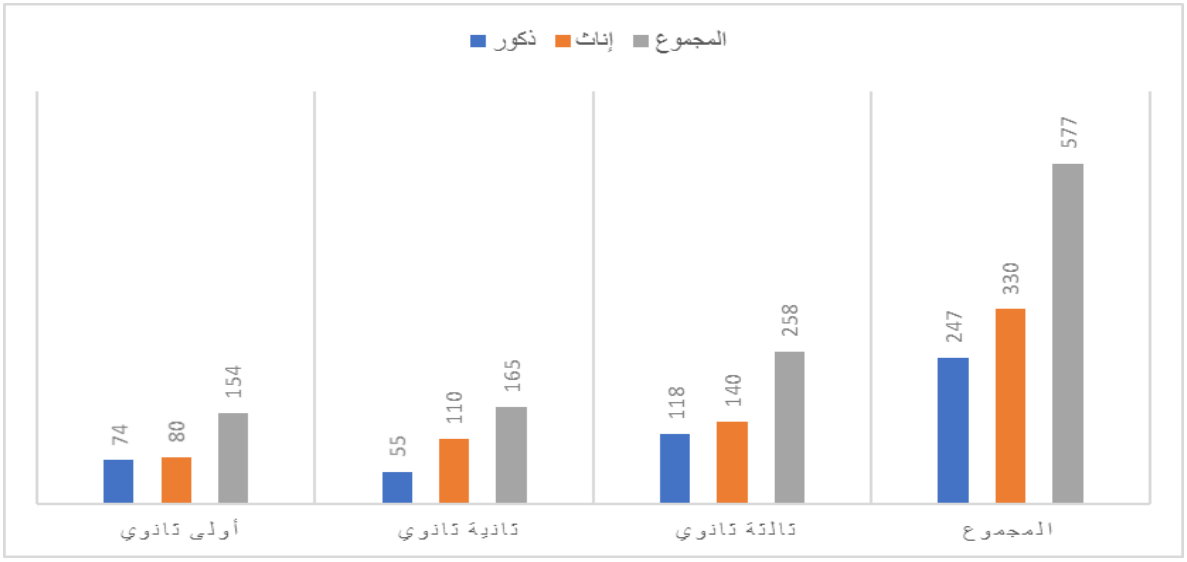
توزيع عينة البحث الأساسية بحسب متغيرى الجنس والمستوى الدراسي:

جدول (09)

يوضح عدد أفراد عينة البحث الأساسية ن=577 موزعين بحسب متغيرى الجنس و المستوى الدراســـي.

الجنس / المستوى	عدد الذكور	النسبة المئوية	عدد الإناث	النسبة المئوية	تعداد التلامذة لكل مستوى	النسبة المئوية
أولى ثانوي	74	%48.1	80	%51.9	154	%26.68
ثانية ثانوي	55	%33.3	110	%66.7	165	%28.59
ثالثة ثانوي	118	%45.7	140	%54.3	258	%44.71
المجموع	247	%42.8	330	%57.19	577	%100

يتضح من الجدول رقم (09) أن ما نسبته 57% من عينة البحث من الإناث، وما نسبته 43% من الذكور، وهذا يدل على أن غالبية عينة البحث إناث، وهو ما يتناسب مع حجم الإناث في المجتمع الاصلي للبحث. كما سجلنا أكبر نسبة وهي 45% من عينة البحث من تلامذة السنة ثالثة ثانوي، ويرجع الباحث ذلك إلى أن تلامذة الثالثة ثانوي كانوا أكثر فهما لتعليمات الاختبار، وجدية في الإجابـــة على مفردات الاختبار، وبالتالي فإن عملية حذف الاستمارات، والتي بلغت 105 استمارة إجابة محذوفة، لم تكن بالقدر الذي تم فيه حذف إجابات من تلامذة السنة الأولى والثانية ثانوي، وهو ما يوضحه الشكل رقم (21):



شكل (21) يوضح توزيع أفراد عينة البحث بحسب المستوى الدراسي والجنس.

توزيع عينة البحث بحسب متغير السن:

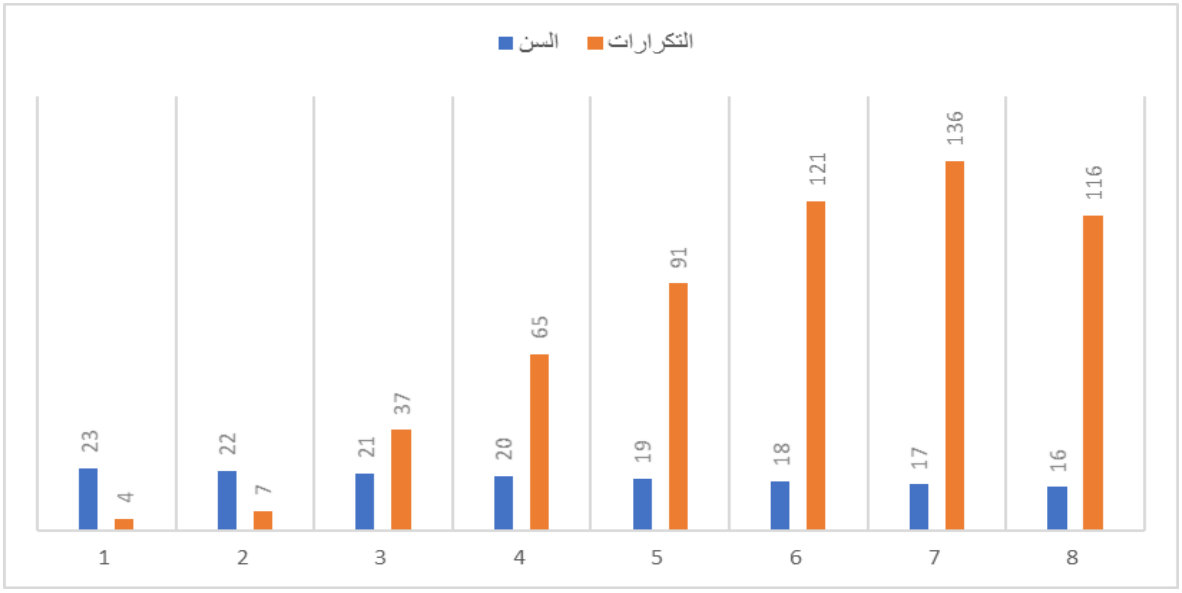
جدول (10)

يوضح توزيع أفراد عينة البحث بحسب متغير السن وتكراره.

السن	أولى ثانوي	ثانية ثانوي	ثالثة ثانوي	التكرارات	النسبة المئوية
23	0	0	4	4	%0.7
22	0	2	5	7	%1.2
21	0	2	35	37	%6.4
20	7	13	45	65	%11.3
19	11	16	64	91	%15.8
18	16	33	72	121	%21
17	48	72	16	136	%23.6

16	72	27	17	116	%20.1
المجموع	154	165	258	577	%100

يوضح الجدول رقم (10)، توزيع عينة البحث الأساسية بحسب متغير السن، حيث توزعت أعمارهم ما بين 16 سنة إلى 23 سنة، ونسبة التكرارات من 0.7 % إلى 23.6 %، حيث سجل التلامذة الذين سنهم 23 و أقل، تكرار بـــ 4 تلامذة، أي ما نسبته 0.7 % من العينة الكلية، بينما سجل التلامذة الذين سنهم 16 و أكبر، تكرار بلغ 116 تلميذاً وتلميذة، بنسبة تقدر بـــ 20.1 %. أما المتوسط الحسابي لاعمار أفراد العينة الكلية، كان 19.5 سنة، و بانحراف معياري 2.45، كما يوضحه الشكل رقم (22):



شكل (22) يوضح توزيع أفراد عينة البحث بحسب متغير السن وتكراره.

توزيع عينة البحث الأساسية بحسب الشعب الدراسية:

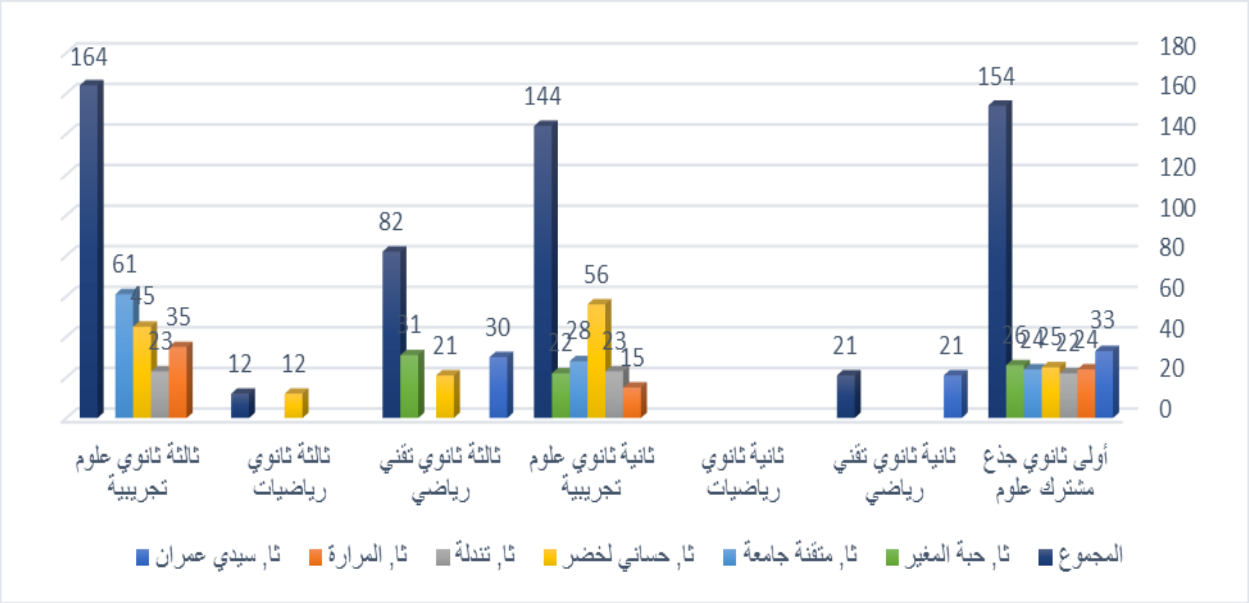
جدول (11)

يوضح توزيع عينة البحث حسب الشعب العلمية والمستويات الدراسية على الثانويات الستة.

الثانويـــــات	عدد التلامذة في الشعب العلميـــة و المستويات الدراسيـــة									المجمـــوع
	أولـــى ثانـــوي		ثانية ثانـــوي			ثالثـــة ثانـــوي				
	علـــوم جذع مشترك	رياضي تقني	رياضيات	علـــوم	رياضي تقني	رياضيات	علـــوم			
ثا. سيدي عمران	33	21	–	–	30	–	–			84
ثا. المـــرارة	24	–	–	15	–	–	35			74

ثا. تندلـــة	22	–	–	23	–	–	23	68
ثا. حساني لخضر	25	–	–	56	21	12	45	159
ثا. متقنة جامعـــة	24	–	–	28	–	–	61	113
ثا. حبة المجيد المغير	26	–	–	22	31	–	–	79
المجمـــوع	154	21		144	82	12	164	577

يتضح من الجدول رقم (11)، والذين يبين توزيع تلامذة عينة البحث الأساسيـــة بحسب الشعب الدراسية التي ينتمون إليها، في كل ثانوية من الثانويات الستة التي سحبت منها عينة البحث الأساسية، إلى أن تلامذة عينة البحث توزعوا على المستويات الثلاثة بالتساوي تقريباً، حيث بلغ عدد تلامذة الأولى ثانوي جذع مشترك علوم 154 تلميذا وتلميذة، وبلغ عدد تلامذة الثانية ثانـــوي 144 تلميذا وتلميذة، في حين بلغ عدد تلامذة السنة الثالثة ثانوي 164 تلميذا وتلميذة، وهي أعداد متقاربة، كما يتضح أن أغلب تلامذة عينة البحث ينتمون إلى شعبة العلـــوم التجريبية، والذين توزعوا كالتالي: 154 تلميذاً وتلميذة من السنة أولى ثانوي، و144 تلميذاً وتلميذة من السنة الثانية ثانوي، و164 تلميذاً وتلميذة بالنسبة لتلامذة السنة الثالثة ثانوي، والذين يمثلون ما مجموعه 462 تلميذاً وتلميذة أي ما نسبته 80% من عينة البحث الكليـــة، وهي نسبة تتناسب مع أعداد التلامذة في المجتمع الأصلـــي للبحث، والذين يغلب عليهم تلامذة شعب العلوم التجريبية. في حين سجلنا العدد 103 تلميذاً وتلميذة من شعبة تقني رياضي للسنة الثانية والثالثة ثانوي، أي ما نسبته 17.85% من عينة البحث الأساسية، وأقل عدد من التلامذة كان 12 تلميذاً وتلميذة من شعبة الرياضيات في السنة الثالثة ثانوي، والذين يمثلون ما نسبته 2.07% من العينة الأساسيـــة. والشكل رقم (23) يوضح توزيع عينة البحث بحسب الشعب العلميـــة (تقني رياضي؛ رياضيات؛ علوم تجريبية) والمستويات الدراسية التي ينتمون إليها في الثانويات الستة المسحوبة منها العينة الأساسية.



الشكل (23) **يوضح توزيع عينة البحث بحسب الشعب العلمية والمستويات الدراسية على الثانويات الستة.**

3- أداة البـــحث:

أداة البحث في بحثنا الحالي، هي عبارة عن اختبار تحصيلـــي موضوعي مرجعي المحك، تم بناءه لقياس مستويات التفكير الهندسي وفقا لنموذج (فان هيل) لتطور التفكير الهندســـي، والمتضمن مكونات المعرفة الهندسية الواردة في كتب مادة الرياضيات جزء الهندســـة، والتي تم تعلمها من قبل التلامذة خلال فترة دراستهم في المراحل الدراسية قبل الجامعية، أي المرحلة الإبتدائية، والمرحلة المتوسطة، والمرحلة الثانوية. اقتصرت عملية بنـــاء الاختبار على المستويات الأربعة الأولى المكونة لنموذج (فان هيل)، وهي (المستوى الادراكي أو البصري، والمستوى التحليلي أو الوصفي، والمستوى شبه الاستدلالي، والمستوى الاستدلالي (البرهان)).وسيتعرض الباحث لإجراءات وخطوات بناء وتصميم الاختبار، وكذلك التحقق من المؤشرات السيكومترية لمفرداته، والاختبار ككل، بأكثر تفصيل في الفصل القادم من هذا البحث.

4- خطوات تطبيق اختبار مستويات التفكير الهندســـي على أفراد عينة البحث:

للتحقق من الخصائص السيكومترية لاختبار مستويـــات التفكير الهندســـي ومفرداته، وفقاً للمقاربتين النظرية الكلاسيكية في القياس، ونظرية الاستجابة للمفردة الاختبارية (نموذج راش)، قام الباحث بإتخاذ مجموعة من الإجراءات لتطبيق الأداة (الاختبار) موضع البحث، وذلك بإتباع الخطوات التالية:

1. بعد مراسلة مصلحة ما بعد التدرج بقسم علم النفـــس وعلوم التربيـــة والأرطوفونيا بجامعة وهران2، وحصول الباحث على الترخيص بإجراء البحث الميداني من طرف مديرية التربية لولاية الوادي.

2. قام الباحث بالإتصال بمديري الثانويات، الذين تأكد من استعدادهم لتعاونهم معه في تطبيق الاختبار، بداية من الفصل الثاني للموسم الدراســـي 2016-2017، والذي امتد إلى نهاية السنة الدراسية، وإطلاعهم على كيفية تطبيق الاختبـــار، وتوضيح لهم ما هي الفئـــات المعنية بهذا الاختبار، وتحديد أوقات التطبيق المناسبة للتلامذة، مع مراعاة خصوصية الجدول الزمني لكل أستاذ.

3. قيام الباحث بطبع كراسات (كتيبات) الاختبار، وكذلك أوراق الإجابـــة، بالعدد الكافي وتحضير بعض الأدوات المساعدة، كأقلام الرصاص والممحاة لكل ممتحن.

4. اعد الباحث ورقة إجابة منفصلة عن كراس الاختبار، لتقليل الكلفة المادية على الباحث، وتسهيل عملية تصحيح الاختبار، وتضمنت ورقة الإجابة معلومـــات عن التلميذ (الإسم واللقب، تاريخ الميلاد، الجنس، القسم)، وتسلسل مفردات الاختبار، وبدائل الإجابة لكل مفردة اختبارية، أنظر الملحق رقم (04).

5. اختيرت الأقســـام الدراسية للشعب العلمية (علوم تجريبية؛ رياضيات؛ تقني رياضي)، التي طبق عليها الاختبار عشوائياً في كل ثانويـــة، تكونت العينة في البداية من 682 تلميذاً وتلميذة، وبعد استرجاع الباحث لأوراق الإجابـــة وتصحيحها، تقرر إلغاء 105 ورقة إجابة، نظراً لعدم الجدية في الإجابـــة، أو نقص في المعلومات، لتستقر العينة الأساسية في الأخير على 577 تلميذاً وتلميذة.

6. قيام الباحث بتطبيق الاختبار على عينة استطلاعية أولى مكونة من 34 تلميذاً وتلميذة من تلامذة السنة الأولى ثانـــوي بثانوية حسانـــي لخضر بجامعة، وذلك للتأكد من وضوح تعليمات الاختبار، ووضوح مفرداته، وسهولة فهمهما من قبل التلامذة، وتعديل المفردات غير الواضحة، وكذلك معرفة الزمن اللازم الذي سيستغرقه تطبيق الاختبار، ومعرفة الصعوبات التي قد تظهر أثناء التطبيق حتي يتم ضبطها وتلافيها عند التطبيق اللاحق للاختبار. حيث تم تقدير الزمن اللازم لـــلاختبار والذي بلغ 60 دقيقة، أي حصة دراسية كاملة، وذلك بجمع الزمن الذي استغرقه أول تلميذ لينهي الإجابة عن مفردات الاختبار، مع الزمن الذي استغرقه آخر تلميذ لينهي الإجابة عن مفردات الاختبار، وقسمة المجموع على إثنين.

7. طبق الباحث الاختبار على العينة الإستطلاعيـــة الثانية، المكونة من تلامذة ثلاثة أقسام دراسية (أولى ثانوي، ثانية ثانوي، ثالثة ثانوي)، وعددها 104 تلميذاً وتلميذة، المأخوذة من الثانويات الستة، حيث تم

توزيع كراســـات الاختبار عليهم وإعطائهم تعليمات في كيفية الإجابة على الاختبـــار، وطلب منهم كتابة البيانات الأولية على كراس الاختبار، وكذلك على ورقة الإجابـــة المستقلة المرفقة مع الاختبار، والتي تتضمن 28 مفردة وبدائلها الأربعة، كما طلب منهم وضع دائرة حول رمز الإجابـــة الصحيحة على كل مفردة، بعد أن تم شرح لهم طريقة الإجابة وتدعيمها بمثال، وبأن يجيبوا على جميع مفردات الاختبار، كما تم إعلامهم بالزمن المخصص للاختبار. وهذا للتحقق من مؤشرات صعوبة المفردات وتمييزها، والتحقق من فعالية المشتتات (المموهات)، وكذلك من فعاليـــة الاختبار ككل، من خلال التحقق من مؤشرات الثبات والصدق، باستخدام النظرية الكلاسيكية في القياس.

8. اشرف الباحث بنفسه على تطبيق الاختبـــار على العينة الأساسيـــة المكونة من 577 تلميذاً وتلميذة، موزعين على ستة ثانويات بمنطقة وادي ريغ بولاية الوادي، وبمساعدة بعض الأساتذة وبالتنسيق والتعاون مع مديري الثانويات، حيث زود كل تلميذ بكراس الاختبار وورقة إجابة كي يحل عليها، أنظر الملحق رقم (03 و04) على التوالي. كما حرص الباحث على تهيئـــة الجو المناسب للتلامذة قبل الإجابة، وحثهم على الجدية في الاجابة وعدم الاعتماد على الآخرين، وتوضيح لهم أن الاختبار من واقع ما تم دراسته في الجزء الخاص بالهندســـة في مادة الرياضيات، ابتدأ من السنة الخامسة إبتدائي إلى غاية السنة الثالثة ثانوي أي التعليم قبل الجامعي، وأن نتائج البحث تستخدم للأغراض العلمية فقط.

والغرض من هذه الدراسة، هو تحليل مفردات اختبار مستويات التفكير الهندســـي من خلال التحقق من مؤشرات الصعوبة والتمييز، وكذلك التحقق من فعالية الاختبار ككل، والمقارنة بين مؤشرات ثباته وصدقه باستخدام النظرية الكلاسيكية في القياس، و تدريجه وفق نظرية الاستجابة للمفردة باستخدام نموذج (راش) أحادي البارامتر.

9. تم إعطاء الأساتـــذة المتعاونين تعليمـــات بشأن التصرف في المواقف التي قد تحدث أثناء تطبيق الاختبار، وحثهم على عدم قراءة الأسئلة (المفردات) أو توضيح الصياغة للتلامذة.

10. قيام الباحث بمراجعة الاستمارات واستبعاد بعض أوراق الإجابة التي اتضح عدم جدية التلامذة في الاجابة على مفرداتها، مثل الاجابة على المفردات في أول الاختبار وترك البقية دون إجابة.

5- تجهيز بيانات اختبار مستويات التفكير الهندســـي للتحليل الإحصائـــي:

مرت عملية تجهيز بيانـــات البحث لمرحلة التحليل الإحصائي بثلاثة مراحل أساسية، وهي: مرحلة ادخال البيانات؛ ثم تصحيح استجابات التلامذة على مفردات اختبار مستويات التفكير الهندســـي؛ وانتهت بتجهيز ملف البيانات للتحليل الإحصائي وفق المقاربتين القياسيتين النظرية الكلاسيكة في القياس، ونموذج (راش)، وفيما يلي توضيح لكيفية تجهيز البيانات:

قام الباحث بعملية تحويل استجابات التلامذة بعد تطبيق اختبار مستويات التفكير الهندســـي على عينة التقنين، من كراسات الإجابـــة إلى الحزمة الإحصائيـــة للعلوم الإجتماعية (Spss v25)، وذلك بالاستعانة بالمفتاح الورقي المثقوب في التصحيح، ثم إعادة ترميز الاجابات لتصبح إما صفر للإجابة الخطأ، أو 1 للإجابة الصحيحة، وبهذا يكون مجموع اجابات التلامذة الصحيحة لمفردات الاختبار هي الدرجة الكلية للاختبار ولكل تلميذ، والتي تتراوح ما بين 0 و28، وفي حالة ترك التلميذ المفردة دون إجابة أو ضع أكثر من دائرة على رموز الإجابة فيعطى الدرجة صفر على هذه المفردة. ومن أجل تحديد مستويات التفكير الهندســـي لـــ(فان هيل) لكل تلميذ، تم جمع إجابات التلامذة حسب كل مستوى وتصحيحها حسب المعايير التالية:

1. يحقق التلميذ مستوى إتقان (المحك المتفائل)، إذا حصل على 5 إجابات صحيحة من 7 كحد أدنى، يعني أن يجيب التلميذ على خمسة من سبعة مفردات المكونة لكل مستوى، من اختبار مستويات التفكير الهندسي المبني وفق نموذج (فان هيل).
2. تحقيق المستوى الأول كحد أدنـــى كي يتم تصنيف التلميذ على مستويات (فان هيل)، وغير ذلك اعتبر التلميذ أنه غير مصنف.
3. تحقيق المستوى السابق للمستوى الذي يجري فحصه، فمثلاً، إذا تحقق المستوى الثالث عند تلميذ ما ولم يحقق المستوى الثاني لا يؤخذ بهذه النتيجة.

وبذلك أصبح ملف البيانـــات المصحح جاهزاً للتحليل، وفقاً للنظرية الكلاسيكيـــة في القياس، وذلك باستخدام البرنامج الإحصائي (spss v25)، ووفقاً لنظرية الاستجابة للمفردة (نموذج راش)، باستخدام برنامج وينستبس (winsteps v4.0.0) الذي قام بتطويره مايك لينكر (2017) في تدريج الاختبار. ويبدأ

التحليل بمصفوفة الاستجابات الإصلية (فرد/مفردة) أي، استجابة كل فرد على كل مفردة من مفردات الاختبار. وقد راعى الباحث في هذا بأن البرنامج يحذف الاستجابات التامة والصفرية.

6- الأساليب والبرامج الإحصائيـــة المستخدمة في تحليل بيانات البحث:

ولتحقيق أهداف البحث، وتحليل بياناته وفقاً للمقاربتين نظرية القياس الكلاسيكية، و نظرية الاستجابة للمفردة الاختبارية، باستخدام نموذج راش اللوغاريتمـــي أحـــادي البارامتر، استخدم الباحث عدة أساليب وبرامج إحصائية في تحليل البيانات، وهي:

1. الحزمة الإحصائيـــة في العلوم الإجتماعية (SPSS v25)، وذلك لعمل الإحصـــاءات المتعلقة بالنظرية الكلاسيكية في القياس، وقد استخدمت في ذلك الأساليب الإحصائية المناسبة التالية:

- القياسات الوصفية (التكرارات، النسب المئوية).
- قياسات النزعة المركزية (المتوسط الحسابي).
- مقاييس التشتت (الإنحراف المعياري).
- مؤشر صعوبة المفردة (Item Difficulty coefficient)؛ لاستخراج النسبة الإجمالية للإجابة الصحيحة لمفردة ما لمجموعة من المفحوصين.
- معادلة بيرنان The **B** index Approach ؛ لاستخراج:

- مؤشر التمييز أو حساسية مفردات الاختبار(Discrimination Index).

- معامل البدائل (المموهات) الخاطئة (Effectiveness of Destructors)؛ لاستخراج قيم البدائل الخاطئة لاختبار مستويات التفكير الهندســـي.
- معادلـــة ليفينجستون (Livingston Index (1972))، في استخراج معامل الثبات للاختبار المحكـــي لمستويات التفكير الهندســـي.
- معادلـــة لاوش (Lawshe, 1975)، لاستخراج: الصدق الوصفي أو (المحتوى).
- معامل إرتباط بيرسون (R)(Person correlation coefficient)؛ لاستخراج:

- الاتساق الداخلي بين درجة كل مفردة والدرجة الكلية لاختبار مستويات التفكير الهندســـي، ودرجة كل مفردة في كل مستوى من مستويات (فان هيل) الأربعة والدرجة الكلية لكل مستوى، والدرجة الكلية لكل مستوى مع الدرجة الكلية للاختبار الكلي.

– لحساب الصدق المحكي (التلازمي) بين درجات التحصيل الدراسي للتلامذة في مادة الرياضيات و درجاتهم على اختبار مستويات التفكير الهندســـي.

- معاملات الارتباط الثنائي المتسلسل أو الأصيل (ρ_{pbis}) Point Biserial Correlation Coefficient لحساب مؤشرات التمييز لمفردات اختبار مستويات التفكير الهندســـي.
- التحليل العاملي الاستكشافي (التحليل العاملي الموازي) بطريقة المكونات الأساسيـــة مع تدوير المحاور بطريقة الفاريمكس (varimax).
- اختبار ملاءمة المعاينة (MSA).
- اختبار بارتلت (Bartlett's Test)
- اختبار كاي مربع (χ^2)
- تقنية العامل المتسارع (Acceleratiion Factor).
- تقنية الجذر الكامن (Eigenvalues).
- تقنية الإحداثيات المثلى (Optimal Coordinates).
- معامل كودر– ريتشاردسون (KR_{20})، لاستخراج الثبات بطريقة الاتساق الداخلي لمفردات الاختبار.
- معامل ألفا كرونباخ (α).
- معامل سبيرمان– وبراون لتصحيح الطول.
- الخطأ المعياري للقياس.
- الاختبار التائي (T) للفروق بين متوسطي عينتين مستقلتين (الذكور والإناث).
- اختبار تحليل التباين الأحادي (One Way ANOVA) الفروق بين مستويات المرحلة الثانوية (أولى، ثانية، ثالثة ثانوي).
- الدرجات المعيارية والميئنيات و الدرجات التائية المعدلة.

2. برنامج (مايكروسوفت إكسال) (Microsoft Excel)، للاستعانة به في استخراج بعض الأشكال التوضيحية ، كما أن مخرجات برنامج (Winsteps v4.0.0) تكون من خلال هذا البرنامج .

3. البرنامج الإحصائـــي للنمذجة وفق نموذج راش اللوغاريتمي أحادي البارامتر (وينستبس) (Winsteps v4.0.0)، الإصدار (2017) والذي طوره مايكل لينكر (John Michael Linacre). الذي يستخدم طريقة

الأرجحية العظمـــى غير المشروطة (Un conditional maximum likelihood estimation (UCON)) أو طريقة تقدير الأرجحيـــة القصوى المشتركة (JMLE)، في إشتقاق المعادلتين الخاصتين بتقدير صعوبة المفردات، وقدرات الأفراد لتدريج اختبار لمستويات التفكير الهندســـي باعتماد نموذج راش، وكذلك المؤشرات الإحصائية التالية:

- تحليل نموذج (راش) للمكونات الأساسية المعتمدة على البواقـــي (Rasch-residual based Principal Components Analysis (PCAR))
- دالـــة استجابة المفردة Item Characteristic Curve (متوسط قيم تمييز (الميل، Slope))
- المؤشر الإحصائي Q3 لين (Yen,1984) للتحقق من الاستقلال الموضعي.
- الإنحدار غير الخطي أو الخط التقاربـــي الأدنـــى (Lower Asymptote) للتحقق من مؤشر التخمين.
- إحصائي المطابقة الداخلية (The Standardized Information Weighted Fit Statistics for Persons Infit ZSTD).
- إحصائي متوسط المربعات التقاربي (Mean Square Infit Statistics,MNSQ).
- إحصائي المطابقة الخارجية (The Standardized Information Weighted Fit Statistics for Persons Outfit ZSTD).
- إحصائي متوسط المربعات التباعدي (Mean Square Outfit Statistics, MNSQ).
- دالـــة معلومات الاختبار.
- دالـــة معلومات المفردات.
- وحدة اللوجيت (logites) ووحدة الواط (Watt).
- معامل ثبات الأفراد (Person Reliability).
- معامل ثبات المفردات (Item Reliability).
- معامل الفصل بين المفردات (Item Separation Index (GI)).
- معامل الفصل بين الأفراد (Item Separation Index (GP))
- الجذر التربيعي لمتوسط مربعات الأخطاء المعيارية للقياس (RMSE).

اضافة لهذا استخدم الباحث للكشف عن الأداء التفاضلي لمفردات اختبار مستويات التفكير الهندســـي المؤشرات الإحصائية التالية:

- الإحصائي اللامعلمي مربع كاي لــ مانتل–هانزل (Mantel-Haenzel).
- مؤشر لوغريتم نسبة الأرجحية التركمية لــ مانتل – هانزل (Mantel-Haenzel log Odds-Ratio MH-LOR) بإستخدام طريقة الإنحدار اللوجستي.
- إحصائي فرق المساحة بين منحنيات خصائص المفردة (Item Characteristic Curve, ICC) أو حجم التباين (DIF contrast).
- الإحصائي المعلمي اختبار راش– ويلش (Rasch-Welch t test).

المحكات التي تم بها الحكم على دقة القيــاس باستخدام برمجية وينستبس (Winsteps v4.0.0):

للحكم على دقة القيــاس اعتمد الباحث مجموعة من المحكات أثناء، وبعد عملية تدريــج اختبار مستويات التفكير الهندســـي، والتي قسمناها إلى نوعين:

- **المحكات التي تم مراعاتها أثناء التدريج:**

– إحصاءات الملاءمة لإستجابات الأفراد.

– إحصاءات تمييز وتخمين المفردات.

– إحصاءات ملاءمة المفردات.

- **المحكات التي تم مراعاتها بعد الإنتهاء من التدريج:**

– مؤشرات ثبات التدريج الخاصة بكل من المفردات والأفراد.

– التحقق من أحاديــة البعد.

– إرتفاع قيمة دالــة المعلومات الخاصة بالاختبار.

– إنخفاض قيمة الخطأ المعياري لصعوبة المفردات.

– إنخفاض قيمة الخطأ المعياري لقدرات الأفراد.

4. كما استعان الباحث ببرنامج (بيلوغ) (BILOG-MG3) من تطوير (Michele Zimowski, Eiji Muraki, Robert Mislevy & Da, 2002)، في استخراج شكل دالــة معلومات الاختبار، والخطأ المعياري للقدرة الكامنة للتلامذة على اختبار مستويات التفكير الهندســـي في صورته النهائية.

ملخص الفصل:

تم في هذا الفصل تحديد المنهج المتبع في البحث الحالـــي، وكذلك طريقة اختيار العينة وحجمها، وتوزيعها على المتغيرات الديمغرافية لعينة التلامذة، وكذلك التعريف المختصر بأداة البحث الرئيسة والتي ستكون محور حديثنا و بأكثر تفصيل في الفصل المقبل من هذا البحث. كما تم توضيح خطوات تطبيق اختبار مستويات التفكير الهندســـي وفق نموذج (فان هيل) على العينة الاستطلاعية الأولى والثانية، وكذلك على العينة الأساسية للبحث، ثم تجهيز بيانـــات البحث بعد تصحيحها وإدخالهـــا إلى الحزمة الإحصائية في العلوم الإجتماعية (SPSS v25) استعداداً لتحليلها وفقاً للنظرية الكلاسيكيـــة في القياس، ونظرية الاستجابة للمفردة، من خلال نموذج راش أحـــادي البارامتر باستخدام البرنامج الإحصائـــي وينستبس (Winsteps v4.0.0)، وبيلوغ (Bilog-mg3). واخيراً تم توضيح للأساليب والبرامج الإحصائية المستخدمة في البحث الحالي.

الفصـــل السابع

إجراءات بناء اختبار مستويـــات التفكير الهندســـي وفقاً لنموذج (فان هيل)

تمهيد:

يتضمن هذا الفصل إجراءات بناء اختبار مستويات التفكير الهندســـي وفقاً لنموذج (فان هيل)، من حيث توضيح خطوات كيفية بناء الاختبار وفقاً لنموذج (فان هيل)، ثم عملية التحليل الإحصائي لمفرداته بعد تطبيقه على عينـــة استطلاعية من تلامذة المرحلة الثانوية، والتحقق من مؤشرات الصعوبة والتمييز أو حساسية مفرداته، ثم تحليل فعالية بدائله (مموهاته)، وتقويم صدقه التقاربي والتباعدي (التمييزي) لمفرداته، ثم تقييم فعالية اختبار مستويات التفكير الهندســـي ككل من خلال التحقق من صدقه وثباته، قبل تطبيقه على العينة الأساسية للبحث.

1- الخطوات الإجرائية لبناء اختبار مستويات التفكير الهندســـي:

1-1- الإطلاع على بعض اختبارات التفكير الهندســـي المبنية وفقاً لنموذج (فان هيل):

لتحقيق أهداف البحث كان لابد للباحث من الإطلاع على النظرية التي استند اليها البحث الحالي في بناء الاختبار، وكذلك على الأدبيات والبحوث والدراسات السابقة ذات الصلة بموضوع البحث، حيث ترى أبو علام (2006) في هذا السياق بأن: " مراجعة الأدبيات والبحوث والدراسات السابقـــة، تساعد في أن يكتسب الباحث بصيرة بموضوع البحث وفهمه لأسباب ما يوجد في مجال البحث من تناقضات، فضلاً عن مساعدة الباحث على تحديد وضع دراسته تاريخياً، وكيف أو ماذا سيضيف بحثه للمعرفة القائمة" (ص.93)

ومن بين أهم الدراسات التي تم الركيز عليها من قبل الباحث دراسة يوسيسكين، التي اجراها بمشروع جامعة شيكاغو (1982)، والرائدة في هذا الموضوع، والتي كان من بين أهم أهدافها بناء أول اختبار لقياس مستويات التفكير الهندسي وفق نموذج (فان هيل) لتطور التفكير الهندســـي، إضافة إلى ذلك تواصل الباحث مع يوسيسكين شخصياً أنظر الملحق رقم (09)، واستشارته في بعض خصوصيات بناء الاختبار موضوع البحث في البيئة الجزائرية، في محاولة للاستفادة من خبرتـــه الرائدة في هذا المجال.

1-2- خطوات بناء اختبار مستويات التفكير الهندســـي:

قام الباحث بعدة خطوات في عملية تصميم وبناء الاختبار التحصيلي المرجعـــي المحك المبني وفق نظرية (فان هيل) للتفكير الهندســـي. بعد الإطلاع على التراث العلمي وعديد الدراسات والبحوث السابقة في هذا المجال على غرار اختبار (فان هيل) المطور من طرف يوسيسكين في دراسة مشروع جامعة شيكاغو (1982)، و المترجم إلى اللغة العربية من قبل جهاد الشويخ في دراسته (2005)، و كذلك دراسة

رفاء جمال الرمحي (2006)، وميمي السيد (2014)، و بوسالم (2014)، وأبو ناهية صلاح الدين (1994)، و(Larry R , 2017). جاءت الخطوات كالتالي:

أولاً: مرحلـــة التحليـــل: وتتكون من الخطوات التالية:

1. تحديد الهدف من الاختبـــار، وهو إعداد اختبار لقياس قدرة تلامذة المرحلة الثانوية للمعرفة الهندسية، وكيفية تعلم الهندســـة وفقاً لنظرية مستويات التفكير الهندســـي الأربعة لـــ(فان هيل)، وهي (الإدراكي أو البصري، التحليلي أو الوصفي، شبه الاستدلالي، الاستدلالي)[1]. وبالتالي فهو اختبار تحصيلي محكي المرجع، تقويمي وتشخيصي للتفكير الهندســـي لدى التلامذة مبني وفق نظرية أو نموذج (فان هيل).

2. تحديد المهارات والكفايات الهندسية الرئيسة التي يتطلبها كل مستوى مستويات فان هيل للتفكير الهندسي، وتصنيفها إلى أربعة مهارات أوكفايات، اعتمد الباحث في تحديد المهارات الرئيسة لكل مستوى من مستويات التفكير الهندســـي على نظرية (فان هيل) في تطور مستويات التفكير في الهندسة، والتي يفترض فيها وجود خمسة مستويات للتفكير في الهندســـة، وهي:(الادراكي أو البصري، التحليلي أو الوصفي، شبه الاستدلالي، الاستدلالي).(Van Hiele, 1986, p.25)

اقتصرت عملية بناء الاختبار على الأربعة مستويـــات الأولى لنموذج (فان هيل)، بناءً على نتائج عدة دراسات سابقة في هذا الموضوع، من أهمها دراسة يوسيسكين (1982) في مشروع جامعة شيكاغو، حيث كان من أبرز نتائجه، أن المستوى الخامس يشكل معضلة بالنسبة للنظرية، وإلغائه يعطي نتائج أفضل، فهو إما غير موجود أو لا يمكن قيـــاسه أو فحصه. (الشويخ، 2003، ص.19)

3. **صياغة الأهداف:** تم صياغة أربعة أهداف رئيســـة، وكل هدف يمثل الكفاية أو المهارة اللازمة في كل مستوى من مستويات (فان هيل) الأربعة الأولى.

4. **صياغة أهداف سلوكية مساعدة:** ويرى في هذا السياق صلاح الدين علام (1990) أن كل هدف سلوكي يتطلب أن يعرف تعريفاً إجرائياً دقيقا من خلال مجموعة متكاملة من المفردات الاختبارية التي تم بنائها، بحيث تشمل جميع الجوانب المتعلقة بالمتغير (الهدف السلوكي)، حيث أن كل هدف من الأهداف السلوكية

---

[1] – استخدم يوسيسكين (Usiskin,1982) في دراسته في مشروع جامعة شيكاغو ترقيم مستويات فان هيل من 1 إلى 5 للتعبير عن مستويات فان هيل الأصلية الخمسة، بدلا من ترقيم النظرية الأصلي من 0 إلى 4.

المحددة للوحدة الدراسية المختارة، يمثل على فترة من الميزان الخطـــي بمجموعة المفردات التي تقيسه. (ميمي السيد، 2014، ص.133)

لهذا قام الباحث بدراسة كل هدف سلوكـــي رئيس على حدة دراسة دقيقة، ثم تحديد لكل هدف رئيس أهداف سلوكية مساعدة والتي عددها 10 أهداف سلوكية، تمثل كل مستوى من المستويات الأربعة الأولى لنظرية (فان هيل) للتفكير الهندســـــي.

5. اعتمد الباحث في صياغة مفردات اختبار مستويات التفكير الهندســـي، على السندات التكوينية الموجهة لنمطي التفتيش والتعليم للمراحل التعليمية الثلاثة (الإبتدائي، والمتوسط، والثانوي) في مادة الرياضيات والجزء الخاص بالهندسة، والتي يصدرها المعهد الوطني لتكوين مستخدمي التربية وتحسين مستواهم بوزارة التربية الوطنية، وكذلك تحليل محتوى كتب مادة الرياضيات لهذه المراحل، والتي شملت المواضيع التالية:

1. هندسة فيثاغورس والهندسة الإقليدية.
2. الإنشاء الهندسي (تضعيف المكعب، تثليث الزاوية، تربيع الدائرة).
3. الهندسة الاسقاطية.
4. الهندسة غير الاقليدية.
5. الهندسة الكسورية.
6. مسلمات الهندسة الاقليدية (المسلمات الخمسة والمسلمة السادسة والبديهيات الخمسة).
7. النقطة والمستقيم والمستوي.
8. المنحنيات.
9. الإنشاء الهندسي: المثلث متساوي الأضلاع، المثلث القائم، المثلث المتساوي الساقين، المتوازي الإضلاع علم مركزه ورأسان من رؤوسه، المعين، المستطيل علم طوله وقطره، شبه المنحرف، المثلث ABC علمت فيه زاوية وضلع مجاور ومجموع الضلعين الآخرين، مثلث علم رأسان من رؤوسه الثلاثة ومركز ثقله. محور قطعة مستقيمة، مستقيم يعامد مستقيماً معلوماً عند نقطة معلومة، مستقيم يعامد مستقيم معلوماً K عند النقطة R لا تقع علىK ، منصف زاوية، زاوية متساوية لزاوية معلومة، مستقيم يمر بنقطة معلومة ويوازي مستقيماً معطى، إنشاء المماسين لدائرة معطاة (مع مركزها)، إنشاء مركز دائرة.
10. الزوايا: حول قياسات الزوايا ووضعياتها، إنشاء زوايا شهيرة.

11. المثلثات: المتساوية الساقين، المتساوية الأضلاع، القائمة.

12. المستقيمات والدوائر الخاصة في المثلث: المتوسطات، الإرتفاعات، المنصفات في المثلث، المحاور، الدائرة المحيطة بالمثلث والمثلث المحيط بالدائرة.

13. التناظر: التناظر المركزي، التناظر المحوري، أشكال لها محاور تناظر.

14. المضلعات: المثلثات، الرباعيات، المضلعات المنتظمة، المثلث المتساوي الأضلاع، المربع، الخماسي المنتظم، السداسي المنتظم، السباعي المنتظم، الثماني المنتظم.

15. إنشاء بعض المضلعات المنتظمة: المربع، السداسي المنتظم، الثماني المنتظم.

16. المجسمات: مسلمات الهندسة الفضائية.

17. بعض خواص التوازي والتعامد.

18. تقاطع مستوى مع متوازي مستطيلات وأسطوانة.

19. مساحات وحجوم بعض المجسمات.

ثانياً: مرحلـــة البنـــاء:

1. صياغة مفردات اختبار مستويات التفكير الهندســـي:

تم صياغة 40 مفردة اختبارية موزعة على أربعة مستويات الأولى لنموذج (فان هيل)، تماشياً مع طبيعة المجال السلوكي أو فكرة المستوى حسب (فان هيل) للتفكير الهندســـي، لتكون منسجمة مع التعريف النظري لكل منها، وقد أخذ الباحث بعين الاعتبار الأهداف التي يستخدم من أجلها اختبار مستويات التفكير الهندســـي وخصائص مجتمع الاختبار. وقد صيغت المفردات بحيث تتلاءم مع مستويات التلامذة ومستوى قدرة إستيعابهم العقليـــة، حيث حددت بـــ 40 مفردة موزعة على مستويات التفكير الهندســـي الأربعة لـــ(فان هيل)، وهي: (الإدراكي أو البصري؛ التحليلي أو الوصفي؛ شبه الاستدلالي؛ الاستدلالي)، كالتالي:

المستوى الأول: المستوى الإدراكـــي أو البصري 10 مفردات (التعرف على الأشكال من مظهرها العام).

المستوى الثاني: المستوى التحليلـــي أو الوصفي 10 مفردات (معرفة خصائص الشكل) .

المستوى الثالث: المستوى شبه الاستدلالـــي 10 مفردات (العلاقات أو الاستنتاج غير الرسمي).

المستوى الرابع: المستوى الاستدلالـــي 10 مفردات (الاستنتاج الرسمي/ الشكلي أو الإثبات (البرهان)).

2. تحديد نوع مفردات الاختبار:

تعتبر مفردات اختبار الاختيار من متعدد من أفضل أنواع الاختبارات وأكثرها صدقاً وثباتاً واستخداماً في الاختبارات التحصيلية في وقتنا الحاضر، والتي يختار فيها الفرد الاستجابة الصحيحة أو الأفضل من عدة اختيارات، حيث تتكون من جذر في شكل عبـــارة غير كاملة أو سؤال أو رسم تخطيطي أو حساب أو تعليمات مطلوب تنفيذها، وعدد من الاستجابات (اختيارات أو بدائل)، وإحدى هذه الاستجابات المقترحة صحيحة، بينما باقي الاستجابـــات غير صحيحة أو متداخلة مع الاستجابة الصحيحة، ويطلب من المتعلم اختيار الإجابة الصحيحة. (بوسالم، 2008، ص.126)

واعتماداً على ما سبق، والدراسات والبحوث السابقة في هذا الموضوع، قام الباحث ببناء مجموعة من المفردات الاختبارية من نوع اختيار من متعدد، حيث أن كل مجموعة من المفردات تقيس أهداف سلوكية معينة بحسب كل مستوى من مستويات (فان هيل) الأربعة الأولى. وقد أخذ بعين الاعتبار الأسس الفنية في كتابة المفردات ومطابقتها للهدف السلوكي لكل مستوى من مستويات التفكير الهندسي الأربعة. اشتملت كل مفردة من مفردات الاختيار من متعدد على أربعة خيارات (A ,B,C,D)، إحداها يمثل الإجابـــة الصحيحة على المفردة. بلغ عدد المفردات الاختبارية التي تمت صياغتها في بادئ الأمر 40 مفردة اختبارية موزعة على أربعة اختبارات فرعية (مستويات). ويضيف فاروق عبد الفتاح موسى (1990) في هذا السياق، بأن اختبارات الاختيار من متعدد من أكثر أنواع الاختبارات الموضوعية انتشاراً، فهي تناسب أكثر المخرجات التربوية أهمية مثل: المعرفـــة والفهم والحكم والقدرة على حل المشكـــلات والتوصية بأعمال معينة وعمل تنبؤات.(ميمي السيد، 2014، ص.133)

ويرى الباحث إن استخدام مفردات الاختيار من متعدد يناسب قياس الأهداف السلوكية التي تهدف إليها نظرية مستويات التفكير الهندسي لـــ(فان هيل)، كما أن هذا النوع من المفردات الاختبارية يتناسب مع أساليب التحليل المتعلقة بنماذج الاستجابة للمفردة، حيث أنها تحقق شروطاً وافتراضات لهذه النماذج.

3. تقييم المفردات الاختباريـــة:

بعد صياغة الباحث لمفردات الاختبـــار، اعد استمارة لتقييم هذه المفردات، اشتملت الإستمارة على التعريف بنظرية مستويات التفكير الهندســـي لـــ(فان هيل)، وتعريف مفصل بالأهداف السلوكية لكل مستوى من مستويات (فان هيل) الأربعة الأولى المعنية بالاختبـــار، حيث يقابل كل مستوى من مستويات التفكير

الهندسي مجموعة المفردات الاختبارية وعددها 10 مفردات لكل مستوى. قام الباحث بعرضها أنظر الملحق رقم (01) على 13 محكماً أو خبير موزعين بين مفتشين وأساتذة ذوي خبرة ومن أساتذة جامعيين حاملين لدرجة الدكتوراه والماجستير تخصص الرياضيات، استجاب منهم 10 فقط، حيث أعطى لكل محكم نسخة من الاستمارة المعدة للتحكيم بالإضافة إلى الاختبار، وطلب منهم القراءة الجيدة لتعريف النظرية المرفق مع مستوياتها الأربعة، ثم قراءة كل مفردة من مجموعة المفردات الاختبارية التي تقابل كل مستوى من المستويات الأربعة للتفكير الهندسي وفق نموذج (فان هيل)، ثم إبداء الرأي حولها من خلال الخيارات الثلاثة المطروحة أمام كل مفردة من المفردات 40 المقترحة، وهذا بوضع علامة (X) في العمود الذي يراه مناسباً: (أ) رفض المفردة؛ (ب) تعديل المفردة؛ (ج) قبول المفردة.

كما طلب الباحث من كل محكم أيضاً كتابة أي ملاحظات أو إقتراحات يراها مناسبة بشأن تعديل أي مفردة اختبارية بغرض إثراء المفردات الاختبارية وتحسين كفاءتها، وبعد مراجعة آراء المحكمين تم وضع الاختبار في صورته الأولية، حيث نتج عن هذه العملية استبعاد 12 هدفاً سلوكياً مساعداً، والإبقاء على 28 هدفاً سلوكياً بواقع 7 أهداف سلوكية مساعدة لكل مستوى من مستويات (فان هيل) الأربعة الأولى. أنظر ملحق رقم (01)، وقد أظهر المحكمين نسبة إتفاق مقدارها 74.74% باستخدام معادلة لاوش (1975)، حيث أن عدد المفردات الاختبارية المتحصل عليها كان كافي لمثل هذه الاختبارات التي تبنى وفق نظرية معينة. والجدول رقم (12) يوضح الهدف السلوكي لكل مفردة من مفردات كل مستوى من مستويات (فان هيل) الأربعة الأولى، و الخيارات المقترحة للمحكمين حول الحكم على كل مفردة وهدفها السلوكي، وكذلك نسبة الموافقة على كل مفردة من مفردات اختبار مستويات التفكير الهندسي 40 مفردة المقترحة.

جدول (12)

يوضح تقويم المحتوى والأهداف السلوكية لكل مفردة من مفردات اختبار مستويات التفكير الهندسي في ضوء أراء المحكمين باستخدام معادلة لاوش Lawshe (1975) .

المستوى	الهدف السلوكي لكل مفردة	رفض	تعديل	قبول	CVR_s	القرار
المستوى البصري أو الإدراكي	1. التعرف على المضلع	–	2	8	%60.00	يحتفض بها
	2. التعرف على الدائرة	5	3	3	%40.00-	ترفض
	3. التعرف على المربع	–	–	10	%99.00	يحتفض بها
	4. التعرف على المضلع المنتظم	1	3	6	%20.00	ترفض
	5. التعرف على المثلث	–	1	9	%80.00	يحتفض بها

	6. التعرف على متساوي الأضلاع	1	1	8	%60.00	يحتفض بها
	7. التعرف على متوازي الأضلاع	–	–	10	%99.00	يحتفض بها
	8. التعرف على المثلث متساوي الساقين	–	2	8	%60.00	يحتفض بها
	9. التعرف على المستطيل	1	–	9	%80.00	يحتفض بها
	10. التعرف على المكعب	3	1	6	%20.00	ترفض
النتائج	عدد المفردات المحتفظ بها 7 مفردات.				CVR_s = 76.85%	
المستوى التحليلي أو الوصفي	11. خصائص المعين	6	1	3	%40.00–	ترفض
	12. خصائص متوازي الأضلاع	–	–	10	%99.00	يحتفض بها
	13. خصائص المستطيل	1	4	5	%00.00	ترفض
	14. خصائص متوازي الأضلاع	–	2	8	%60.00	يحتفض بها
	15. خصائص المثلث المتساوي الأضلاع	1	2	6	%20.00	ترفض
	16. خصائص شبه منحرف متساوي الساقين	–	–	10	%99.00	يحتفض بها
	17. خصائص الرباعي	–	1	9	%80.00	يحتفض بها
	18. خواص المربع	–	–	10	%99.00	يحتفض بها
	19. خواص زوايا مركزية	1	1	8	%60.00	يحتفض بها
	20. تحويلات النقطية	1	–	9	%80.00	يحتفض بها
النتائج	عدد المفردات المحتفظ بها 7 مفردات .				CVR_s = 82.42%	
المستوى شبه الإستدلالي	21. علاقة مستطيل بمتوازي الأضلاع	1	1	8	%60.00	يحتفض بها
	22. خصائص شكل رباعي ناتج عن تقاطع دائرتين	–	1	9	%80.00	يحتفض بها
	23.علاقة الرباعي الدائري بالمربع والمستطيل ومتوازي الأضلاع وشبه المنحرف المتساوي الساقين	4	1	5	%00.00	ترفض
	24. إستدلال منطقي حول المستطيل والمثلث	–	2	8	%60.00	يحتفض بها
	25. علاقة متوازي الأضلاع بالمعين	–	–	10	%99.00	يحتفض بها
	26. المربع والمستطيل ومتوازي الأضلاع	6	1	3	%40.00–	ترفض
	27. علاقة المستطيل بالمربع	–	–	10	%99.00	يحتفض بها
	28. علاقة المعين المربع متوازي الأضلاع المستطيل	3	–	7	%40.00	ترفض
	29. علاقة زوايا نتيجة تقاطع ثلاث مستقيمات متوازية	–	2	8	%60.00	يحتفض بها
	30.علاقة زوايتان متقابلتان بالرباعي الدائري والمعين والمستطيل ومتوازي الأضلاع	1	1	8	%60.00	يحتفض بها
النتائج	عدد المفردات المحتفظ بها 7 مفردات .				CVR_s = 74%	

المستوى الإستدلالي	31. إستنتاج حول التناسب لـــ"طاليس"	–	1	9	%80.00	يحتفض بها
	32. إستنتاج حول المثلث القائم	1	4	5	%00.00	ترفض
	33. إثبات حول المثلث المتساوي الساقين	1	1	8	%60.00	يحتفض بها
	34. إثبات حول متوازي الأضلاع	2	–	8	%60.00	يحتفض بها
	35. إثبات تشابه مثلثين	6	1	3	%40.00-	ترفض
	36. إثبات حول الشكل الرباعي الدائري	–	2	8	%60.00	يحتفض بها
	37. إستنتاج حول المثلث المتقايس الأضلاع	1	1	8	%60.00	يحتفض بها
	38. إثبات تساوي أضلاع مثلث محصور بين دائريتين داخلية وخارجية	7	1	2	%60.00-	ترفض
	39. عبارات منطقية حول خصائص المربع والمعين والقطرين	–	1	9	%80.00	يحتفض بها
	40. إثبات حول المستطيل وقطريه	–	2	8	%60.00	يحتفض بها
النتائج	عدد المفردات المحتفظ بها 7 مفردات.				CVR_s = 65.71%	

* (CVR_s) نسبة صدق المحتوى لكل مفردة.

وبعد استرجاع الباحث لاستمارات التحكيم من المحكمين، الذين اقترحوا حذف المفردات ذات الأفكار المكررة والمتشابهة مثل المفردة 4 "التعرف على المضلع المنتظم"، والمفردة 23 "علاقة المربع بالمستطيل ومتوازي الأضلاع" في المستوى الأول والثالث على التوالي، وكذلك المفردات التي لا تقيس الهدف السلوكي بحسب نظرية (فان هيل) بشكل مباشر مثل المفردة 10 "التعرف على المكعب" حيث أن أفكار (فان هيل) كانت مبنية أساساً على الهندســـة المستوية وليس الهندســـة الفضائية (ثلاثية الأبعاد). قام الباحث بمراجعة المفردات الاختبارية التي أشار إليها المحكمين بالتعديل، وحذفت المفردات التي اقترح المحكمين حذفها، حتى صار عدد المفردات الاختبارية 28 مفردة ، أي 7 مفردات في كل مستوى من مستويات (فان هيل) الأربعة المكونة للاختبار، ثم اعيدت لكل محكم استمارة التحكيم السابقة والمعدلـــة وفق ملاحظتهم الأولى، وطلب منهم الإجابة عن نفس الخيارات الثلاثة السابقة. ووجد الباحث نسبة الإتفاق بين المحكمين حول إجابة الخيارات السابقة، فكانت كما يوضحها الجدول رقم (13):

جدول (13)

يلخص نتائج مؤشر صدق المحتوى (CVIs).

المستوى	العدد الإجمالي للمفردات	عدد المفردات المحتفظ بها.	مؤشر CVIs
المستوى البصري	10	7	%77.14
المستوى التحليلي	10	7	%82.85
المستوى شبه الاستدلالي	10	7	%74.28
المستوى الاستدلالي	10	7	%65.71
المجموع	40	28	%74.74

يتضح من الجدول رقم (13)، أن نسبة قبول المفردات والتي تعبر على قيمة مؤشر صدق المحتوى للاختبار تساوي 0.74 وهي قيمة جيدة ومقبولة، ودالـــة عند مستوى 0.05 حسب (لاوش) . استخدم الباحث لحسب صدق المحتوى أو الصدق الوصفـــي المعادلة رقم (29) لـــ لاوش (1975) Lawshe:

$$CVRs = \frac{n_e - N_2}{N_2} \quad \ldots\ldots\ldots\ldots\ldots\ldots (29)$$

حيث:(n_e) عدد المحكمين (الخبراء) الذين أجمعوا على أن المفردة (مقبولة)؛ و (N_2) العدد الإجمالي للمحكمين مقسوم على إثنين. وحسب لاوش (1975) فإن القيمة المقبولة لـــمؤشر (CVR_s) في حالة 10 محكمين هي القيمة 0.62. كما أعطى (لاوش) بعض الإرشادات، كالتالي:

1. عندما يعبر أقل من نصف المحكميـــن بأن المفردة (مقبولة) فإن قيمة (CVR_s) تكون سالبة.
2. عندما يعبر نصف المحكميـــن بأن المفردة (مقبولة) والنصف لآخر بأنها غير مقبولة فإن قيمة (CVR_s) تكون صفر.
3. عندما يعبر كل المحكميـــن على أن المفردة (مقبولة) تكون قيمة (CVR_s) 1.00 ولذلك يتم تعديلها إلى القيمة 0.99.
4. عندما يعبر أكثر من نصف المحكمين على أن المفردة (مقبولة) فإن قيمة (CVR_s) تكون بين الصفر و 0.99. (Lawshe, 1975, pp.563-575).

قام الباحث وللمرة الثانية بعد إجراء جميع التعديلات التي اقترحها المحكمين، بعرض مفردات الاختبار مرة أخرى بعد إجراء التعديلات عليها على بعض من المحكمين أنفسهم مرة أخرى فوافقو عليها وأقروها.

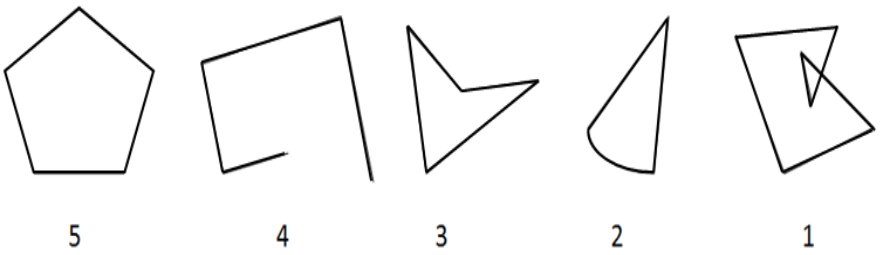
وبهذا أصبحت مواصفات الاختبار جاهزة، وهي موضحة في الملحق رقم (03)، وفيما يلي عرض لجزء من محتوى سلوكي الخاص بالمستوى الأول (البصري أو الإدراكـــي)؛ والمستوى الرابع (الاستدلالي) لمستويات (فان هيل) للتفكير الهندســـي:

- **الهدف الإجرائي للمستوى الأول (البصري أو الإدراكي):**

يتعرف التلميذ فيه على المضلع من خلال مجموعة من الأشكال الهندسيـــة المقترحة.

عينة المفردة:

- **المثال الأول:** أي من الأشكال التالية مضلع؟



A). 1 و3 و5 فقط.

B). 3 و5 فقط.

C). 2 و4 فقط.

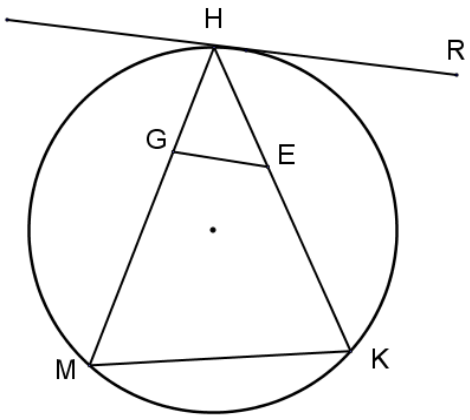
D). جميعها مضلعات.

- **الهدف الإجرائي للمستوى الرابع (الاستدلالي):**

إثبات أو برهان حول الشكل الرباعـــي الدائري.

عينة المفردة:

- **المثال الثاني:** إليك المثلث HKM مثلث مرسوم داخل دائرة و (HR) مماس لها . المستقيم (EG) يقطع (HK) في E ، و (HM) في G. من المعطيات السابقة **يمكن إثبات** أن الشكل EKMG رباعي دائري **إذا كان:**



A). HE = HG .

B). HR // GE .

C). KM // EG .

D). KM = RH .

4-تحديد درجة القطع (المحك أو المعيار):

تبنى الباحث وحدد محك لكل مستوى (مستوى إتقان)، بناءً على دراسات وأبحاث سابقة في هذا المجال وخاصة دراسة يوسيسكين باعتبارها من الدراسات الرائدة في هذا المجال، في تحديد الدرجة الفاصلة للتمييز بين المتقنين وغير المتقنين نظراً لكفائتها وسهولـــة فهمها ويسر تطبيقها، والتي استخدمها بمشروع جامعة شيكاغو (1982)، حيث استخدم هذا الأخير عند تصميم اختبار لمستويات التفكير الهندســـي وفق نموذج فان هيل 25 مفردة في خمسة مستويات، و 5 مفردات لكل مستوى، محك أو معيار 5/3 أي 60%، وكذلك معيار 5/4 أي 80%، ويرى يوسيسكين أن اختيار أي من المحكين أو المعيارين يتم استناداً لطبيعة الاختبار، وما يرغب فيه الباحث في التقليل من الخطـــأ الأول أو الخطـــأ الثاني، حيث يمكن الوقوع في الخطأ الأول وهو الحكم على أن التلميذ حقق مستوى الإتقان أو المحك في كل مستوى وهو في الحقيقة لم يحققه، أما الإحتمال الوقوع في الخطأ الثانـــي، وهو الحكم على أن التلميذ لم يحقق مستوى الإتقان أو المحك وهو في الحقيقة حققه، والذي وضحه بالشكل التالي:

1. احتمال (P) الإجابة الصحيحة على 3 من 5 مفردات عن طريق التخمين العشوائي = 0.05792 عند مستوى دلالة إحصائية 0.01، أو بما يسمى المحك المتفائل.

2. احتمال (P) الإجابة الصحيحة على 4 من 5 مفردات عن طريق التخمين العشوائي= 0.00672 عند مستوى دلالة إحصائية 0.001، أو بما يسمى المحك الصارم . (Usiskin, 1982, p.23)

واستخدم يوسيسكين كذلك اضافة إلى المحك 5/3 و5/4[2]، لتقييم التلميذ من خلال مجموع الدرجات الموزونة أو المرجحة، وفق التالي : (أ) درجة 1 عند تحقيق محك المستوى الأول المفردات من 1–5؛ (ب) درجة 2 عند تحقيق محك المستوى الثاني المفردات من 6–10؛ (ج) درجة 4 عند تحقيق محك المستوى الثالث المفردات من 11–15؛ (د) درجة 8 عند تحقيق محك المستوى الرابع المفردات من 16–20؛ (هـــ) درجة 16 عند تحقيق محك المستوى الخامس المفردات من 21–25. حيث أن القيمة أو الدرجة الكلية الموزونة (المرجحة) تسمح للشخص بتحديد المستويـــات التي تم التوصل إليها من خلال محـــك الدرجات

[2] – استخدم (فان هيل) معيارين: (3) من (5) الذي يقدم صورة متفائلة عن مستوى المتعلم، أو معيار (4) من (5) المحافظ أو الصارم الذي يقلل من فرصة أن يكون المتعلم قد وصل إلى المستوى عن طريق التخمين، أو معيار (5) من (5) الإتقان الكامل.

الموزونة (المرجحة) وحدها، فعلى سبيل المثال: الدرجة 19 تشير إلى أن التلميذ حقق المحك على المستويات الأول والثاني والخامس وبهذه الطريقة فإن الدرجات تتراوح ما بين 0 و31.

واعتماداً على ما سبق تبنى الباحث في البحث الحالي المحـــك المتفائل 7/5 أي 71%، والمحك الصارم 7/6 أي 86% وهي تقريباً نفس المحكات التي اعتمد عليهم يوسيسكين في عملية تطوير اختبار لمستويات التفكير الهندســـي لفان هيل، إلا أن الباحث قام بتكييف هذه المحكات مع طبيعة تكوين مفردات الاختبار الحالي، حيث يحتوى كل مستوى من المستويـــات الأربعة للاختبار على 7 مفردات. كما أن احتمال الإجابة الصحيحة على مفردات الاختبـــار عن طريق التخمين العشوائي في كل من المحك المتفائل والمحك الصارم، يكون كالتالي:

1. احتمال (P) الإجابة الصحيحة على 5 من7 مفردات عن طريق التخمين العشوائي = 0.01287 عند مستوى دلالة إحصائية 0.01.[3]

2. احتمال (P) الإجابة الصحيحة على 6 من7 مفردات عن طريق التخمين العشوائي= 0.00134 عند مستوى دلالة إحصائية 0.001. أو بما يسمى المحك الصارم .[4]

مع الأخذ بعين الاعتبار أن احتمال حصول التلميذ على إجابة صحيحة من أربعة إجابات مقترحة هو 0.25، ويكون احتمال تخمين خمسة إجابات صحيحة –على الأقل– من سبعة مفردات يساوي 0.01287، أي 0.01287 x 577 ، مما يعني أن ما يقارب 7 تلامذة يمكنهم تحقيق أي مستوى بالتخمين الصرف. ويكون احتمال تخمين ستة إجابات صحيحة –على الأقل– من سبعة مفردات يساوي 0.00134، أي 0.00134 x 577، مما يعني أن ما يقارب تلميذ واحد فقط يمكنه تحقيق أي مستوى بالتخمين الصرف.

1. يحقق التلميذ المستوى الأول والثالث إذا حصل على 3 إجابات صحيحة من 5 كحد أدنى، مع الأخذ بعين الاعتبار أن احتمال الحصول على إجابة صحيحة من أربعة خيارات هو 0.25، ويكون احتمال تخمين ثلاثة إجابات صحيحة – على الأقل– من خمسة مفردات يساوي 0.103[5]، أي 0.103 x 566، مما يعني أن ما يقارب 58 تلميذاً وتلميذة، يمكنهم تحقيق أي مستوى بالتخمين الصرف.

---

[3]- C_7^5 $(0.25)^5(0.75)^2+C_7^6(0.25)^6(0.75)^1+C_7^7(0.25)^7(0.75)^0 = 0.01287$

[4]- $C_7^6(0.25)^6(0.75)^1+C_7^7(0.25)^7(0.75)^0 = 0.00134$

[5] - C_5^3 $(0.25)^3(0.75)^2+C_5^4(0.25)^4(0.75)^1+C_5^5(0.25)^5(0.75)^0 = 0.103$

2. يحقق التلميذ المستوى الثاني والرابع إذا حصل على 4 إجابات صحيحة من 6 كحد أدنى، مع الأخذ بعين الاعتبار أن احتمال الحصول على إجابة صحيحة من أربعة خيارات هو 0.25، ويكون احتمال تخمين أربعة إجابات صحيحة – على الأقل– من ستة مفردات يساوي 0.037[6]، أي 0.037 x 566، مما يعني أن ما يقارب 21 تلميذاً وتلميذة، يمكنهم تحقيق أي مستوى بالتخمين الصرف.

3. تحقيق المستوى الأول كحد أدنى كي يتم تصنيف التلميذ على مستويات (فان هيل)، وغير ذلك اعتبر التلميذ أنه غير مصنف.

4. تحقيق المستوى السابق للمستوى الذي يجري فحصه، فمثلاً، إذا تحقق المستوى الثالث عند تلميذ ما ولم يحقق المستوى الثاني لا يؤخذ بهذه النتيجة.

5- صياغة تعليمات الإجابـــة على اختبار مستويات التفكير الهندســـي:

تمت صياغة التعليمات الخاصة بالاختبار، وإعطاء فكرة تامة عن الهدف من الاختبار، ونوعية المفردات وكيفية الإجابة عليها بدقة، وتوضيح للتلامذة أنهم قد لا يستطيعون الإجابة على كل المفردات في هذا الاختبار ولكن عليهم بذل أكبر جهد ممكن للإجابة على كل مفردة.

6- إعداد مفتاح لتصحيح اختبار مستويات التفكير الهندســـي:

اعد الباحث مفتاح الإجابات الصحيحة لمفردات الاختبار، أنظر الملحق رقم (05)، وتعامل المفردة التي لها أكثر من إجابة واحدة معاملة المفردة غير الصحيحة، حيث أعطيت درجة واحدة 1 للإجابة الصحيحة و (صفر) للإجابة الخاطئة، وبذلك تصبح الدرجة الكليـــة لاختبار مستويات التفكير الهندسي تساوي 28 درجة، أي تتراوح بين 0 و28.

ثالثاً: مرحلـــة التجريـــب:

1- التجريب الاستطلاعـــي لمفردات اختبار مستويات التفكير الهندســـي:

تأتي هذه الخطوة قبيل طباعة الاختبار وإخراجه بصورته النهائية ليكون جاهزاً للتطبيق والاستخدام العملي والميداني، إذ ينصح بتطبيق مفردات الاختبـــار على عينة صغيرة من المفحوصين ما بين 20 و30

---

[6] - C_6^4 $(0.25)^4(0.75)^2+C_6^5(0.25)^5(0.75)^1+C_6^6(0.25)^6(0.75)^0 = 0.037$

شخصاً، وربما يكون الرقم المطلوب أكبر من ذلك ما بين 100 و200 إذا كانت النية تتجه إلى استخدام الاختبار على مدى أوسع مستقبلاً أو أغراض تجارية مثلاً. (النبهان، 2013، ص.81)

لهذا قام الباحث بتطبيق الاختبار على عينة استطلاعية أولى مكونة من 34 تلميذاً وتلميذة، وذلك لغرض التحقق من وضوح التعليمات وصياغة المفردات وملاءمة البدائل، وكذلك للتعرف على المفردات الغامضة التي تحتاج إلى إعـــادة صياغة أو تحتاج إلى تعديل، وأيضا للتعرف على البدائل المناسبة وتعديل أو إعادة صياغة البدائل غير المناسبة، وتحديد الزمن المناسب للإجابـــة عن مفردات الاختبار، وكذلك التعرف على أي مشكلات قد تواجه الباحث أثناء التطبيق لكي يمكن تداركها قبل التطبيق على العينـــة النهائية.

اتضح للباحث من خلال التجريب الاستطلاعي وضوح التعليمات عموماً، ورغم ذلك ظهرت بعض الإستفسارات حول طريقة الإجابـــة على الأسئلة الموضوعية (الاختيار من متعدد)، وهذا كان متوقع، حيث أن التلامذة في مدارسنا غير متعودين على هذا النوع من الاختبارات (الموضوعية)، كذلك طريقة طرح الأسئلة المنطقية في المستوى الرابع (الاستدلالي) وكيفية الإجابة عليها، والاستفسار حول بعض البدائل، كما حدد متوسط الزمن المستغرق الذي قدره الباحث من خلال استجابة التلامذة على الاختبار بـــ 60 دقيقة أي حصة دراسية كاملة، وتم حسابه عن طريق حساب المتوسط الحسابي للوقت الذي استغرقه التلامذة للإجابة عن مفردات الاختبار بعد تسجيل الوقت على ورقة إجابة كل تلميذ عند انتهائه من أداء الاختبار.

قيام الباحث بتطبيق الاختبـــار على عينة استطلاعية ثانية، بعد اعادة صياغة المفردات المبهمة وغير الواضحة، اختيرت بطريقة العينة المتوفرة، مكونة من 104 تلميذاً وتلميذة من تلامذة المرحلة الثانوية (الشعب العلمية)، في النصف الأول من السنة الدراسية 2016-2017، كعينة للتحقق من جودة المفردات الاختبارية، من خلال مؤشرات التمييز والصعوبة لمفردات الاختبار، وفعالية المشتتات (المموهات)، وكذلك التحقق من فعالية الاختبار ككل، من خلال مؤشرات صدقه وثباته.

2- التحليل الإحصائي لمفردات اختبار مستويات التفكير الهندســـــــي:

تختلف إجراءات تحليل مفردات الاختبارات المحكية المرجع عن الاختبارات معيارية المرجع، وذلك لأن الاختبارات محكية المرجع تهدف إلى معرفة مدى تحقيق المتعلم لمستويات الإتقان المحددة، وبالتالي فهي تحدد أداء التلميذ أو موقعه في ضوء محكات محددة (مستوى الإتقان المطلوب)، أو في ضوء أهداف سلوكية تصف الأداء المتوقع من التلميذ بعد الإنتهاء من وحدة تدريسية معينة. وبما أن بناء مثل هذه النوع

من الاختبارات يتطلب ربط المفردة في الاختبار بهدف محدد تسعى إلى قياسه، فإن تحليل المفردات في هذه الحالة يكون موجهاً لتحديد فاعلية كل مفردة، وحساسيتها لعملية التدريس، ويعتمد تحليل مفردات هذا النوع من الاختبارات على المؤشرات الثلاثة التالية: (أ) صعوبة المفردة؛ (ب) التمييز أو حساسية المفردة؛ (ج) مؤشر التوافق (النبهان، 2013، ص.247).

وقبل البدء في إجراءات التحليل الإحصائي لمفردات الاختبار، قام الباحث بمراجعة الإحصاءات الوصفية للتأكد من أنه لا توجد أية قيم غريبة أو شاذة، ومنها التأكد من أن جميع متوسطات المفردات تقع بين صفر وواحد، كما تأكد بأنه لا توجد قيم شاذة في تباين المفردات، وأن جميع قيم تباين المفردات تتراوح بين صفر و 0.25، وكذلك أن جميع الارتباطـــات بين مفردات الاختبار موجبة، وهو ما تحقق في البيانات الحالية، حيث أن جميع متوسطات درجات مفردات اختبار مستويات التفكير الهندسي وفقا للعينة الاستطلاعية، وقعت بين 0.14 و0.85، كما أن جميع قيم تباين مفردات الاختبـــار تراوحت بين 0.13 و0.25، وبأن جميع الارتباطات كانت موجبة، ووفقاً لنتائج الإحصاءات الوصفية تبدو أن جميع البيانـــات التي ادخلت للتحليل الإحصائـــي بيانات سليمة يمكن الإطمئنان عليها، و بالتالـــي الاستمرار في الحصول على مؤشرات أخرى لمفردات الاختبار.

2-1- مؤشرات الصعوبـــة لمفردات الاختبار:

لا يتحدد مستوى صعوبة المفردة في الاختبار الإتقان محكي المرجع، وفقاً لقدرة المفردة على التمييز بين الأفراد ذوي التحصيل المرتفع في المجموعة العليا وذوي التحصيل المنخفض في المجموعة الدنيا، وإنما يتحدد وفقا لنتاج التعلم أو المهمة التعليمية التي صممت المفردة لقياسها، وهذا يعني أن كون المفردة سهلة جداً أو صعبة جداً يعتمد بالدرجة الأولى على طبيعة المهمة التعليمية التي تقيسها، فمعامل صعوبـــة مفردات الإختبار المحكـــي المرجع مرتبط بأمرين، وهما:

1. المستوى التعليمي الذي يوجد عليه التلامذة، بمعنى أن أساليب التعليم غير فعالة، في حالة إنخفاض مؤشر صعوبة المفردة.
2. المحك أو نقطة القطع التي يجب اجتيازها من قبل التلامذة.

ولحساب مؤشر صعوبة مفردات اختبار مستويات التفكير الهندسي، والذي هو عبارة عن النسبة المئوية للممتحنيين الذين أجابوا على المفردة إجابة صحيحة. قام الباحث بحساب معامل أو مؤشر الصعوبة لمفردات الاختبار باستخدام المعادلة رقم (03) السابقة الذكر: (النبهان، 2013، ص.247)

جدول (14)

يوضح قيم مؤشر صعوبـــة مفردات الاختبار وفقا للعينة الاستطلاعية ن=104.

الإدراكي أو البصري		التحليلي أو الوصفي		شبه الاستدلالي		الاستدلالي	
رقم المفردة	مؤشر الصعوبة	رقم المفردة	مؤشر الصعوبة	رقم المفردة	مؤشر الصعوبة	رقم المفردة	مؤشر الصعوبة
1	0.51	8	0.37	15	0.31	22	0.41
2	0.48	9	0.44	16	0.36	23	0.44
3	0.77	10	0.26	17	0.43	24	0.60
4	0.44	11	0.78	18	0.68	25	0.41
5	0.73	12	0.85	19	0.73	26	0.39
6	0.63	13	0.55	20	0.14	27	0.21
7	0.82	14	0.45	21	0.51	28	0.31

يتضح من الجدول رقم (14)، أن قيم مؤشرات صعوبة المفردات تراوحت بين 0.14 و0.85 بمتوسط حسابي 0.50، فحسب كيلي (1939) Kelley مؤشر الصعوبة الأقل من 0.20 تعتبر المفردة صعبة جداً و يجب أن تراجع، حيث اشارت النتائج إلى أن هناك 19 مفردة جاءت قيمها مناسبة، والتي تراوحت مؤشرات صعوبتها ما بين 0.3 و0.7، أي ما نسبته 68% من مفردات الاختبار الكلي، بينما كانت 7 مفردات سهلة، والتي تمثل 21% من مفردات الاختبار الكلي، في حين أن ثلاثة مفردات فقط والتي تمثل 11% من مفردات الاختبار كانت صعبـــة.

و يرى أبوناهية (1994) أن مفردات الاختبار المحكـــي المرجع، يجب أن تكون متدرجة في صعوبتهاً، بحيث تبدأ بالمفردات السهلة وتنتهي بالمفردات الصعبة، وبالتالي تتراوح قيمة مؤشرات صعوبتها ما بين 10% إلى 90%، بشرط أن يكون معدل الصعوبـــة للمفردات ككل في الاختبار في حدود 50%، وهو ما تحقق في بحثنا الحالـــي، بتسجيل لقيمة مؤشر متوسط صعوبـــة المفردات يساوي 50%، وهو دليل على تمتع مفردات اختبار مستويات التفكير الهندســـي بمؤشرات صعوبة مناسبة.

كما تم حساب متوسط صعوبـــة المفردات لكل مستوى من مستويات التفكير الهندســـي لـــ(فان هيل)، حيث سجلنا القيمة 0.67 في المستوى الأول، والقيمة 0.55 في المستوى الثاني، والقيمة 0.48 في المستوى

الثالث، وأخيراً القيمة 0.42 في المستوى الرابع، والتي جاءت متدرجة في الصعوبة بحسب كل مستوى من الأدنى إلى الأعلى، وتضيف هذه القيم دليل آخر على هرمية نموذج (فان هيل) لتطور التفكير الهندســـي.

2-2- مؤشر التمييـــز أو حساسية مفردات الاختبار:

يتم حساب مؤشر حساسية المفردات لعملية التعلم أو مؤشر التمييز لمفردات الاختبارات محكية المرجع، من خلال حساب الفرق بين مستويـــات صعوبة المفردة لكل من المفحوصين الذين حققوا مستوى الإتقان والمفحوصين الذين لم يحققوا مستوى الإتقان، والطريقة التي استخدمها الباحث في تحليل تمييز مفردات اختبار التفكير الهندســـي هي طريقة مؤشر التمييز (ب)، والتي اقترحها برينان (1972) في حساب مؤشرات حساسية مفردات الاختبارات محكية المرجع، والتي تعطى بالمعادلات رقم (05) و رقم(05.1) و رقم(05.2) السابقة الذكر:

ويتراوح مؤشر الحساسية لبرينـــان ما بين-1 و+1 وتشير القيمة الموجبة لهذا المعامل إلى أن المفردة تميز بين المفحوصين الذين حققوا مستوى الإتقان المطلوب أو المحـــك، وأولئك الذين لم يحققوا مستوى الإتقان أو المحك. أما القيمة السالبة فتشير إلى وجود مشكلة أو خلل حول المفردة.(Brennan, 1972, p.292) (أبوناهية، 1994، ص.325) (النبهان، 2013، ص.250)، والجدول رقم (15) يوضح مؤشرات تمييز مفردات اختبار مستويات التفكير الهندســـي.

جدول (15)

يوضح قيم مؤشرات الحساسية أو التمييز لمفردات الاختبار وفقا للعينة الاستطلاعية ن=104.

الإدراكي أو البصري		التحليلي أو الوصفي		شبه الاستدلالي		الاستدلالي	
رقم المفردة	مؤشر التمييز	رقم المفردة	مؤشر التمييز	رقم المفردة	مؤشر التمييز	رقم المفردة	مؤشر التمييز
1	0.26	8	0.59	15	0.76	22	0.70
2	0.46	9	0.37	16	0.30	23	0.40
3	0.37	10	0.42	17	0.48	24	0.21
4	0.67	11	0.30	18	0.38	25	0.60
5	0.31	12	0.15	19	0.33	26	0.54
6	0.29	13	0.47	20	0.38	27	0.74
7	0.23	14	0.33	21	0.41	28	0.27

يتضح من البيانات الخاصة بمؤشرات التمييز لمفردات اختبار مستويات التفكير الهندســـي في الجدول رقم (15)، أن جميع مؤشرات تمييز المفردات لها قيم موجبة ولا توجد بينهما مفردات سالبة القيمة، أي أنها

تميز بين المفحوصين الذين حققوا مستوى الإتقان المطلوب أو المحـــك، وأولئك الذين لم يحققوا مستوى الإتقان أو المحك. حيث سجلنا أعلى قيمة لهذا المؤشر 0.76 والتي كانت من نصيب المفردة 15، وأقل قيمة 0.15 كانت من نصيب المفردة 12، أي أن قيم مؤشرات التمييز لمفردات اختبار مستويات التفكير الهندســـي ككل تراوحت ما بين القيمة 0.15 و0.76 بمتوسط حسابي قدره 0.42، حيث أشارت 22 مفردة تمثل ما نسبته 78.57% من مفردات الاختبـــار، بأن مؤشرات تمييزها جيدة وممتازة وهي أكبر من القيمة 0.30، بينما هناك 5 مفردات فقط تمثل 17.85% من مفردات الاختبار لها معاملات تمييز هامشية ما بين القيمة 0.20 و 0.29، ومفردة واحدة فقط أقل من 0.20، ما يمثل نسبة 3.57% من مفردات الاختبار، فحسب كيلي (1939) Kelley فإن قيم مؤشرات التمييز التي تتراوح ما بين 0.10 و 0.30 قيم مقبولة، والقيم ما فوق 0.30 فهي قيم جيدة (Kelley, 1939 , pp.17-24)، مع مراعاة أن مفردات الاختبارات مرجعية المحك تتميز بإنخفاض معامل تمييزها ذلك لأن الهدف من هذه الاختبارات ليس قياس الفروق الفردية بين الأفراد، بل قياس وتحديد ما إذا كان المتعلم قد وصل إلى مستوى الإتقان (Mastery) في المعرفة أو المهارة المحددة، لهذا فإن مؤشرات التمييز الموجبة لمفردات الاختبارات المحكية المرجع لا تحذف، في حين أن المفردات ذات القيم السالبة والتي رغم تحسينها تبقى سالبة فهذه المفردات تحذف من الاختبـــار حسب (أبو ناهية، 1994، ص.326). وهذا ما يدل على تمتع معظم مفردات اختبار مستويات التفكير الهندســـي بمؤشرات تمييز موجبة ومناسبة لمجتمع البحث.

2-3- تحليل فعاليـــة البدائل (مموهات) الاختبار:

يرى كل من أمين وأبو علام (2012) أن فحص بدائل مفردات الاختبـــار من متعدد نلجأ إليه إذا كانت قيمة مؤشر تمييز المفردة سالباً أو قريباً من الصفر، أو إذا كانت إحدى البدائل لم يتخيرها أحد من المختبرين سواء المجموعة العليا أو الدنيا (المتقنين وغير المتقنين)، فربما يكون دليلاً على عدم فاعليـــة هذا المشتت وبالتالي يحذف، لأنه يضيع وقت المشارك في قراءته، أو ربما نلاحظ أن أحد البدائـــل اختاره معظم أفراد المجموعة العليا ولم يختاروا الإجابة التي حددها معد الاختبـــار، فهذا يشير إلى أن المشتت ربما يكون هو الإجابـــة الصحيحة بالفعل أو أكثر جاذبية من الإجابـــة الصحيحة التي حددها معد الاختبار.(ص.321)

ولكي تكون مفـــردات الاختيار من متعدد على درجة عاليـــة من الصلاحية والكفاءة، يجب أن تتسم المموهات أو المشتتات بفعالية عالية، ويهدف تحليل البدائل إلى مراجعة هذه المموهات وتعديلها إن كانت

غير فعالـــة في صرف إنتباه الأفراد عن البديـــل الصحيح، وذلك للتقليل من نسبة التخمين، ويتم ذلك من خلال معرفة عدد الأفراد الذين اختاروا البدائل الخطأ من بين الأفراد ذوي الدرجات العليا (المتقنين)، والأفراد ذوي الدرجات الدنيا (غير المتقنين). والجدول رقم (16)، يوضح نتائج تحليل بدائل مفردات اختبار مستويات التفكير الهندســـي وفقا للعينة الاستطلاعية 104 تلميذاً وتلميذة. بعد اختيار نسبة 27% من طرفي المجموعة، واستبعاد نسبة 46% من المتحصلين على الدرجـــات الوسطـــى، حيث بلغ عدد الأفراد في كلا المجموعتين الطرفيتين 56 تلميذاً وتلميذة .

جدول (16)

يوضح نتائج تحليل بدائل مفردات اختبار مستويات التفكير الهندســـي 28 مفردة وفقا للعينة الإستطلاعية ن= 104.

رقم المفردة	المجموعة	البدائل				رقم المفردة	المجموعة	البدائل			
		D	C	B	A			D	C	B	A
1	العليا	4	0	16	7	15	العليا	9	11	3	5
	الدنيا	4	0	12	11		الدنيا	10	7	3	8
2	العليا	1	0	14	13	16	العليا	12	6	5	5
	الدنيا	0	2	11	15		الدنيا	10	7	5	5
3	العليا	26	0	1	1	17	العليا	3	14	4	6
	الدنيا	22	2	4	0		الدنيا	8	11	4	5
4	العليا	13	4	11	0	18	العليا	17	5	5	0
	الدنيا	12	3	0	13		الدنيا	14	5	6	2
5	العليا	0	2	0	26	19	العليا	2	22	2	2
	الدنيا	0	2	4	22		الدنيا	3	15	5	5
6	العليا	19	3	3	3	20	العليا	21	0	2	5
	الدنيا	10	9	1	7		الدنيا	20	1	2	5
7	العليا	0	2	0	26	21	العليا	3	17	4	4
	الدنيا	1	7	0	20		الدنيا	3	13	9	3
8	العليا	3	2	7	10	22	العليا	7	12	4	4
	الدنيا	6	3	11	6		الدنيا	9	10	4	4
9	العليا	7	0	17	4	23	العليا	3	1	17	6
	الدنيا	7	4	14	3		الدنيا	6	4	13	5
10	العليا	8	11	8	1	24	العليا	2	9	1	16
	الدنيا	3	5	13	5		الدنيا	3	3	1	21
11	العليا	1	1	2	24	25	العليا	3	5	16	3
	الدنيا	3	3	2	20		الدنيا	3	11	13	1

12	العليا	1	25	1	0	26	العليا	7	4	4	13
	الدنيا	3	20	1	4		الدنيا	15	2	6	5
13	العليا	5	7	16	0	27	العليا	10	4	5	9
	الدنيا	5	5	17	1		الدنيا	5	4	6	12
14	العليا	13	1	9	5	28	العليا	4	9	4	11
	الدنيا	10	3	10	5		الدنيا	2	8	8	10

* البديل المضلل هو البديل الصحيح.

يتضح من الجدول رقم (16)، أن المفردات 8،11،15،16،17،19،21،22،23،24،25،26،27، 6 على درجة عالية من الكفاءة في التمييز بين المجموعتين العليا والدنيا في اختيار البديل الصحيح، كما أن جميع البدائل الأخرى (المشتتات) تم اختيارها من قبل بعض الأفراد في المجموعة العليا والدنيا. كما أن المفردات 2،3،4،5،7،9،12،18 فقد ميزت بين الأفراد في المجموعتين العليا والدنيا في اختيارهم البديل الصحيح، إلا أنها تحتوى على بديل لم يتم اختياره مطلقا ضمن أفراد المجموعتين وعلى هذا فيجب إستبداله بمشتت آخر. أما المفردات 10،13،14،20،28 فكانت مفــردات ضعيفة إذ أنها لم تميز جيداً بين الأفراد في المجموعتين، كما أن عدد الأفراد في المجموعة العليا الذين اختاروا البديل غير الصحيح كان أكبر من الذين اختاروا البديل الصحيح، وربما بمراجعة هذه البدائل من جديد، أو تطويرها على نحو آخر ليصبح أدائها أفضل. و تتمثل أهمية تحليــل المفــردات أيضاً في التغلب على مشكلة من أكبر المشكلات التي تواجه الاختبــارات الموضوعية، وهي مشكلة التخمين، والتي اهتم خبراء القيــاس النفســي والتربوي بدراستها للتعرف على مدى تأثيرها سواء في أداء المختبرين، أو الخصائص السيكومترية لهذه الاختبارات.

2-4- تقويم الصدق التقاربــي و التباعــدي (التمييزي) لمفردات الاختبار:

لتقويم الصدق التقاربي والتمييزي لمفردات كل مستوى من مستويات اختبار التفكير الهندسي الأربعة، و ذلك من خلال حساب معامل الارتباط (بيرسون)، بين درجة كل مفردة في المستوى بالدرجة الكلية للمستوى الذي تنتمي إليه، وكذلك بين درجة كل مفردة في كل مستوى والدرجة الكلية للمستويات الأربعة الأخرى، بعد حذف درجة المفردة من البعد أو المستوى. والجدول رقم (17) يوضح نتائج قيم معاملات الارتباط المتحصل عليها.

جدول (17)

يوضح قيم معاملات الارتباط (بيرسون) بين درجة كل مفردة والدرجة الكلية للمستوى الذي تنتمي إليه بعد حذف درجة المفردة من البعد، والمستويات الأخرى ن=104.

رقم		قيم معاملات الارتباط بين مستويات التفكير الهندســي لفان هيل			
المفردة	المستوى	الإدراكي أو البصري	الوصفي أو التحليلي	شبه الاستدلالي	الاستدلالي
01	الإدراكي أو البصري	**0.382	*0.211	**0.274	0.178
02		**0.503	0.017	0.133	0.054–
03		**0.519	*0.227	0.040	0.067
04		**0.386	0.142	0.123	0.005
05		**0.422	0.154	*0.239	0.115
06		**0.401	0.117	0.002–	0.164
07		**0.350	0.117	0.094	0.154
08	الوصفي أو التحليلي	*0.222	**0.430	0.114	0.030
09		0.037	**0.292	*0.225	0.146
10		0.050	**0.369	*0.215	*0.225
11		0.121	**0.544	*0.215	0.038
12		0.122	**0.273	0.150	0.156
13		**0.260	**0.494	0.164	**0.307
14		0.120	**0.440	0.001–	0.091
15	شبه الاستدلالي	*0.216	**0.346	**0.664	**0.273
16		0.057	0.053	**0.431	*0.233
17		**0.428	*0.237	**0.449	0.176
18		0.010	*0.249	**0.390	0.042
19		0.032	*0.208	**0.467	0.123
20		0.132	**0.325	**0.424	**0.259
21		0.121	0.047–	**0.461	0.179–
22	الاستدلالي	0.067	*0.204	0.114	**0.658
23		*0.197	*0.200	*0.224	**0.436
24		0.105	0.031	0.078	**0.357
25		0.071	*0.225	0.135	**0.468
26		0.108	0.088	0.097	**0.418
27		0.071	0.170	**0.288	**0.575
28		0.014	0.194	0.041–	**0.299

** دال إحصائيا عند مستوى دلالة (0.01). * دال إحصائيا عند مستوى دلالة (0.05).

يضح من الجدول رقم (17)، بأن بعض المفردات لها معاملات ارتباط دالـــة إحصائياً في المستوى الذي تنتمي إليه المفردة والمستويات الأخرى، إلا أنها لم تكن بحجم قيم معاملات الارتباط للعلاقة بين درجة كل مفردة والمستوى الذي تنتمي إليه المفردة، وبناءً على هذه النتائج تم الإحتفاظ بمفردات الاختبار المقترحة لكل مستوى، وهذا دليل قوى على تحقق الصدق التقاربـــي، والصدق التباعدي (التمييزي) لمفردات اختبار مستويات التفكير الهندسي، وبالتالي فإن كل مفردة تقيس الأهداف المتوقعة منها في كل مستوى من مستويات الاختبـــار الأربعة.

3- تقييم فاعلية اختبار مستويات التفكير الهندســـــــي:

يمكن تناول التحقق من فاعلية الاختبـــار من جانبين، هما: الجانب الأول يتعلق بمفردات الاختبار وبدرجة فاعلية كل مفردة من هذه المفردات ومدى توفر خصائص الفاعلية لهذه المفردات، وهي خصائص تتعلق بمستوى صعوبتها وقدرتها على التمييز بين المفحوصين، وكذلك فعالية بدائلها المختلفة، أي أنه يتعلق بفاعلية مكونات أو أجزاء أو حدات منفصلة في داخل الاختبار، وهي الإجراءات التي قام بها الباحث سابقاً، أما الجانب الثاني فيتعلق بفاعلية الاختبار ككل، بجميع مفرداته أو مكوناته أو أجزائه، ومدى توفر خصائص فاعلية هذا الاختبار بمجمل مفرداته، وهي الخصائص التي تتعلق بالصدق، والثبات، والقابلية للإستعمال. ويعتمد تحديد صدق وثبـــات الاختبار المعيـــاري المرجع على الإرتباط والتباين في الدرجات، وذلك خلاف طريقة تحديدها في الاختبار المحكي المرجع، الذي يحدد صدقه مدى ارتباطه بالهدف المراد تحقيقه، ويقاس ثباتـــه عن طريق إتساق قرارات التصنيف.

3-1- تقويم صـــدق درجـــات الاختبار:

يعتبر الصدق الخاصية الأساسية الأولى التي يجب أن تتوفر في وسيلة القياس بصفة عامة، والاختبار بصفة خاصة، بل أن حساب صدق الاختبار من أفضل الأساليب التي تستخدم لتقييم فاعلية الاختبار بوصفه أداة قياس، والصدق بتعبير بسيط هو أن يقيس الاختبار ما وضع لقياسه، أي يقيس فعلا الوظيفة التي يفترض أنه يقيسها.

3-1-1- الصـــدق بدلالة المحتوى أوالصدق الوصفـــي:

يتم حساب صدق الاختبارات محكية المرجع باستخدام طريقة صدق بدلالة المحتوى أو صدق المحتوى، والطريقة المستخدمة في حساب صدق المحتوى بالنسبة للاختبارات محكيـــة المرجع هي نفس الطريقة

المستخدمة مع الاختبارات معيارية المرجع، إلا أن التحقق من صدق المحتوى في الاختبارات محكية المرجع أسهل لحد ما منه في الاختبارات معيارية المرجع، والسبب في ذلك هو وجود مجـــال ذو مواصفات محددة مما يسهل تحديد مدى مطابقـــة مفردات الاختبار لذلك المجال. (أبوناهية، 1994، ص.350)

تم تقدير الصدق الوصفي لاختبار مستويات التفكير الهندســـي والمبنى وفق نظرية (فان هيل) للتفكير الهندســـي، بعد عرضه ومعه تعريف مفصل بالنظرية، و مستوياتها في تطور التفكير الهندســـي، على مجموعة من المحكمين المتخصصين و الخبراء في مجال تدريس مادة الرياضيات، بحيث تم الحكم على كل مفردة بمدى مطابقتها مع محتوى وأهداف كل مستوى من مستويات (فان هيل) الأربعة، والتأكد مما إذا كانت مفردات الاختبـــار تمثل فعلا محتوى المجال (المستوى) و أهدافه، وأن هذا التمثيل هو تمثيلا صادقا.[7] حيث يمثل صدق المحتوى أهمية خاصة في التحقق من الاختبـــارات التكوينية التي تهتم بالإتقان في واحدة أو أكثر من الأهداف التعليمية النوعية أو الخاصة، وكذلك الاختبـــارات التحصيلية محكيـــة المرجع.

وسبق وأن اشرنا إلى القيمة المحسوبة لمؤشر صدق المحتوى المحقق للاختبـــار، وهي القيمة 0.74، والتي تعتبر قيمة جيدة ومقبولة، ودالـــة إحصائياً عند مستوى دلالة 0.05 حسب (لاوش).

3-1-2- صدق إتقـــاء النطاق السلوكـــي (التكوين الفرضي):

يعتبر أحد التطبيقات لمحك الاتساق الداخلي ارتباط درجات الاختبارات الفرعية (المستويات) بالدرجة الكلية، فارتباطات الاتساق الداخلـــي سواء اعتمدت على المفردات أو الاختبارات الفرعية هي بدرجة أساسية مقاييس التجانس. ونظرا لأن درجة تجانس اختبار معين تفيد في وصف النطاق السلوكـــي أو السمة التي يعاينها الاختبار، فإن هذه الدرجة توائم إلى حد ما صدق التكوين الفرضـــي للاختبـــار. (أناستازي، 2015، ص.169)

---

[7] – اعتبر (Usiskin) مطور اختبار (فان هيل) الأصلي بمشروع جامعة شيكاغو(1982)، خلال تواصل الباحث معه عبر البريد الالكتروني أن الاختبار صادقاً طالماً أن أسئلته تتناسب مع تفكير التلامذة والرياضيات التي تعلموها، والتي تم الحكم عليها في هذا البحث من خلال خبراء المادة (الرياضيات).

لهذا تأكد الباحث من صدق التكوين الفرضي من خلال الاتساق الداخلـــي لمفردات اختبار مستويات التفكير الهندســـي وفقا للعينة الاستطلاعية، باستخدام معامل الارتباط (بيرسون) بين درجات الاختبارات الفرعية أو المستويات الأربعة بالدرجة الكلية للاختبار، ونتائجه موضحة في الجدول رقم (18):

الجدول (18)

يوضح قيم معاملات الارتباط (بيرسون) بين درجات مستويات الاختبار الأربعة (الفرعية) ودرجة الاختبار الكلي وفقا للعينة الاستطلاعية ن=104.

المستوى	التحليلي- الوصفي	الاستنباط غير الشكلي	الاستنباط الشكلي	الاختبار الكلي
الإدراكي- البصري	0.385**	0.348**	0.248*	0.677**
التحليلي- الوصفي		0.376**	0.414**	0.756**
الاستنباط غير الشكلي			0.351**	0.751**
الاستنباط الشكلي				0.730**

** دال احصائياً عند مستوى دلالة (0.01). * دال إحصائياً عند مستوى دلالة (0.05).

يتبين من الجدول رقم (18)، أن قيم معاملات الارتباط بين الدرجة الكلية للاختبار والدرجة الكلية لكل مستوى من مستويات الاختبار الأربعة، معاملات ارتباطات قوية تراوحت ما بين 0.68 و0.76 وجميعها قيم مرتفعة و دالة إحصائياً عند مستوى معنوية α=0.01، كما يتضح أن هناك معاملات ارتباط بقيم متفاوتة ولكنها جميعها قيم دالـــة إحصائياً عند مستوى معنوية α=0.01 و α=0.05 بين الاختبارات الفرعية (المستويات) الأربعة المكونة للاختبار في ما بين بعضها البعض، ويعتبر هذا مؤشـــراً دالاً وقوى على الاتساق الداخلي لمفردات اختبار مستويات التفكير الهندسي ككل.

3-2- **تقويم ثبـــات درجات الاختبار:**

تعتمد معظم طرق تقدير ثبات الاختبارات المعيارية المرجع، على معامل الارتباط بين متغيرين، وتزداد قيم هذه المعاملات بزيادة تباين الدرجات، بينما في الاختبارات المحكية المرجع يتم تفسير الدرجة بالرجوع إلى مستوى أداء محدد فيتوقع أن تكون مجموعة التلامذة متجانسة، وخاصة عندما يوجه التعليم إلى مستوى معين من الإتقان للمعلومات والمهارات، وبالتالـــي حساب الثبات بالطرق التقليدية للاختبارات محكية المرجع فإنه من المتوقع أن تكون القيم منخفضة، وبالتالي فإن طرق تقدير الثبات الخاصة بالاختبارات المعيارية المرجع لا يصلح لقياس ثبات الاختبارات محكيـــة المرجع. لهذا اختار الباحث أساليب تتناسب مع هذا النوع من الاختبارات بالإضافة إلى الأساليب التي اعتمدها يوسيسكين في دراسته.

3-2-1- حساب مؤشر ليفينجستون و معامل كودر- ريتشاردسون (KR_{20}):

يعتمد مؤشر ليفينجستون على مفاهيم النظرية الكلاسيكية في تقدير ثبـــات الاختبارات معيارية المرجع، والتي تركز على معرفة انحراف درجة الفرد عن متوسط العينة، إلا أنه يستخدم الدرجة الفاصلـــة بدلاً من المتوسط، فبدلاً من إيجاد انحراف درجـــات الأفراد عن المتوسط تحسب انحرافات درجات الأفراد عن الدرجة الفاصلة المحددة. ويتوقع أن مؤشر ليفنجستون يطابق معامل كيودر- ريتشاردسون (KR_{20}) في حالة أن تكون درجة القطع مطابقة لمتوسط درجات الأفراد على الاختبار. (النبهان، 2013، ص.314) (أبو علام، 1995، ص.264)

جدول (19)

يوضح قيم معاملات ثبات الاختبار باستخدام كيودر-ريتشاردسون (KR_{20})، ومعامل سبيرمان-وبراون لتصحيح الطول، و مؤشر ليفنجستون (Livingston,1972) وفقا للعينة الإستطلاعية ن=104.

مستويات(فان هيل) للتفكير الهندسي	معامل كودر- ريتشاردسون KR_{20}	معامل سبيرمان - وبراون لتصحيح الطول	مؤشر ليفنجستون (Livingston)	
			محك 7/5	محك7/6
الإدراكي- البصري	0.24	0.65	0.36	0.68
التحليلي- الوصفي	0.23	0.72	0.60	0.80
الاستنباط غير الشكلي	0.46	0.77	0.78	0.88
الاستنباط الشكلي	0.48	0.60	0.83	0.90
الاختبار الكلـــي	0.71			

يتضح من الجدول رقم (19)، أن قيم الثبـــات باستخدام معامل كودر- ريتشاردسون (KR_{20}) للاختبار الكلـــي قد بلغت القيمة 0.71، وبالتالي فهي قيمة مقبولة للثبات، إلا أن قيم الثبات حسب كل مستوى من المستويات الأربعة كان منخفضاً، وهذا ما سجلته البحوث السابقـــة المشابهة لبحثنا، وأهمها دراسة يوسيسكين (1982)، ويمكن أن يرد ضعف الثبـــات لكل من مستويـــات الاختبار الأربعة، باستخدام مؤشر كودر- ريتشاردسون (KR_{20})، والذي يعتمد على معامل الارتباط، حيث تزداد قيمه بزيادة تباين الدرجات، في حين أن الاختبارات المحكية المرجع- كالاختبار المصمم في بحثنا الحالي- يتم تفسير الدرجة فيها بالرجوع إلى مستوى أداء محدد، فيتوقع أن تكون مجموعة التلامذة متجانسة، كذلك قلة المفردات 7 مفردات مكونة لكل مستوى، ويمكن أن يرد كذلك إلى عامل صعوبة المفردات، فالمفردات السهلة جداً والصعبة جداً، تكون فيها

درجات التلامذة متقاربة، فكما نعلم أن تباين درجات الاختبـــار يرتبط مع ثبات الاختبار بعلاقة طردية، أي أنه كلما زاد تباين الدرجات زاد ثبـــات الاختبار والعكس صحيح.

وبتطبيق نفس الأسلوب الذي طبقه يوسيسكين في دراسته السابقة عند تصميم اختبار فان هيل لمستويات التفكير الهندســـي، وهو استخدام معادلة سبيرمان- براون لتصحيح الطول في التنبؤ بثبات الاختبار في حالة زيادة أو تخفيض عدد المفردات. بلغت معاملات ثبـــات الاختبارات الفرعية أو (المستويات)، إذا كان كل مستوى يحتوي على 28 مفردة، كالتالي: 0.54؛ 0.54؛ 0.77؛ 0.78 من المستوى الأول إلى المستوى الرابع على التوالي.

و تشير نتائج الجدول رقم (19)، كذلك إلى أن قيم مؤشر الثبات لـــ ليفنجستون في المستوى الأول وعند درجة قطع 7/5 أي ما نسبته 71% بأنها قيمة منخفضة 0.36، في حين تشير إلى القيمة 0.68 عند المحك الصارم 7/6 أي ما نسبته 86%، وبلغت قيم الثبات بهذا المعامل للمستوى الثاني 0.60 عند المحك 7/5، والقيمة 0.80 عند المحك 7/6، كما سجلنا القيم التالية 0.78 و0.83 للمستوى الثالث والمستوى الرابع على التوالي عند المحك 7/5، والقيم 0.88 و0.90 عند المحك 7/6 وهي قيم مرتفعة، وتشير جميعها إلى أن مجموعة المفردات التي تشتمل عليها هذه المستويات تتميز بدرجة مقبولة من الثبات.

فاشتقاق معامل ليفنجستون، حسب محمود علام (2005) كما ورد عن بوسالم (2014، ص.184) لا يعتمد على تقدير مدى اتساق تصنيف الأفراد بالنسبة لدرجة قطع معينة، إذا أعيد تطبيق نفس الاختبـــار أو صورة موازية له على نفس مجموعة المتعلمين، بل يعتمد على مدى اتساق التصنيف لمدى تمكن الأفراد من نطاقات سلوكية أو عدم تمكنهم، والذي يعد الغرض الرئيســـي من استخدام هذا النوع من الاختبارات.

ملخص الفصل:

قام الباحث بإتباع خطوات إجرائية في بناء اختبار تشخيصي محكي المرجع يقيس التفكير الهندســـي، اعتماداً على نظرية (فان هيل) لمستويات التفكير الهندســـي. مروراً بثلاثة مراحل أساسية، وهي مرحلة التحليل؛ ثم مرحلة البناء؛ ثم مرحلة التجريب، لينتقل بعدها الباحث إلى التحليل الإحصائي لمفردات الاختبار من خلال التحقق من مؤشرات الصعوبـــة، والتمييز، وفعاليـــة البدائل، وتقويم الصدق التقاربـــي، و الصدق التباعدي، ثم إلى تقويم فعالية الاختبار ككل، من خلال مؤشرات صدقه، وثباته درجاته. وفي الفصل الموالي سوف يتم عرض وتحليل وتفسير ومناقشة النتائج التي توصلنا إليها في بحثنا الحالي.

الفصل الثامن

عرض وتحليل نتائـج البحث

تمهيد:

بعد عرضنا لفصل إجراءات بناء اختبار مستويات التفكير الهندسي وفقا لنموذج (فان هيل)، وتجريبه على العينة الاستطلاعية، والتي أشارت فيه نتائج التحليل الأولى لمفردات الاختبار وفعالية الاختبار ككل، إلى خصائص سيكومترية مقبولة، مما يؤهله إلى التطبيق النهائي على العينة الأساسية للبحث.

وقبل عرض النتائج للتحقق من فرضيات البحث، قام الباحث ببحث الفروق بين متوسطي مجموعتي الذكور والإناث في أدائهم على اختبار مستويات التفكير الهندسي قبل التعامل مع العينة كمجموعة واحدة. من خلال استخدم اختبار (ت) للفروق بين متوسطي عينتين مستقلتين (الذكور، والإناث)، والنتائج يوضحها الجدول رقم (20):

الجدول (20)

يوضح نتائج اختبار (ت) للفروق بين المتوسطات المستقلة تبعاً لمتغير الجنس.

المجموعات	العدد	المتوسط الحسابي	الإنحراف المعياري	قيمة (ت)	القيمة الاحتمالية	الدلالة الإحصائية
الذكور	247	14.17	3.983	1.609	0.108	غير دالة احصائياً
الإناث	330	14.74	4.278			

يتضح من الجدول رقم (20) أن قيمة اختبار (ت) للفروق بين متوسطي استجابات الذكور و الإناث على اختبار مستويات التفكير الهندسي بلغت 1.609 وهي قيمة غير دالة إحصائياً، أي أنه لا توجد فروق معنوية بين متوسطي الذكور و الإناث في الأداء على اختبار مستويات التفكير الهندسي.

ولبحث الفروق بين متوسطات الاستجابة على اختبار مستويات التفكير الهندسي في المستويات الدراسية الثلاثة (أولى، ثانية، ثالثة ثانوي) قبل التعامل مع العينة كمجموعة واحدة. استخدم الباحث اختبار تحليل التباين الأحادي الإتجاه للفروق بين المتوسطات تبعاً لمتغير المستوى الدراسي، والجدول رقم (21) يوضح هذه النتائج:

الجدول(21)

يوضح نتائج اختبار تحليل التباين الأحادي (ف) تبعاً لمتغير المستوى الدراسي.

التفكير الهندسي تبعاً لمتغير المستوى الدراسي	المتوسط الحسابي	الانحراف المعياري	قيمة (ف)	القيمة الاحتمالية	الدلالة الإحصائية
أولى ثانوي	14.08	4.151			غير دالة
ثانية ثانوي	14.48	4.111	1.305	0.272	إحصائياً
ثالثة ثانوي	14.76	4.194			

يتضح من الجدول رقم (21) أن قيمة اختبار تحليل التباين أحادي الإتجاه (ف) للفروق بين متوسطات المستويات الدراسية الثلاثة (أولى، وثانية، وثالثة ثانوي) في أدائهم على اختبار مستويات التفكير الهندسي، بلغت قيمته 1.305 وهي قيمة غير دالة إحصائياً، أي أنه لا توجد فروق معنوية بين متوسطات المستويات الدراسية الثلاثة (أولى، وثانية، وثالثة ثانوي).

وبناءً على نتائج اختبار (ت) للفروق بين المتوسطات المستقلة، ونتائج اختبار(ف) لتحليل التباين. فإن الباحث سيتعامل مع العينة كمجموعة واحدة في عملية التحليل الإحصائي لمفردات اختبار مستويات التفكير الهندسي سواء باستخدام أساليب النظرية الكلاسيكية في القياس أو باستخدام نموذج راش. وفي ما يلي عرض و تحليل لنتائج البحث وفقاً لفرضياته، كالتالي:

1- عرض وتحليل نتائج فرضيات البحث:

1-1- عرض وتحليل نتائج الفرضية الأولى:

والتي تنص على أن:" يتوفر اختبار مستويات التفكير الهندسي على خصائص سيكومترية مقبولة باستخدام النظرية الكلاسيكية في القياس".

من أجل اختبار هذه الفرضية، قام الباحث بإجراء عملية التحليل الإحصائي لبيانات عينة البحث الأساسية على عدة مستويات، على النحو التالي:

أولاً: تقويم مؤشرات الصعوبة لمفردات اختبار مستويات التفكير الهندسي:

استخدم الباحث قيمة النسبة المئوية للممتحنين الذين أجابوا على كل مفردة اجابة صحيحة من مفردات الاختبار، و النتائج يوضحها الجدول رقم (22):

الجدول (22)

يوضح قيم مؤشرات صعوبة مفردات اختبار مستويات التفكير الهندسي وفقا للعينة الأساسية ن=577.

الإدراكي أو البصري		التحليلي أو الوصفي		شبه الاستدلالي		الاستدلالي	
رقم المفردة	مؤشر الصعوبة	رقم المفردة	مؤشر الصعوبة	رقم المفردة	مؤشر الصعوبة	رقم المفردة	مؤشر الصعوبة
1	0.51	8	0.34	15	0.33	22	0.42
2	0.53	9	0.53	16	0.45	23	0.54
3	0.85	10	0.32	17	0.48	24	0.56
4	0.49	11	0.78	18	0.65	25	0.41
5	0.79	12	0.82	19	0.73	26	0.44
6	0.64	13	0.63	20	0.21	27	0.27
7	0.84	14	0.43	21	0.49	28	0.30

يتضح من الجدول رقم (22)، أن قيم مؤشرات صعوبة مفردات اختبار مستويات التفكير الهندسي تراوحت ما بين 0.21 و0.85 وبمتوسط حسابي بلغ 0.53، حيث اشارت مؤشرات الصعوبة لـ20 مفردة من الاختبار بأنها مناسبة وتراوحت ما بين 0.3 و0.7، أي ما نسبته 71.42% من مفردات الاختبار ككل، بينما اشارت 6 مفردات لمؤشر صعوبة سهلة، والتي تمثل ما نسبته 21.43% من مفردات الاختبار ككل، في حين أن مفردتين فقط تمثل ما نسبته 7.14% من مفردات الاختبار اشارت لمؤشرات صعوبتها بأنها صعبة جداً. فحسب كيلي (1939) Kelley فإن مؤشر الصعوبة الأقل من 0.20 يؤشر على أنها مفردة صعبة جداً و يجب أن تراجع، كما ينصح بعض الباحثين وخبراء القياس بأن مفردات الاختبار يجب أن تكون متدرجة في صعوبتها، بحيث تبدأ بالمفردات السهلة وتنتهي بالمفردات الصعبة خاصة في الاختبارات محكية المرجع، وبالتالي يجب أن تتراوح قيمها مؤشرات صعوبتها ما بين 10% و 90% بشرط أن يكون معدل الصعوبة للمفردات ككل في الاختبار في حدود 50% (أبوناهية، 1994، ص.310). وللتأكد من توفر هذا الشرط تم حساب متوسط صعوبة مؤشرات مفردات اختبار مستويات التفكير الهندسي، والتي بلغت 53%، وهي في الحدود التي أوصى بها الخبراء، وهذا ما يدل على تمتع مفردات اختبار مستويات التفكير الهندسي بمؤشرات صعوبة مقبولة.

ثانياً: تقويم مؤشرات التمييز (الحساسية) لمفردات اختبار مستويات التفكير الهندسي:

ولتقويم مؤشرات تمييز مفردات الاختبار، استخدم الباحث أسلوب مؤشر التمييز (ب)، والتي اقترحه برينان (1972) لحساب مؤشرات تمييز مفردات الاختبارات محكية المرجع، والنتائج يوضحها الجدول رقم(23):

الجدول (23)

يوضح قيم مؤشرات التمييز لمفردات اختبار مستويات التفكير الهندسي وفقا للعينة الإساسية ن=577.

الإدراكي أو البصري		التحليلي أو الوصفي		شبه الاستدلالي		الاستدلالي	
رقم المفردة	مؤشر التمييز	رقم المفردة	مؤشر التمييز	رقم المفردة	مؤشر التمييز	رقم المفردة	مؤشر التمييز
1	0.45	8	0.47	15	0.60	22	0.59
2	0.45	9	0.41	16	0.46	23	0.41
3	0.21	10	0.45	17	0.49	24	0.41
4	0.38	11	0.28	18	0.38	25	0.45
5	0.29	12	0.24	19	0.33	26	0.47
6	0.43	13	0.40	20	0.50	27	0.49
7	0.22	14	0.50	21	0.36	28	0.39

يتضح من الجدول رقم (23)، بأن مؤشرات التمييز لمفردات اختبار مستويات التفكير الهندسي جميعها مؤشرات ذات قيم موجبة. وهذا ما تتطلبه مؤشرات التمييز لمفردات الاختبارات المحكية، حيث سجلنا أعلى قيمة لهذا المؤشر 0.60 والتي كانت من نصيب المفردة 15، وأقل قيمة 0.21 من نصيب المفردة 3، أي أن قيم مؤشرات التمييز لمفردات اختبار مستويات التفكير الهندسي تراوحت ما بين 0.21 و0.60 بمتوسط حسابي قدره 0.41، حيث اشارت 23 مفردة من الاختبار والتي تمثل ما نسبته 82.14% من مفردات الاختبار الكلي، لمؤشرات تمييز جيدة وممتازة أكبر من القيمة 0.30، بينما اشارت 4 مفردات فقط والتي تمثل 14.28% من مفردات الاختبار لمؤشرات تمييز هامشية ما بين 0.20 و 0.29، في حين لم تشيرأي قيمة لمؤشرات تمييز مفردات الاختبار لأقل من 0.20، فحسب كيلي (1939) فإن قيم مؤشرات التمييز التي تتراوح ما بين 0.10 و 0.30 قيم مقبولة، والقيم ما فوق 0.30 فهي قيم جيدة (Kelley, T. L. 1939 , pp.17-24)، وهذا دليل على تمتع مفردات اختبار مستويات التفكير الهندسي بمؤشرات تمييز مناسبة ومقبولة وتتماشى مع افتراضات الباحث.

ثالثاً: تقويم ثبات اختبار مستويات التفكير الهندسي:

تم تقدير ثبات درجات اختبار مستويات التفكير الهندسي بعد تطبيقه على العينة الأساسية للبحث، باستخدام أساليب النظرية الكلاسيكية في القياس، وهي: الاتساق الداخلي، معادلة كودر– ريتشاردسون (KR_{20})، والخطأ المعياري للقياس(SEM)، وكذلك معادلة سبيرمان – وبراون لتصحيح الطول، اضافة إلى مؤشر ليفنجستون (Livingston,1972) والذي يستخدم في تقدير قيم ثبات الاختبارات محكية المرجع وبناءً على محك أداء معين.

1. التحقق من مؤشر الاتساق الداخلي لمفردات الاختبار:

تم تقدير مؤشر الاتساق الداخلي لمفردات اختبار مستويات التفكير الهندسي وفقا للعينة الأساسية، باستخدام معامل الارتباط لبيرسون بحساب العلاقة بين الدرجة الكلية للاختبار ودرجة كل مفردة من مفردات الاختبار ككل، والتي يعتبرها البعض مؤشرات أولية للثبات (تيغزة، 2009، ص.644)، والقيم يوضحها الجدول رقم (24):

الجدول (24)

يوضح قيم معامل الارتباط (بيرسون) بين درجة كل مفردة والدرجة الكلية وفقاً للعينة الأساسية ن=577.

معامل الارتباط بين درجة كل مفردة من الاختبار والدرجة الكلية وفقا للعينة الاساسية.							
المفردة	م . الارتباط	المفردة	م . الارتباط	المفردة	م . الارتباط	المفردة	م . الارتباط
1	**0.344	8	**0.303	15	**0.463	22	**0.371
2	**0.344	9	**0.317	16	**0.308	23	**0.341
3	**0.329	10	**0.373	17	**0.369	24	**0.333
4	**0.347	11	**0.393	18	**0.319	25	**0.259
5	**0.337	12	**0.359	19	**0.337	26	**0.311
6	**0.389	13	**0.441	20	**0.510	27	**0.314
7	**0.263	14	**0.338	21	**0.200	28	**0.165

**دالة احصائيا عند مستوى (0.01).

يتضح من الجدول رقم (24) أن قيم معاملات الارتباط لبيرسون تراوحت ما بين 0.51 كأعلى قيمة، و0.17 كأدنى قيمة، وجميعها قيم موجبة ومقبولة وذات دلالة إحصائية عند مستوى α=0.01. ويعد هذا دليلاً على مدى فاعلية مفردات اختبار مستويات التفكير الهندسي، وبأنها تعمل في إتجاه واحد، في تحديد

التكوين الفرضي المقاس الذي ننوي قياسه، أي أنها تقيس سمة كامنة أحادية البعد. والتي يعتبرها البعض مؤشر للثبات.

2. التحقق من الثبات باستخدام الأساليب التالية: (مؤشر ليفنجستون؛ معامل كودر- ريتشاردسون؛ الخطأ المعياري للقياس؛ معامل سبيرمان- وبراون لتصحيح الطول):

استخدم الباحث مجموعة من الأساليب والتي نريد من خلالها تقدير مؤشرات ثبات درجات اختبار مستويات التفكير الهندسي بعد تطبيقه على العينة الكلية، والجدول رقم (25) يوضح هذه المؤشرات:

الجدول (25)

يوضح قيم معامل الثبات لاختبار مستويات التفكير الهندسي وفقاً للعينة الأساسية 577 تلميذاً وتلميذة.

مستويات(فان هيل) للتفكير الهندسي	معامل كودر- ريتشاردسون KR_{20}	الخطأ المعياري للقياسSEM	معامل سبيرمان - وبراون لتصحيح الطول	مؤشر ليفنجستون (Livingston)	
				محك 7/5	محك7/6
الإدراكي- البصري	0.33	1.15	0.65	0.36	0.65
التحليلي- الوصفي	0.40	1.16	0.72	0.62	0.80
الاستنباط غير الشكلي	0.47	1.18	0.77	0.74	0.85
الاستنباط الشكلي	0.28	1.25	0.60	0.75	0.86
الاختبار الكلي	0.71	2.35			

يتضح من الجدول رقم (25)، أن قيم معامل كودر- ريتشاردسون (Kr_{20}) أو بما يسمى تحليل التباين والذي يعتمد على مدى الإتفاق بين الاستجابات على كل مفردة في الاختبار وغيرها من المفردات في الاختبارات الفرعية أو (المستويات) الأربعة المكونة للاختبار، والذي بلغ القيم التالية: 0.33؛ 0.40؛ 0.47؛ 0.28 على التوالي. والتي تشير إلى إنخفاض قيمه نوعاً ما بعكس قيمة معامل الثبات للاختبار الكلي والذي بلغ 0.71.

في حين اشارت تقديرات مؤشر الثبات لـ ليفنجستون والذي يعتمد على الدرجة الفاصلة أو المحك بدلاً من المتوسط الحسابي وانحراف درجة الفرد عن هذا المتوسط الذي تعتمد عليه الاختبارات معيارية المرجع، فالاختبارات محكية المرجع تركز على انحراف درجة الفرد عن المحك المحددة، وبالتالي تتأثر قيم الثبات بهذا المحك، لهذا أشارت قيم هذا المؤشر في المستوى الأول وعند المحك 7/5 أي 71% إلى 0.38 وهي قيمة منخفضة نوعاً ماً، في حين بلغت 0.65 باستخدام المحك 7/6 أي 86% أو بما يسمى المحك الصارم،

وهي قيم متوسطة للثبات، في حين بلغت قيمة الثبات بهذا المعامل للمستوى الثاني 0.62 عند المحك 7/5، والقيمة 0.80 باستخدام المحك 7/6، وهي قيم مقبولة و جيدة لمعامل الثبات، كما سجلنا القيم 0.74 و0.75 للمستوى الثالث والمستوى الرابع على التوالي باستخدام المحك 7/5، والقيم 0.85 و0.86 باستخدام المحك 7/6، وجميعها تشير إلى أنها قيم مرتفعة وجيدة لمعامل ثبات المستويات الأربعة المكونة الاختبار مستويات التفكير الهندسي.

كما استخرج الباحث الخطأ المعياري للقياس باعتباره مؤشراً آخر على ثبات الاختبار وفق التفسير المعياري المرجع، حيث تراوحت قيم الأخطاء المعيارية للقياس للمستوى الأول 1.15؛ و1.16 للمستوى الثاني؛ و1.18 للمستوى الثالث؛ و1.25 للمستوى الرابع، كما بلغ الخطأ المعياري للقياس للاختبار ككل 2.35 وهي قيم منخفضة نوعاً ماً، مما يدل على أن مؤشر ثبات الاختبار ككل، ومستوياته الأربعة مقبول عموماً.

رابعاً: تقويم صدق اختبار مستويات التفكير الهندسي:

تم التحقق من صدق الاختبار بعدة أساليب، وهي:

1. الصدق الوصفي أو الصدق بدلالة المحتوى:

تم تقدير صدق محتوى اختبار مستويات التفكير الهندسي، بعد إتمام إجراءات بناءه وفقاً لنظرية (فان هيل) للتفكير الهندسي، و عرض مفرداته المكونة له في صورته الأولية ومعها تعريف مفصل لنموذج (فان هيل) ولكل مستوى من مستوياته الأربعة الأولى لتطور التفكير الهندسي، على مجموعة من المحكمين الخبراء في مجال تدريس مادة الرياضيات، منهم الحاملين لشهادة الدكتوراه والماجستير تخصص الرياضيات، ومفتشين وأساتذة رئيسيين وأساتذة لهذه المادة في المرحلة الثانوية أنظر الملحق رقم (02)، حيث خضعت المفردات إلى تحكيم وتعديل في عملية تبادلية من خلال الحكم على كل مفردة بمدى مطابقتها مع محتوى وأهداف كل مستوى من مستويات (فان هيل) الأربعة، والتأكد مما إذا كانت مفردات الاختبار تمثل فعلا محتوى المجال أو (المستوى) وأهدافه، وهل أن هذا التمثيل هو تمثيلاً صادقاً. وصولاً إلى حد كبير ودرجة مقبولة من الصدق الوصفي (أو صدق المحتوى)، كما توضحه نتائج الجدول رقم (13) المشار إليه سابقاً، باستخدام معادلة لاوش والتي بلغت قيمتها 0.74 وهي قيمة جيدة ومقبولة ودالة إحصائياً عند مستوى 0.05

حسب (لاوش)، مع الإبقاء على 28 مفردة مكونة للاختبار الذي تم تطبيقه على عينة استطلاعية ومن ثم العينة الأساسية.

2. الصدق الوظيفي أو المرتبط بالمحك:

استخدم الباحث المعدل السنوي للتحصيل الدراسي في مادة الرياضيات كمحك للتحقق من الصدق التنبوئ لاختبار مستويات التفكير الهندسي، وهو من أكثر المحكات استخداماً لقياس ارتباطه باختبارات الذكاء والقدرات العقلية، نظراً للتأثير الكبير للخبرات التعليمية في نمو وتطور أنماط التفكير المختلفة. اقتصرنا على بيانات عينة مكونة من 459 تلميذاً وتلميذة من العينة الأساسية من تلامذة المرحلة الثانوية الشعب العلمية، بمتوسط حسابي لأعمارهم يساوي 19 سنة، وانحراف معياري قدره 2.41 ، ممن استطاع الباحث الحصول على معدلاتهم السنوية في مادة الرياضيات، وبحساب معامل الارتباط بيرسون بين المعدل السنوي للتلامذة في مادة الرياضيات، ودرجاتهم بعد تم تطبيق اختبار مستويات التفكير الهندسي عليهم بداية الفصل الثاني للسنة الدراسية، والذي بلغ القيمة 0.26 وهي قيمة منخفضة نوعاً ما، ولكنها قيمة موجبة ودالة إحصائياً عند مستوى دلالة 0.01.

3. صدق انتقاء النطاق السلوكي أو صدق التكوين الفرضي:

يعتبر أحد التطبيقات لمحك الاتساق الداخلي ارتباط درجات الاختبارات الفرعية (المستويات) بالدرجة الكلية، فارتباطات الاتساق الداخلي سواءً اعتمدت على المفردات أو الاختبارات الفرعية هي بدرجة أساسية مقاييس للتجانس. ونظراً لأن درجة تجانس اختبار معين تفيد في وصف النطاق السلوكي أو السمة التي يعاينها الاختبار، فإن هذه الدرجة توائم إلى حد ما صدق التكوين الفرضي للاختبار. (أناستازي وسوزان، 2015، ص.169) (الكبيسي، 2011، ص.267)

استخدم الباحث معامل الارتباط بيرسون بين درجة كل مفردة والدرجة الكلية لكل مستوى من مستويات (فان هيل) للتفكير الهندسي بعد حذف درجة المفردة من البعد، والنتائج يوضحها جدول رقم (26):

الجدول (26)

يوضح قيم معاملات الارتباط بيرسون بين درجة كل مفردة والدرجة الكلية لكل مستوى من مستويات (فان هيل) للتفكير الهندسي بعد حذف درجة المفردة من البعد. ن=577.

الإدراكي		التحليلي		شبه الاستدلالي		الاستدلالي	
المفردة	م . الارتباط	المفردة	م . الارتباط	المفردة	م . الارتباط	المفردة	م . الارتباط
1	**0.500	8	**0.420	15	**0.572	22	**0.492
2	**0.501	9	**0.456	16	**0.466	23	**0.460
3	**0.365	10	**0.430	17	**0.534	24	**0.496
4	**0.451	11	**0.468	18	**0.491	25	**0.426
5	**0.410	12	**0.445	19	**0.453	26	**0.403
6	**0.526	13	**0.503	20	**0.504	27	**0.396
7	**0.332	14	**0.518	21	**0.387	28	**0.338

** دالة احصائيا عند مستوى 0.01.

يتضح من الجدول رقم (26) أن قيم معاملات الارتباط بيرسون لمفردات اختبار مستويات التفكير الهندسي جميعها ذات دلالة إحصائية عند مستوى α0.01، والتي وقعت قيمها في المدى ما بين 0.33 إلى0.57 أي أنها تجاوزت جميعاً محك إيبل (Ebell) والمحدد بالقيمة 0.30. وتشير هذه المعاملات إلى درجة مقبولة من التجانس أو الاتساق الداخلي للمفردات المكونة للاختبار. ويعد هذا دليلاً آخر على مدى فاعلية مفردات الاختبار، في قياسها ما يقيسه البعد أو المستوى الذي تقع فيه، أي أن هناك انسجاماً كبيراً بين ما تقيسه المفردة وما يقيسه البعد أو المستوى. اضافة إلى ذلك تم حساب معامل الارتباط بيرسون بين الدرجة الكلية لكل مستوى من المستويات الأربعة التي يقيسها الاختبار، و الاختبار ككل.

الجدول (27)

يوضح قيم معاملات الارتباط بيرسون بين الدرجات الكلية لمستويات اختبار مستويات التفكير الهندسي الاربعة، ودرجة الاختبار ككل ن=577.

المستوى	التحليلي- الوصفي	الاستنباط غير الشكلي	الاستنباط الشكلي	الاختبار الكلي
الإدراكي- البصري	**0.385	**0.200	**0.288	**0.635
التحليلي- الوصفي		**0.381	**0.330	**0.736
الاستنباط غير الشكلي			**0.330	**0.709
الاستنباط الشكلي				**0.693

** دال إحصائياً عند مستوى دلالة (0.01).

يتبين من الجدول رقم (27)، أن قيم معامل الارتباط بيرسون بين الدرجة الكلية للاختبار والدرجة الكلية لكل مستوى من مستويات اختبار مستويات التفكير الهندسي الأربعة، بأنها معاملات ارتباط قوية والتي تراوحت في المدى بين 0.64 إلى 0.74 وجميعها قيم مرتفعة و دالة إحصائياً عند مستوى معنوية α0.01. أي أنها تجاوزت جميعاً محك إيبل (Ebell) والمحدد بالقيمة 0.30. وتشير هذه المعاملات مرة آخرى إلى درجة مقبولة من التجانس أو الاتساق الداخلي للمستويات الأربعة المكونة للاختبار. كما يتضح أن معاملات الارتباط بين المستويات الأربعة في ما بينها منخفضة بالرغم من دلالتها الإحصائية، وهذا دليل على أن الاختبار يقيس فعلاً أربعة مستويات متمايزة حسب عملية بنائه وفقاً لنظرية (فان هيل).

1-2- عرض وتحليل نتائج الفرضية الثانية:

والتي تنص على أن:" تتوفر لبيانات اختبار مستويات التفكير الهندسي المستمدة من استجابات العينة على ملاءمة مقبولة لنموذج راش".

من أجل اختبار صحة هذه الفرضية، قام الباحث بتفحص الافتراضات الرئيسة التي يتطلبها نموذج راش، وهي: (أولاً) أحادية البعد؛ (ثانياً) الدالة المميزة لكل مفردة والتي تصف العلاقة الوتيرية بين القدرة والأداء على المفردة، والتي تسمى اختصاراً " دالة خصائص المفردة "؛ (ثالثاً) الاستقلال الموضعي؛ (رابعاً) التخمين؛ (خامساً) التحرر من عامل السرعة.

أولاً: التحقق من افتراضات نموذج راش أحادي البارامتر:

1- تقويم افتراض أحادية البعد:

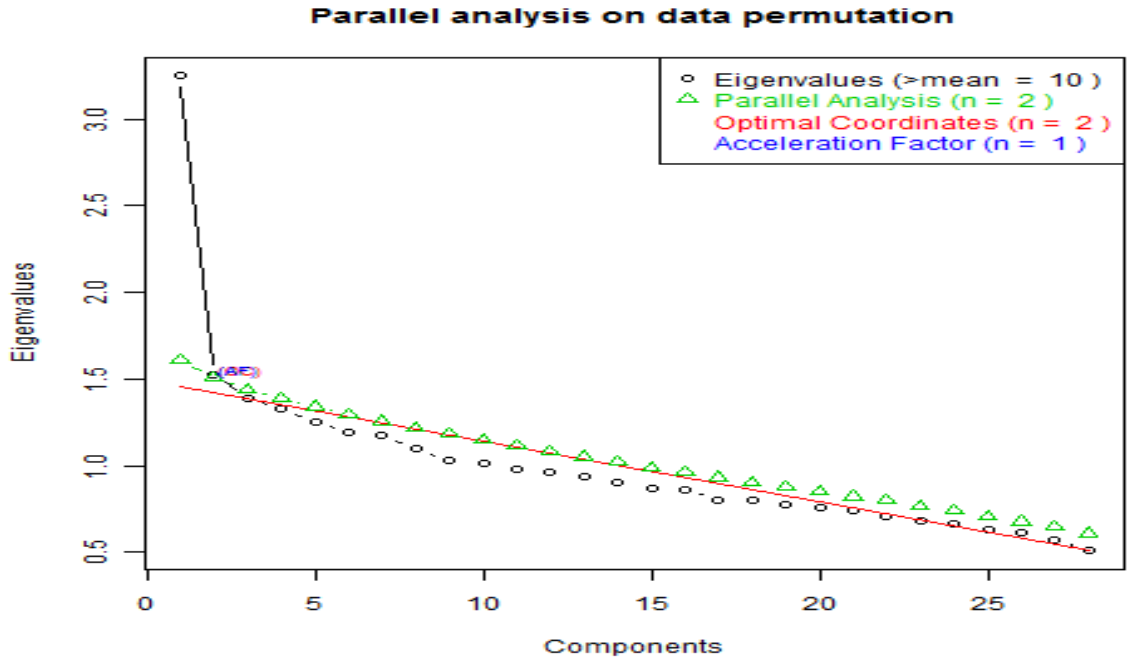
قام الباحث باختبار هذا الافتراض بطريقتين أومنهجيتين مختلفتين، وهما: طريقة التحليل العاملي الاستكشافي (النظرية الكلاسيكة)؛ وطريقة تحليل نموذج (راش) للمكونات الأساسية المعتمدة على البواقي (PCAR).

1-1- طريقة التحليل العاملي الاستكشافي (النظرية الكلاسيكة):

باستخدام التحليل العاملي الاستكشافي (التحليل العاملي الموازي)، بطريقة المكونات الأساسية مع تدوير المحاور بطريقة الفاريمكس لدرجات التلامذة، باستخدام البرنامج الإحصائي (Spss v25)، والتي تم فيها تسجيل قيمة لمحدّد مصفوفة معاملات الارتباط المحسوبة بين استجابات أفراد عينة البحث وبين مفردات الاختبار بلغت 0.108، وهي قيمة تزيد عن الحد الأدنى المقبول وهو0.00001، في حين بلغت قيمة مؤشر"

كايزر– ماير أولكين " (KMO) للكشف عن كفاية حجم العينة 0.73، وهي قيمة تزيد عن الحد الأدنى المقبول لاستخدام أسلوب التحليل العاملي وهو0.50، كما تم قياس ملاءمة المعاينة (MSA) والموجودة في قطر مصفوفة معاملات الارتباط الصورية، والملاحظ بالنسبة لنتائج العينة الحالية أن جميع القيم الحرجة أكبر من 0.50 والتي تتراوحت ما بين 0.55 و0.83، كما أشارت قيم اختبار بارتلت لدلالتها الإحصائية عند مستوى 0.0001، و قيمة لاختبار كاي مربع (كا^2) تساوي 846.474 عند درجة حرية 378. (تيغزة، 2011، ص.294).

أسفر التحليل العاملي الاستكشافي، وباعتماد أسلوب التحليل الموازي وتقنية العامل المتسارع (Accel-eratiion Factor) لكلٍ من رايش وروبيل وبليس (Raiche, Roipel, and Blais, 2006)، والتي تظهر نتائجه كما يوضحها الشكل رقم (24):



الشكل(24) يوضح نتائج التحليل العاملي المتوازي، و الجذر الكامن، والإحداثيات المثلى، والعامل المتسارع.

يظهر من الشكل رقم (24)، تغير في ميل المنحنى بشكل مفاجئ محدداً لنا لعامل واحد متسارع (Matthew and el, 2013, p.03). كما أن الجذر الكامن الأول كانت قيمته 3.251، فسر 11.61% من التباين الكلي، وأن ناتج قسمة الجذر الكامن للعامل الأول على العامل الثاني الذي جذره الكامن 1.524 والذي فسر ما نسبته 5.444% من التباين الكلي، وهي قيمة أكبر من 2. وهو ما يمكن اعتباره عاملاً سائداً حسب همبلتون (Hamblet, 2004).

1-2- طريقة تحليل نموذج (راش) للمكونات الأساسية المعتمد على البواقي:

نظراً لأن نموذج راش يفترض أحادية البعد إلا أن أحادية البعد ليست مطلقة، ويجب علينا أن ننظر إلى مفهوم أحادية البعد كما يعنيها التحليل العاملي، حيث يختلف الهدف بينهما، فالتحليل العاملي يهدف إلى تحديد العوامل التي يتكون منها الاختبار، ولكن نظرية الاستجابة للمفردة تهدف إلى التعرف على هل أن الانحرافات عن السمة المقاسة ترقى إلى أن تكون عامل مستقل أم لا؟ لذلك يوفر برنامج وينستبس تحليل نموذج (راش) للمكونات الأساسية المعتمدة على البواقي (PCAR)، والذي يظهر الاختلافات بين الأبعاد، والجدول رقم (28) يلخص أهم مؤشرات نتائج هذا التحليل:

الجدول (28)

نتائج التحليل العاملي للمكونات الأساسية للبواقي(PCAR) باستخدام نموذج (راش) قبل التدريج.

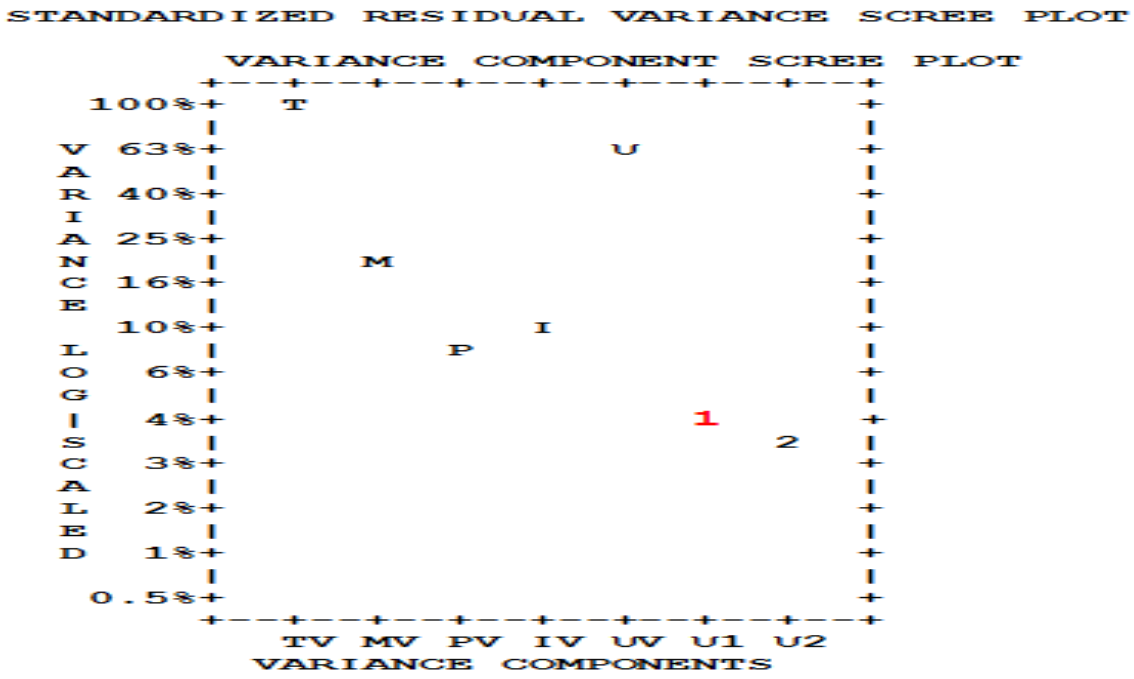
حجم تباين البواقي المعيارية مقدرا بوحدات القيم Eigen.	الجذر الكامن	الملاحظ	المتوقع
التباين الكلي في الاستجابات	35.8189	%100	%100
التباين الذي فسره العامل الرئيسي، تقديرات نموذج (راش)	7.8189	% 21.8	%22
التباين المفسر بواسطة الأفراد	3.3385	%9.3	%9.4
التباين المفسر بواسطة المفردات	4.4804	%12.5	%12.6
مجموع التباين غير المفسر	28.0000	%78.2	%78.0
التباين الذي فسره العامل الثاني (الأول في البواقي)	1.6894	% 4.7	–

يظهر الجدول رقم (28) مؤشرات التحليل العاملي بالمكونات الأساسية للبواقي باستخدام نموذج (راش)، وللحكم على أحادية البعد تم التحقق من بعض المحكات، وهي: (Linacre, 2008, p.272)

1. المحك الأول، و كقاعدة عامة إذا كانت قيمة التباين المفسر بواسطة التقديرات (Raw variance explained by measures) أكبر من أو تساوي 50 %، واسترشاداً بما أكده إلينا لينكر مطور برنامج وينستبس- أنظر الملحق رقم (10) - حيث حدد المجال من 20 % إلى 80 % بأنه مجال مقبول ويعد مؤشراً قوياً لأحادية البعد. وهو ما تحقق في نتائج بحثنا الحالي حيث أشار مؤشر التباين الذي فسره العامل الرئيسي، لتقديرات نموذج (راش) للقيمة الملاحظة 21.8 % والفرق بينه وبين القيمة المتوقعة للنموذج 0.2% فقط.

2. المحك الثاني، والذي يحدد قيمة نسبة التباين الذي يفسره العامل الثاني (الأول في البواقي) (Unexplained Variance in 1st contrast) بأقل من القيمة 5%، وهذا ما تحقق في البحث الحالي بالقيمة 4.7%، وهذا دليل آخر على أحادية البعد.

3. المحك الثالث، والذي يحدد بقيمة الجذر الكامن لنسبة التباين الذي يفسره العامل الثاني (Unexplained Variance in 1st contrast)، بأن لا تتعدى أو تكون أقل من القيمة 3، وهذا ما تحقق في البحث الحالي بتسجيلنا للقيمة 1.6885 وهي أقل من القيمة 3 المحددة. والشكل رقم (25) يوضح نتائج التحليل العاملي بالمكونات الأساسية للبواقي باستخدام نموذج (راش):



الشكل (25) يوضح خريطة اختبار (Scree Plot) لتباين البواقي المعيارية بواسطة برنامج (winsteps).

يتضح من خلال المؤشرات الإحصائية في الجدول رقم (28)، وكذلك الشكل رقم (25) بأن محكات التحليل المستخدمة تحققت كلياً، وهو ما يؤكد على أن اختبار مستويات التفكير الهندسي، له تكوين فرضي ذو فضاء كامن أحادي البعد.

ثانياً: التحقق من افتراض دالة استجابة المفردة:

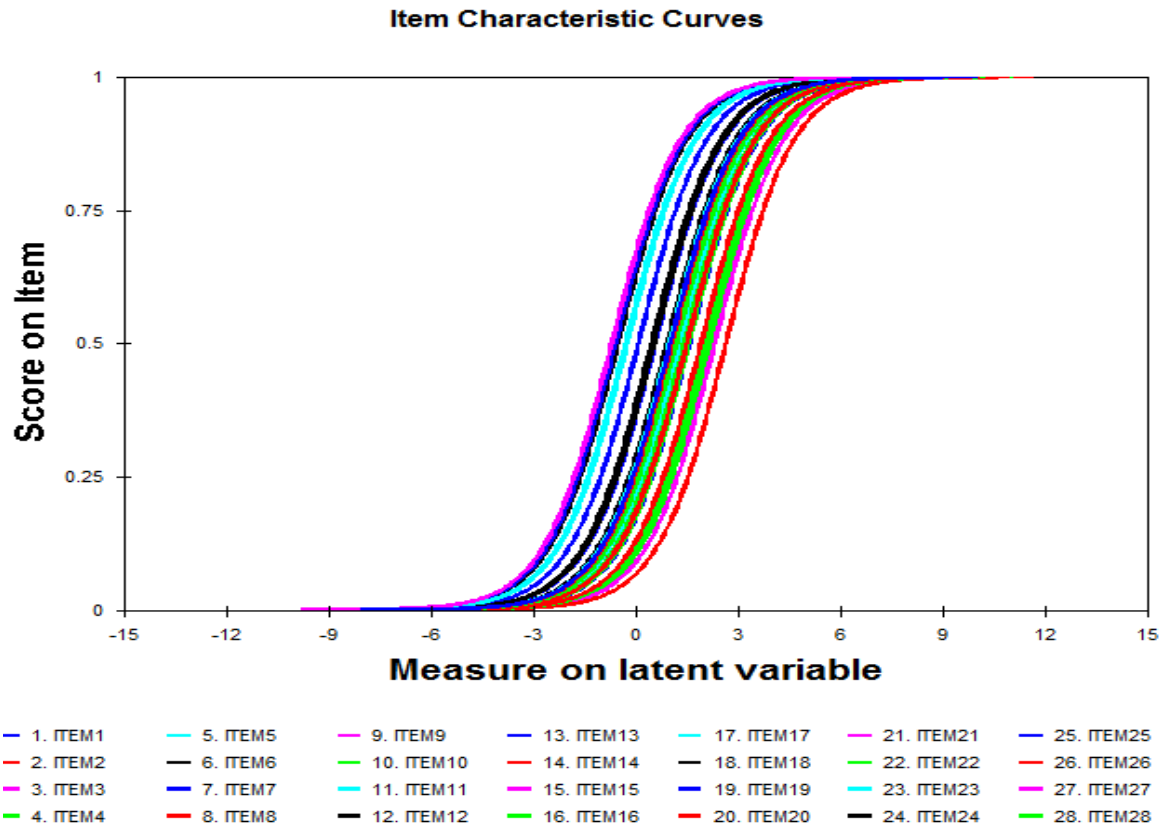
استخدم الباحث البرنامج الإحصائي وينستبس والذي يوفر تقدير معلم التمييز، ولكن فقط في حالة طلب ذلك منه، والجدول رقم (29) يوضح نتائج مؤشرات تمييز مفردات اختبارمستويات التفكير الهندسي.

الجدول (29)

يظهر قيم مؤشرات التمييزلاختبار مستويات التفكير الهندسي.

رقم المفردة	مؤشر التمييز	رقم المفردة	مؤشر التمييز	رقم المفردة	مؤشر التمييز
1	0.94	11	1.10	21	0.56
2	0.59	12	1.09	22	0.98
3	1.05	13	1.30	23	1.06
4	0.56	14	0.97	24	0.91
5	1.09	15	1.17	25	0.87
6	1.12	16	0.89	26	0.77
7	1.03	17	1.15	27	1.01
8	0.83	18	0.95	28	0.74
9	1.06	19	1.04		
10	1.01	20	1.19		

يتضح من الجدول رقم (29) أن مؤشرات تمييز مفردات اختبار مستويات التفكير الهندسي، كلها قيم أعلى من القيمة 0.50، والتي تراوحت مابين 0.56–1.30، وحسب كلٍ من (لينكر، ماسترز، ووليد)، فإن قيم التمييز المقبولة يجب أن تكون ما بين 0.5 – 2، و يمكن أن تكون من 0.5 إلى ما لانهاية. (Linacre, 2012, p.596) (Masters, 1988 ,pp.15-29) (وليد، 2014، ص.67). ويرى لينكر أن متوسط قيم تمييز (الميل، Slope) المفردات في نموذج (راش) يجب أن يكون قريب من 1.0 تقريباً. (Linacre, 2012, p.597). وهذا ما أشارت إليه القيم في بحثنا الحالي حيث أن متوسط قيم مؤشرات التمييز كان 0.98، وهي قيمة قريبة جداً من الواحد. و تظهر مؤشرات تمييز المفردات على شكل منحنيات بيانية توضح احتمال التوصل إلى الاستجابة الصحيحة على المفردة كدالة للسمة الكامنة في ضوء الأداء على مفردات الاختبار، وأنه في تطبيقات نظرية الاستجابة للمفردة يفترض أن منحنيات خصائص المفردة لها شكل حرف (S)، وهو ما يوضحه الشكل رقم (26)، لخصائص مفردات الاختبار المستخرجة بواسطة برنامج وينستبس.



الشكل (26) منحنى خصائص مفردات اختبار مستويات التفكير الهندسي 28 مفردة.

يتضح من الشكل رقم (26) أن الدرجة تتزايد بزيادة السمة الكامنة، وبالتالي تزداد احتمالية التوصل إلى الاستجابة الصحيحة. ومن المؤشرات التي تدل كذلك على تحقق تكافؤ مؤشرات تمييز المفردات ومطابقتها للنموذج، حسب ما أشار إليه هامبلتون وسوامينثان (Hambelton and Swuminathan, 1985) هو أن تكون قيمها واقعة ضمن حدود المدى لمتوسط معاملات التمييز يساوي 0.15، وبما أننا وجدنا قيمة متوسط معاملات التمييز 0.32، وبالتالي فهي في حدود المدى ما بين 0.18–0.50.

كما أسفرت نتائج التحليل الموضحة في الجدول رقم (34) لقيم لمعاملات الارتباط الثنائي المتسلسل (ρ_{pbis})، والتي كان جميعها ضمن المدى المقبول، و توزيعها إلى حد ما متجانس، مما يدل على تساوي مؤشرات تمييز مفردات الاختبار، مع تسجيل قيم صغيرة للانحراف المعياري لهذه المؤشرات، وهو مؤشر آخر على تحقق هذا الافتراض.

ثالثاً: التحقق من افتراض الاستقلال الموضعي (المحلي):

للتحقق من هذا الافتراض، استخدم البرنامج الإحصائي وينستبس في استخراج المؤشر الاحصائي (Q3)، والنتائج يوضحها الجدول رقم (30):

الجدول (30)

يوضح قيم المؤشر الإحصائي (Q3) من خلال قيم ارتباط البواقي المعيارية.

أزواج المفردات	الارتباط	أزواج المفردات	الارتباط
المفردة 15 مع المفردة 10	0.17−	المفردة 13 مع المفردة 4	0.12−
المفردة 15 مع المفردة 2	0.16−	المفردة 26 مع المفردة 20	0.12−
المفردة 19 مع المفردة 4	0.13−	المفردة 24 مع المفردة 15	0.11−
المفردة 18 مع المفردة 4	0.12−	المفردة 27 مع المفردة 23	0.11−
المفردة 17 مع المفردة 2	0.12−	المفردة 19 مع المفردة 2	0.11−
المفردة 14 مع المفردة 1	0.12−	المفردة 26 مع المفردة 18	0.11−
المفردة 21 مع المفردة 6	0.12−	المفردة 13 مع المفردة 8	0.11−
المفردة 23 مع المفردة 2	0.12−	المفردة 15 مع المفردة 12	0.11−
المفردة 16 مع المفردة 3	0.12−	المفردة 21 مع المفردة 9	0.11−
المفردة 16 مع المفردة 6	0.12−	المفردة 25 مع المفردة 14	0.11−

يتضح من الجدول رقم (30)، أن قيم مؤشر الاحصائي (Q3) تراوحت ما بين −0.11 و−0.17، بمتوسط حسابي يساوي −0.12، وعموما فإن هذه القيم لم تتجاوز مستوى الإنتهاك 0.30 حسب زينسكي وهامبلتون وسيريسي (Zenisky and Hammbleton and Sireci, 2002)، وكذلك المجال الذي اقترحه تشن وتيسن (Chen and Thissen, 1997) وهو أن تكون القيمة ما بين 0.20 و−0.20. ومنه يمكننا الحكم على أن مفردات اختبار مستويات التفكير الهندسي لا توجد فيها مفردات تتأثر الإجابة عنها بمفردات أخرى من مفردات نفس الاختبار، مما يعني عدم تداخل منحنيات الخصائص المميزة للمفردات كذلك، وهو ما يشار إليه بالاستقلال الموضعي.(Linacre, 2012, p.405).

رابعاً: التحقق من افتراض تدني عامل التخمين:

تم استخدام البرنامج الإحصائي وينستبس، والذي يوفر تقدير معلمة التخمين للمفردات، ولكن فقط في حالة طلب ذلك منه، وتظهر قيم معلمة التخمين تحت العمود "التخمين الأدنى"، والجدول رقم (31) يوضح نتائج تقدير مؤشرات التخمين لمفردات اختبار مستويات التفكير الهندسي.

الجدول (31)

يوضح مؤشرات التخمين لمفردات اختبار مستويات التفكير الهندسي.

رقم المفردة	مؤشر التخمين		رقم المفردة	مؤشر التخمين		رقم المفردة	مؤشر التخمين	
	الأعلى	الأدنى		الأعلى	الأدنى		الأدنى	الأدنى
1	1.00	0.08	11	1.00	0.09	21	0.90	0.10
2	0.88	0.06	12	1.00	0.00	22	0.93	0.00
3	1.00	0.00	13	1.00	0.00	23	1.00	0.00
4	0.92	0.12	14	0.96	0.00	24	0.98	0.02
5	1.00	0.00	15	1.00	0.00	25	0.87	0.00
6	1.00	0.00	16	0.99	0.03	26	0.99	0.07
7	1.00	0.00	17	1.00	0.00	27	1.00	0.00
8	0.82	0.03	18	0.99	0.10	28	0.83	0.06
9	1.00	0.02	19	1.00	0.09			
10	1.00	0.01	20	1.00	0.00			

يتضح من الجدول رقم (31)، أن جميع قيم مؤشرات التخمين الأدنى لمفردات اختبار مستويات التفكير الهندسي أقل من القيمة 0.10، ما عدا المفردة 4 والتي سجلت القيمة 0.12، وهي على العموم قيم منخفضة وقريبة من الصفر وفي المجال المقبول لقيم التخمين الذي حدده كل من (Linacre, 2012, p.112) (Steven, 2003, pp.164-184) (وليد، 2014، ص.68) (Baker, 2001, p.28).

خامساً: التحقق من التحرر من عامل السرعة:

تم التأكد عملياً من هذا الافتراض من خلال تطبيق الباحث للاختبار، و أن فشل التلامذة في الوصول إلى الإجابة الصحيحة عن مفردات الاختبار يعود فقط لإنخفاض قدراتهم، ولا يعود إلى ضيق وقت الاختبار أو لعوامل السرعة في الاستجابة، وذلك بملاحظة الإجابات المتروكة والتي توزعت عشوائياً على طول مفردات الاختبار ولم تتمركز في منطقة معينة، مما يشير إلى أن زمن الاختبار لم يؤثر على استجابات التلامذة، وأن مدة (1) ساعة كانت كافية للإنتهاء من الإجابة على جميع مفردات الاختبار بالنسبة لأغلبية التلامذة، كما أنه لم يظهر أي تلميذ أي شكوى أثناء التطبيق من ضيق الوقت وعدم كفايته، فإن هذا يعني أن الاختبار موضع البحث يقيس القوة و لا يقيس السرعة.

1-3- عرض وتحليل نتائج الفرضية الثالثة:

والتي تنص على أن:" تتباين تقديرات صعوبة مفردات اختبار مستويات التفكير الهندسي وتدرجها باستخدام نموذج راش".

للتحقق من صحة هذا الفرض، قام الباحث بتحليل استجابات التلامذة اعتماداً على برنامج و ينستبس والذي يستخدم طريقة تقدير الأرجحية القصوى المشتركة (JMLE) في تقدير بارامترات النموذج، حيث يقوم البرنامج بعملية " التحرير" لتدريج أو معايرة الاختبار من خلال حذف الاستجابة التامة، وهي حذف استجابات الأفراد الذين أجابوا إجابة صحيحة على جميع مفردات الاختبار، وكذلك حذف الأفراد الذين لم يجيبوا على أي مفردة إجابة صحيحة، كما يقوم البرنامج كذلك بحذف المفردات التي نجح جميع التلامذة في الإجابة عليها إجابة صحيحة، أو التي لم يجب عليها أي تلميذ إجابة صحيحة، ولم تظهر عملية "التحرير" أي حالة من الحالات السابقة، لذلك نجحت مفردات اختبار مستويات التفكير الهندسي جميعها في تحليل استجابات أفراد العينة 577 تلميذاً وتلميذة، حيث تم تحليل استجابات جميع الأفراد عن جميع مفردات اختبار مستويات التفكير الهندسي بهدف الكشف عن الأفراد غير المطابقين للنموذج، ولتحقيق ذلك تم تقدير قدرة كل فرد، والخطأ المعياري في مقياس القدرة، وقيم إحصائي المطابقة الداخلية (Infit ZSTD)، أو إحصائي متوسط المربعات التقاربي (Infit MNSQ)، وهو مؤشر إحصائي للسلوكات غير المتوقعة التي تؤثر في الاستجابات عن المفردات التي تكون قريبة من مستوى قدرة الفرد. كما تم تقدير قيم إحصائي المطابقة الخارجية (Outfit ZSTD)، أو إحصائي متوسط المربعات التباعدي (Outfit MNSQ)، وهو مؤشر إحصائي أكثر حساسية للسلوكات غير المتوقعة التي تؤثر في الاستجابات عن المفردات التي تكون بعيدة عن مستوى قدرة الفرد، وذلك لكل تقدير من تقديرات القدرة، والجدول رقم (32) يوضح ملخص لنتائج تحليل قدرات الأفراد.

الجدول (32)

ملخص نتائج التحليل لقدرات الأفراد.

الإحصاءات	الدرجة الخام Raw Score	القدرة Measure	الخطأ المعياري للنموذج Model S.E	إحصائي المطابقة الداخلية Infit		إحصائي المطابقة الخارجية Outfit	
				متوسط المربعات MNSQ	قيمة الإحصائي ZSTD	متوسط المربعات MNSQ	قيمة الإحصائي ZSTD
المتوسط الحسابي	14.5	0.17	0.44	1.00	0.0	0.99	0.0
الإنحراف المعياري	4.1	0.77	0.04	1.15	0.9	0.22	0.9
أعلى علامة	26	2.87	0.75	1.50	2.8	1.82	2.5
أدنى علامة	4	2.02–	0.41	0.62	2.7–	0.51	2.6–

يتضح من الجدول رقم (32)، أن قيم المتوسط الحسابي لمتوسطات المربعات الداخلية والخارجية (MNSQ) يساوي الواحد بالنسبة لمتوسطات المربعات الداخلية، وقريب جداً من الواحد بالنسبة لمتوسطات المربعات الخارجية، وأن المتوسط الحسابي لمتوسط إحصائي المطابقة الداخلية والخارجية (ZSTD) تساوي الصفر، وهو الوضع المثالي والذي يفترضه النموذج (0، 1)، كما أن الإنحراف المعياري قريب من الواحد.

وعند تفحص قيم إحصائي المطابقة الخارجية والداخلية للأفراد، التي تشير إلى مطابقة قدرة الفرد مع قدرات مجموعة الأفراد التي ينتمي إليها في قياس السمة التي يقيسها الاختبار، فإذا كانت قيمة هذان الإحصائيان تزيد على +2 تعتبر قدرة الفرد غير مطابقة لقدرات مجموعة الأفراد (Alastair and Hutchinson,1987 ; julian,1988)، حيث تبين أن 11 فرداً غير مطابقين للنموذج، تزيد قيم إحصائي المطابقة الخارجية والداخلية لدرجاتهم على القيمة +2، أو قيم متوسطات المربعات المقابلة لدرجاتهم تزيد على 1، وهي القيم التي يتوقعها النموذج.(Wright and Stone, 1979).

بعد استبعاد الأفراد غير المطابقين للنموذج، تمت إعادة التحليل للكشف عن المفردات غير المطابقة للنموذج، حيث تم تقدير معلم الصعوبة لكل مفردة، والخطأ المعياري في قياس معلم الصعوبة، وقيم إحصائي المطابقة الداخلية للمفردات (Infit ZSTD) ، والذي يعبر عنه بإحصائي متوسطات المربعات للمفردات (Infit MNSQ). كما تم تقدير قيم إحصائي المطابقة الخارجية للمفردات (Outfit ZSTD)، والذي يعبر عنه

بإحصائي متوسطات المربعات للمفردات (Outfit MNSQ)، لكل معلم من معالم الصعوبة، والجدول رقم (33) يوضح ملخص نتائج التحليل لمعالم صعوبة مفردات اختبار مستويات التفكير الهندسي.

الجدول (33)

ملخص نتائج التحليل لمعالم الصعوبة لمفردات اختبار مستويات التفكير الهندسي .

الإحصاءات	الدرجة الخام Raw Score	الصعوبة Measure	الخطأ المعياري للنموذج Model S.E	إحصائي المطابقة الداخلية Infit		إحصائي المطابقة الخارجية Outfit	
				متوسط المربعات MNSQ	قيمة الإحصائي ZSTD	متوسط المربعات MNSQ	قيمة الإحصائي ZSTD
المتوسط الحسابي	298.6	0.00	0.10	1.00	0.2	0.99	0.1
الإنحراف المعياري	104.2	0.90	0.01	0.06	1.6	0.09	1.6
أعلى علامة	485	1.62	0.12	1.12	3.4	1.22	3.8
أدنى علامة	123	1.72–	0.09	0.86	2.7–	0.79	2.8–

يتضح من الجدول رقم (33)، أن المتوسط الحسابي لمتوسطات المربعات الداخلية والخارجية (MNSQ) يساوي الواحد بالنسبة لمتوسطات المربعات الداخلية، وقريب جداً من الواحد بالنسبة متوسطات المربعات الخارجية 0.99، بينما المتوسط الحسابي لمتوسط إحصائي المطابقة الداخلية والخارجية (ZSTD) بعيدة نوعاً ما عن الصفر، فالوضع المثالي الذي يفترضه النموذج هو (0، 1)، كما أن الإنحراف المعياري قريب من الواحد. وفيما يتعلق بالمفردات، تبين عدم مطابقة ستة (6) مفردات لنموذج راش، حيث زادت قيم متوسطات المربعات الموزونة لها الواحد الصحيح، والجدول رقم (34) يوضح قيم احصائيات المطابقة الداخلية والخارجية ومتوسطات المربعات الموزونة.

الجدول (34)

يوضح المفردات غير المطابقة، وقيم إحصائيات المطابقة الداخلية والخارجية، ومتوسطات المربعات الموزونة لهذه المفردات في التدريج الأولى لاختبار مستويات التفكير الهندسي.

إحصائي المطابقة الخارجية Outfit		إحصائي المطابقة الداخلية Infit		الخطأ المعياري للنموذج Model S.E	الصعوبة Measure	إرتباط بونت بيسريال PTBIS-CORR	الدرجة الخام Raw Score	المفردة Item
قيمة الإحصائي ZSTD	متوسط المربعات MNSQ	قيمة الإحصائي ZSTD	متوسط المربعات MNSQ					
3.0-	.77	2.5-	.85	.11	1.64	.50	120	20
0.5-	.96	.0	1.00	.10	1.30	.34	146	27
3.4	1.23	2.5	1.12	.10	1.16	.18	156	28
.1	1.00	0.1-	1.00	.10	.99	.34	172	10
1.5-	.92	2.0-	.92	.10	.93	.43	186	15
1.6	1.09	1.6	1.07	.10	.91	.26	173	8
1.1	1.04	1.0	1.03	.09	.54	.30	226	25
.2	1.01	0.2-	.99	.09	.54	.34	229	22
.3	1.01	.2	1.01	.09	.46	.33	241	14
.7	1.03	.5	1.02	.09	.41	.32	247	16
1.2	1.04	1.9	1.06	.09	.40	.28	246	26
0.9-	.97	1.1-	.97	.09	.24	.38	266	17
2.6	1.09	3.4	1.09	.09	.20	.23	278	4
4.3	1.17	3.2	1.09	.09	.18	.22	273	21
1.9	1.07	2.7	1.07	.09	.07	.25	294	2
0.8-	.97	0.7-	.98	.09	.05	.36	294	9
0.4-	.98	0.6	1.02	.09	.01	.32	295	1
0.1-	.99	0.3-	.99	.09	0.05-	.34	303	23
1.3	1.05	0.9	1.02	.09	0.08-	.30	310	24
2.2-	.89	2.4-	.92	.09	0.44-	.41	354	13
0.5-	.97	1.0-	.97	.09	0.48-	.35	361	6
.4	1.02	.5	1.02	.09	0.52-	.29	355	18
0.5-	.96	0.4-	.98	.10	0.95-	.31	404	19
1.5-	.87	0.9-	.95	.11	1.27-	.34	441	11
1.3-	.88	0.9-	.95	.11	1.32-	.33	442	5

1.4–	.85	0.9–	.94	.12	1.55–	.34	465	12
0.7–	.91	0.3–	.98	.12	1.68–	.27	475	7
1.0–	89	0.6–	.95	.12	1.70–	.30	475	3
0.1	0.99	2.0	1.00	0.10	0.00		293.8	Mean
1.6	0.10	1.5	0.06	0.01	0.90		102.2	S.D

بعد استبعاد الأفراد غير المطابقين، والمفردات غير المطابقة، ولغرض التأكد من مطابقة البيانات الناتجة عن استجابات أفراد العينة للاختبار والتحقق من موضوعية الاختبار بصورته النهائية 22 مفردة، والتحقق من افتراضات نموذج (فان هيل) لمستويات التفكير الهندسي، تمت إعادة التحليل للحصول على تقديرات متحررة من صعوبة المفردات وقدرات الأفراد، والجدول رقم (35) يبين ملخص نتائج التحليل للقيم المتحررة من قدرات الأفراد بعد استبعاد المفردات والأفراد غير المطابقين لنموذج راش.

الجدول (35)

ملخص نتائج التحليل للقيم المتحررة من قدرات الأفراد بعد استبعاد الأفراد والمفردات غير المطابقة.

الإحصاءات	الدرجة الخام Raw Score	القدرة Measure	الخطأ المعياري للنموذج Model S.E	إحصائي المطابقة الداخلية Infit		إحصائي المطابقة الخارجية Outfit	
				متوسط المربعات MNSQ	قيمة الإحصائي ZSTD	متوسط المربعات MNSQ	قيمة الإحصائي ZSTD
المتوسط الحسابي	11.9	0.30	0.50	1.00	0.0	0.99	0.0
الإنحراف المعياري	3.4	0.84	0.06	0.17	0.9	0.26	0.9
أعلى علامة	21	3.38	1.04	1.56	2.8	1.89	2.6
أدنى علامة	3	2.10–	0.47	0.57	2.7–	0.53	2.6–

يتضح لنا من الجدول رقم (35)، أن التقديرات النهائية المتحررة من قدرات الأفراد تراوحت ما بين الدرجتين 3– 21، وبمتوسط توزيع القدرة بلغ 0.30 لوجيت، وبانحراف معياري بلغ 0.84 لوجيت، أما الخطأ المعياري للمتوسط الحسابي لتقديرات القدرة فقد بلغ 0.50، وهي قيمة تقترب مما يفترضه النموذج. كما يتضح كذلك أن متوسط توزيع القدرة كانت قيمته قبل حذف المفردات الستة من الاختبار، القيمة 0.17 لوجيت، وبانحراف معياري 0.77، أما الخطأ المعياري للمتوسط الحسابي لتقديرات القدرة بلغ 0.44 لما كان متوسط صعوبة المفردات يساوي صفر، كما هو مبين في الجدولين رقم (33) ورقم (34) وبالتالي فإن

متوسط قدرة الأفراد إرتفع من القيمة 0.17 قبل الحذف، إلى القيمة 0.30 بعد حذف المفردات والأفراد غير المطابقين لنموذج راش، وبالتالي فهذه القيمة أعلى من صعوبة المفردات، أي أن مستوى الاختبار بشكل عام في متناول مستوى أفراد العينة مما يقلل من عوامل التخمين كما تؤكد ذلك (كاظم، 1988ب). مما يشير إلى دقة تحديد مواقع الأفراد على متصل القدرة، وهو ما يوضحه الجدول رقم(36):

الجدول (36)

يوضح ملخص نتائج التحليل المتحرر من صعوبة المفردات بعد التدريج.

الإحصاءات	الدرجة الخام Raw Score	الصعوبة Measure	الخطأ المعياري للنموذج Model S.E	إحصائي المطابقة الداخلية Infit		إحصائي المطابقة الخارجية Outfit	
				متوسط المربعات MNSQ	قيمة الإحصائي ZSTD	متوسط المربعات MNSQ	قيمة الإحصائي ZSTD
المتوسط الحسابي	306.9	0.00	0.10	1.00	0.1	0.99	0.1
الإنحراف المعياري	103.4	0.93	0.01	0.04	1.1	0.08	1.2
أعلى علامة	475	1.46	0.12	1.09	1.2	1.13	2.4
أدنى علامة	146	1.61–	0.09	0.90	2.6–	0.84	2.0–

يشير الجدول رقم (36)، إلى أن متوسط توزيع القدرة بلغ (صفراً) لوجيت، والانحراف المعياري بلغ 0.93 لوجيت، وقيم التقديرات المتحررة من صعوبة المفردات تراوحت ما بين 1.61– و1.46، أما الخطأ المعياري للمتوسط الحسابي لتقديرات الصعوبة فقد بلغ 0.10 وهي قيمة متدنية نوعا ما، مما يشير إلى دقة تقديرات الصعوبة، فالمتوسط الحسابي لمتوسطات المربعات الداخلية والخارجية (MNSQ) يساوي الواحد بالنسبة لمتوسطات المربعات الداخلية، وقريب جداً من الواحد بالنسبة متوسطات المربعات الخارجية 0.99، بينما المتوسط الحسابي لمتوسط إحصائي المطابقة الداخلية والخارجية (ZSTD) تقريباً يساوي الصفر، فالوضع المثالي الذي يفترضه النموذج هو (0، 1)، كما أن الإنحراف المعياري قريب من الواحد. وقد تم تقدير قيم معالم مفردات اختبار مستويات التفكير الهندسي بصورته النهائية 22 مفردة، باستخدام طريقة الأرجحية العظمى غير المشروطة (UCON)، لتقدير أدق القدرات والصعوبة، ولتقليل الخطأ في التقدير لصعوبة المفردات. والجدول رقم (37) يبين صعوبة المفردات والخطأ المعياري للقياس في تقدير الصعوبة مرتبة تنازلياً بحسب صعوبتها بعد التدريج النهائي.

الجدول (37)

يوضح صعوبة المفردات وأخطاؤها المعيارية مرتبة تنازلياً وفق صعوبة المفردات.

إحصائي المطابقة الخارجية Outfit		إحصائي المطابقة الداخلية Infit		الخطأ المعياري للنموذج Model S.E	الصعوبة Measure	ارتباط بونت بيسريال PTBIS-CORR (r_{bis})	الدرجة الخام Raw Score	المفردة Item
قيمة الإحصائي ZSTD	متوسط المربعات MNSQ	قيمة الإحصائي ZSTD	متوسط المربعات MNSQ					
0.0	1.00	0.5	1.03	.10	1.46	.34	146	27
0.8	1.04	0.9	1.04	.10	1.15	.33	172	10
2.0–	.89	2.6–	.90	.10	1.08	.47	186	15
2.1	1.13	2.0	1.09	.10	1.06	.27	173	8
2.2	1.10	1.4	1.05	.09	0.69	.31	226	25
0.1	1.01	00.0	1.00	.09	0.68	.37	229	22
0.3	1.01	0.2	1.00	.09	0.60	.36	241	14
0.4	1.02	0.4	1.01	.09	0.54	.35	247	16
1.6	1.07	2.2	1.07	.09	0.54	.29	246	26
0.8–	.97	1.0–	.97	.09	0.37	.39	266	17
1.0–	.96	0.7–	.98	.09	0.18	.38	294	9
0.5	1.02	1.4	1.04	.09	0.15	.32	295	1
0.3	1.01	0.3–	.99	.09	0.07	.36	303	23
2.4	1.10	1.7	1.05	.09	0.05	.29	310	24
0.4–	.98	0.5–	.98	.09	0.36–	.36	361	6
0.2	1.01	.1	1.00	.10	0.40–	.33	355	18
0.3–	.97	0.3–	.99	.10	0.84–	.32	404	19
1.6–	.86	1.0–	.94	.11	1.16–	.36	441	11
1.4–	.87	0.9–	.95	.11	1.21–	.35	442	5
1.5–	.84	0.8–	.95	.12	1.45–	.34	465	12
0.5–	.94	0.1–	.99	.12	1.58–	.27	475	7
0.1–	.98	0.4–	.97	.12	1.61–	.28	475	3
0.1	0.99	0.1	0.1	0.10	0.00		306.9	Mean
1.2	0.08	1.1	0.04	0.01	0.93		103.4	S.D

يتضح لنا من الجدول رقم (37)، أن مؤشرات المطابقة لكل مفردة من مفردات اختبار مستويات التفكير الهندسي المطور كانت ضمن حدود المطابقة الداخلية (التقاربية) والخارجية (التباعدية)، والشكل رقم (27) يوضح قيم كل من صعوبة المفردات ومجال قيم الإحصائي لمتوسط المربعات التباعدي للمفردات (Outfit MNSQ)، وقيم الإحصائي متوسط المربعات التقاربي للمفردات (Infit MNSQ).

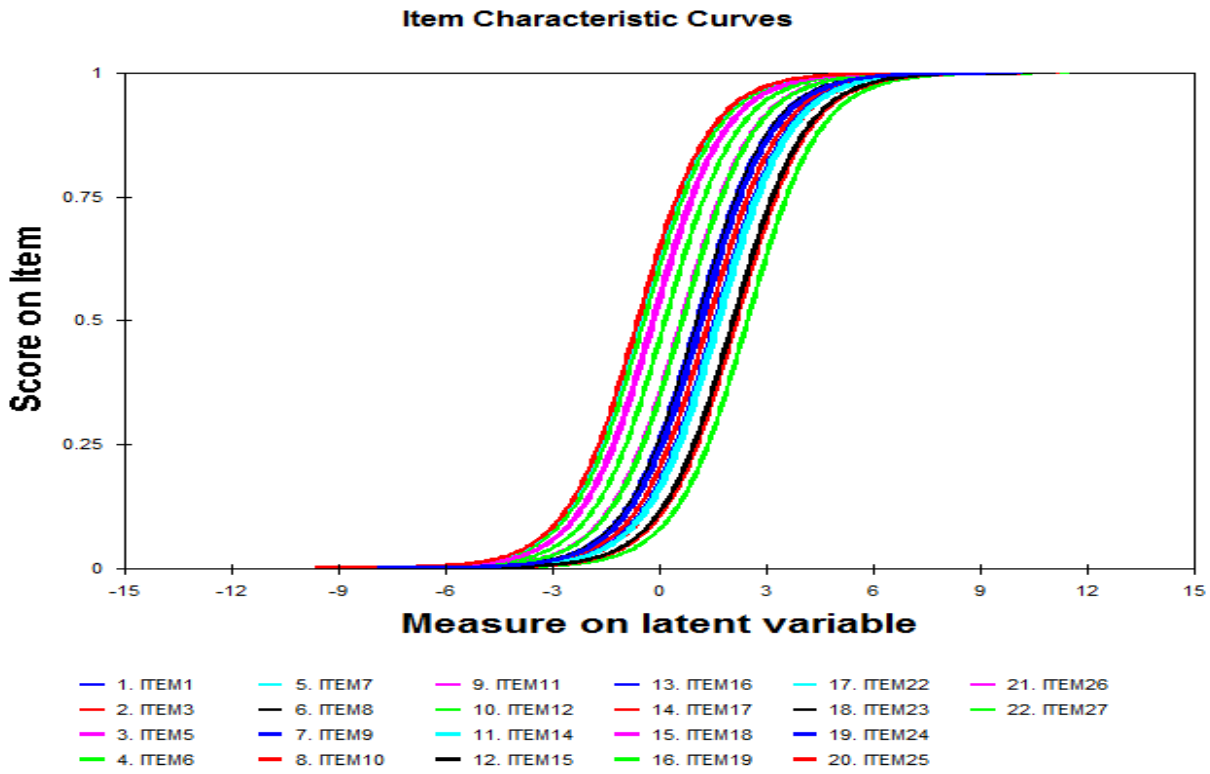
ITEM FIT GRAPH: MISFIT ORDER

ENTRY NUMBER	MEASURE - +	INFIT MEAN-SQUARE 0.0 1 2	OUTFIT MEAN-SQUARE 0.0 1 2	ITEM
6	*	: *	: .*	ITEM8
19	*	: *	: .*	ITEM24
20	*	: *	: *	ITEM25
21	*	: *	: *	ITEM26
1	*	: *	: *	ITEM1
8	*	: *	: *	ITEM10
22	*	: *	: *.	ITEM27
13	*	: *	: *	ITEM16
11	*	: *	: *	ITEM14
15	*	: *	: *	ITEM18
17	*	: *.	: *	ITEM22
18	*	: *.	: *	ITEM23
5	*	: *.	: *.	ITEM7
16	*	: *.	: *.	ITEM19
2	*	: *.	: *.	ITEM3
4	*	: *.	: *.	ITEM6
7	*	: *.	: *.	ITEM9
14	*	: *.	: *.	ITEM17
3	*	: *.	: * .	ITEM5
10	*	: *.	: * .	ITEM12
9	*	: *.	: * .	ITEM11
12	*	: * .	: * .	ITEM15

الشكل (27) يوضح مؤشرات المطابقة للمفردات وحدود المطابقة ومواقع للمفردات من هذه الحدود وفقاً لإحصائي متوسط المربعات التباعدي والتقاربي (MNSQ).

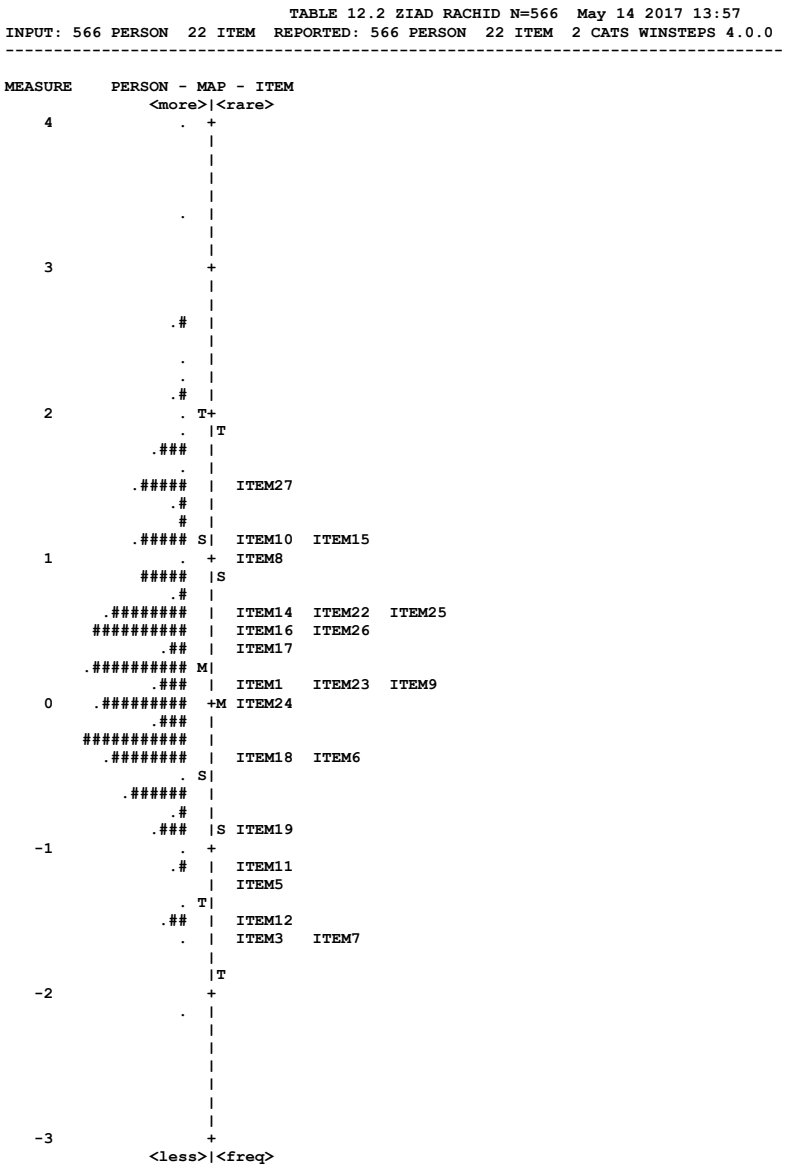
يتضح من الشكل رقم (27)، أن مؤشرات المطابقة للمفردات وحدود المطابقة ومواقع المفردات من هذه الحدود وفقاً لإحصائي متوسط المربعات التباعدي والتقاربي (MNSQ). تشير إلى أن صعوبة مفردات اختبار مستويات التفكير الهندسي في صورته النهائية 22 مفردة، والمعبر عنها في العمود (Measure) بأنها في المجال المقبول للصعوبة، كما تشير قيم مؤشرات الملاءمة الإحصائية لمتوسط المربعات التباعدية والتقاربية لمفردات الاختبار إلى أنها قريبة من الوضع المثالي الذي يفترضه النموذج وهو القيمة (1)، والذي تراوحت قيمه ما بين 0.84 و 1.13 بمتوسط حسابي 1 وإنحراف معياري 0.04 بالنسبة لمؤشرات الملاءمة الإحصائية لمتوسط المربعات التقاربية (الداخلية)، وبمتوسط حسابي 0.99 وإنحراف معياري 0.08 بالنسبة

لمؤشرات الملاءمة الإحصائية لمتوسط المربعات التباعدية (الخارجية). وهي قيم في مجال القياس المنتج الذي حدده كلٍ من (Wright and Linacre, 1994) ما بين 0.70 و1.30 أو ما بين (0.50 و1.5) عموماً والقريبة من الواحد، وهو مؤشر إحصائي أكثر حساسية للسلوكيات غير المتوقعة من النموذج (المتطرفة مثل التخمين و الأخطاء غير المتوقعة) التي تؤثر في الاستجابات عن المفردات التي تكون بعيدة عن مستوى قدرة الفرد. لذلك فالقيمة الأكبر من 1.0 لمؤشر (Outfit-Underfit) أو مؤشر (MNSQ <0.7 /-2) تعني وجود تباين كبير أو تداخل في البيانات غير مبرر. والقيمة أقل من 1.0 لمؤشر (Infit-Overfit) أو مؤشر (MNSQ >1.3 /2) يعني أن النموذج مفرط في البيانات وقد يتسبب في تضخم إحصاءات الثبات (William j et al, 2014, p.166) . ويدعم هذه النتائج الشكل رقم (28) الذي يظهر منحنيات خصائص مفردات اختبار مستويات التفكير الهندسي بعد التدريج النهائي.



الشكل (28) منحنى خصائص مفردات اختبار مستويات التفكير الهندسي بعد التدريج باستخدام نموذج راش.

يلاحظ من الشكل رقم (28)، بأن المنحنيات المميزة لمفردات اختبار مستويات التفكير الهندسي بعد تدريجه باستخدام نموذج (راش) أحادي البارامتر لا تتقاطع، ولكن تختلف عن بعضها في الصعوبة فقط وتتلاقى فقط عند المحور الأفقي الذي يمثل متصل القدرة أو السمة. وهو ما تؤكد عليه خريطة فرد– بند لرايت في الشكل رقم (29):



الشكل (29) يوضح خريطة رايت (Wright) لتدريج الأفراد ومفردات اختبار مستويات التفكير الهندسي.

تشير خريطة رايت في الشكل رقم (29) وبصورة جلية إلى وجود تناسب تام بين توزيع تقدير مفردات اختبار مستويات التفكير الهندسي من الجانب الأيسر للخريطة، وتدريج المفردات من الجانب الأيمن. وبالرغم من إنتظام التدريج لصعوبة المفردات بصورة واضحة من المدى −1.61 و1.46 لوجيت، فقد تبدو بعض الفراغات الضيقة عند بعض المستويات الأخرى من المتغير والتي تظهرها خريطة (رايت) والتي يمكن التحقق منها من خلال العلاقة بين مسافة الفراغ بين صعوبة المفردتين ومجموع الخطأ المعياري لهما، والتي يجب أن يكون الفرق بين صعوبتي المفردتين المحددتين للفراغ أقل من مجموع الخطأ المعياري للقياس لكلاهما والفراغات، هي:

الفراغ الأول: الفرق بين صعوبتي المفردتين 18 و6 يساوي 0.04 ومجموع الخطأ المعياري للقياس 0.19، أي أن الخطأ المعياري للقياس يغطي الفراغ.

الفراغ الثاني: الفرق بين صعوبتي المفردتين 24 و6 يساوي 0.31 ومجموع الخطأ المعياري للقياس 0.18، أي أن الخطأ المعياري للقياس لا يغطي الفراغ.

الفراغ الثالث: الفرق بين صعوبتي المفردتين 8 و25 يساوي 0.37 ومجموع الخطأ المعياري للقياس 0.19، أي أن الخطأ المعياري للقياس لا يغطي الفراغ.

الفراغ الرابع: الفرق بين صعوبتي المفردتين 27 و10 يساوي 0.31 ومجموع الخطأ المعياري للقياس 0.20، أي أن الخطأ المعياري للقياس لا يغطي الفراغ.

وهكذا فان الخطأ المعياري للقياس لصعوبة كل مفردتين مكونتين للفراغ، يغطي مسافة الفراغ الأول بينهما، ولكنه لا يغطي مسافة الفراغات الثلاثة المتبقية و بفرق ضئيل، وبالرغم من ذلك يمكن اعتبار أن هناك انتظام في تدرج المفردات على مدى ميزان القياس، بما يعني تعريف مفردات الاختبار للمتغير الذي يمثل بمدى هذا المتصل وهو القدرة على التفكير الهندسي. مما يدل على جودة تدريج مفردات الاختبار.

وللتحقق من القضية الثانية من قضايا الصدق البنائي، يجب الإجابة على السؤال التالي وهو: هل تظهر مستويات التفكير الهندسي وفق نظرية (فان هيل) ترتيب هرمي متسلسل أم أنها منفصلة؟ حيث أظهرت خريطة (رايت) بعد التدريج النهائي لاختبار مستويات التفكير الهندسي، بأن المفردات المصممة لاختبار كل مستوى من مستويات التفكير الهندسي الأربعة لـ(فان هيل) تميل إلى أن تكون متسلسلة، فالزيادة في متوسط عتبة صعوبة المفردات في كل مستوى، تدعم الأدلة البحثية الأخرى أن المستويات تشكل التسلسل الهرمي. على الرغم من وجود بعض التداخل في عتبات الصعوبة بالنسبة لبعض المفردات والتي تشير إلى أن المستويات ليست منفصلة تماماً.

1-4- عرض وتحليل نتائج الفرضية الرابعة:

والتي تنص على أن:" تتباين تقديرات قدرات الأفراد ودالة المعلومات المقابلة لكل درجة كلية خام محتملة على مفردات اختبار مستويات التفكير الهندسي باستخدام نموذج راش".

ولاختبار صحة هذا الفرض، تم تقدير كمية المعلومات التي يقدمها اختبار مستويات التفكير الهندسي عند مستويات القدرة المختلفة، والعلاقة بين الدرجة الخام والقدرة بصورته النهائية، باستخدام نموذج راش

أحادي البارامتر، وتقدير دالة المعلومات المقابلة لكل درجة خام وكل قدرة، وتقدير الخطأ المعياري والقدرة، والجدول رقم (38) يوضح هذه المؤشرات:

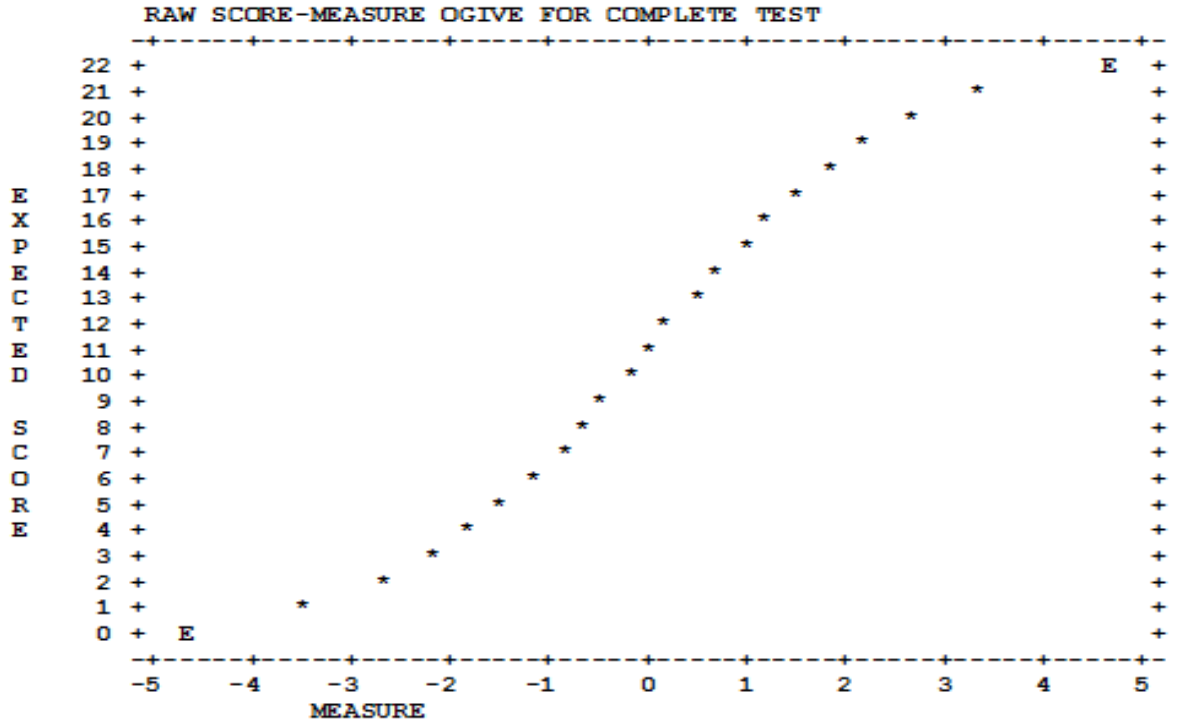
الجدول (38)

يتضمن تقديرات قدرة الفرد المقابلة لكل درجة كلية محتملة على الاختبار، وإنحرافاتها المعيارية، وكذا كمية المعلومات التي يقدمها اختبار مستويات التفكير الهندسي عند كل مستويات القدرة المختلفة، والعلاقة بين الدرجة الخام والقدرة بوحدة اللوجيت.

الدرجة الخام	تكرار الأفراد	التكرار التراكمي للأفراد	القيم المتحررة لقدرات الأفراد بوحدة اللوجيت	الخطأ المعياري	دالة المعلومات
1	0	0	3.44-	1.04	0.92
2	0	0	2.65-	0.77	1.68
3	1	1	2.15-	0.66	2.32
4	1	2	1.77-	0.59	2.86
5	13	15	1.44-	0.55	3.3
6	10	25	1.16-	0.52	3.67
7	19	44	0.89-	0.5	3.98
8	38	82	0.65-	0.49	4.22
9	45	127	0.42-	0.48	4.4
10	70	197	0.2-	0.47	4.52
11	**60**	**257**	**0.02**	**0.47**	**4.58**
12	64	321	0.24	0.47	4.58
13	57	378	0.46	0.47	4.51
14	47	425	0.69	0.48	4.37
15	30	455	0.92	0.49	4.15
16	32	487	1.17	0.51	3.85
17	37	524	1.44	0.54	3.47
18	21	545	1.75	0.58	2.99
19	10	555	2.12	0.64	2.42
20	7	562	2.61	0.76	1.73
21	4	566	3.38	1.04	0.93

يوضح الجدول رقم (38) تقديرات قدرات الأفراد، وكذا الأخطاء المعيارية المقابلة لكل تقدير للقدرة المقابل للدرجات الكلية التي حصل عليها أفراد عينة التدريج، وكذلك المحتملة على الاختبار، حيث امتدى مدى الدرجات المحتملة(الحقيقية) من الدرجة الكلية واحد، إلى الدرجة الكلية 21، حيث يحذف من التدريج

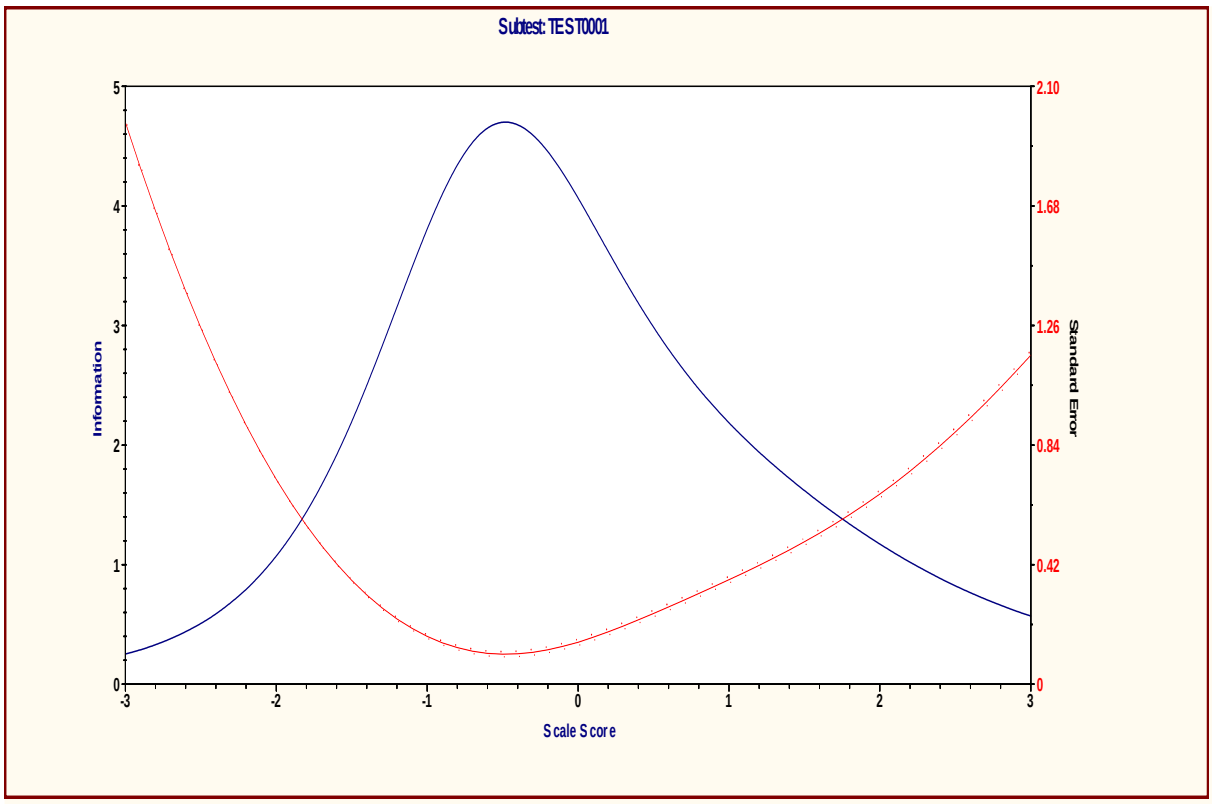
التلامذة الحاصلون على الدرجة (صفر) والدرجة النهائية 22، وتراوحت تقديرات قدرات التلامذة من –3.44 إلى 3.38 لوجيت. و يمكننا إيجاد الدرجة الحقيقية المقابلة لأي مستوى محدد من القدرة من خلال المنحنى المميز للاختبار، والذي يعتمد شكله على عدد من العوامل منها، عدد مفردات الاختبار، ونموذج المنحنى المميز للمفردات المستخدمة، وقيم معالم المفردات، فالمنحنى المميز للاختبار هو العلاقة الوظيفية (التقييسية) بين الدرجة الحقيقية ومقياس القدرة، كما يوضحه الشكل رقم (30) للعلاقة المتجمعة بين الدرجة الكلية الخام المحتملة على الاختبار، والقدرة المقابلة لها باللوجيت.



الشكل (30) العلاقة التقييسية بين الدرجة الكلية المحتملة على الاختبارعند كل مستوى من مستويات القدرة المختلفة للأفراد.

يوضح الشكل رقم (30) رسم بياني للدرجات الحقيقية المقابلة لمستويات القدرة، كدالة للقدرة على المحور الأفقي، في حين يمثل المحور العمودي الدرجات الحقيقية المتوقعة، و التي تراوحت ما بين 1 و21 وهو عدد المفردات المكونة للاختبار وهو مقياس القدرة، بعد حذف الدرجة التامة والصفرية، كما يتضح أن العلاقة بين الدرجة الكلية الخام والمحتملة على الاختبار، والقدرة المقابلة لها علاقة طردية موجبة في شكل دالة منتظمة متزايدة، حيث تزيد القدرة كلما ازدادت الدرجة الكلية الخام المقابلة لها. فالدور الأساسي للمنحنى المميز للاختبار في نظرية الاستجابة للمفردة، هو إمدادنا بوسائل لتحويل درجات القدرة إلى درجات حقيقية. كما يتضح من الجدول رقم (38)، أن قيم دالة معلومات اختبار مستويات التفكير الهندسي تراوحت ما بين

0.92 و0.47 في النصف الأول من مدى الاختبار، وبعد ذلك بدأت في الإنحدار تدريجياً، وهذا يدل على أن الدالة تعبر عن الوضع الحقيقي والواقعي لمتطلبات أفراد العينة، كما تتفق تماماً مع القدرات المتوقعة لأفراد العينة، مما يؤكد دقة الاختبار في عملية القياس والثقة في نتائجه، وقدرته على التمييز، وتحديد القدرات المختلفة لأفراد العينة. فمتوسط القدرة يساوي تقريباً صفراً، أي أنها عند أقصى دالة للمعلومات، فالاختبار يعطي أقصى دالة للمعلومات عند مستوى القدرة المتوسط حيث بلغت قيمتها 4.58 وخطأ معياري للقياس 0.47، المقابلة لمنتصف الدرجة الحقيقية 11 والتي تقابل صفر، أي صفر التدريج، أو القيمة المتوسطة لمعالم صعوبة المفردات وهو مؤشر على المكان الذي سيقوم الاختبار فيه بوظيفته على مقياس القدرة، كما يتضح من الشكل رقم (31):

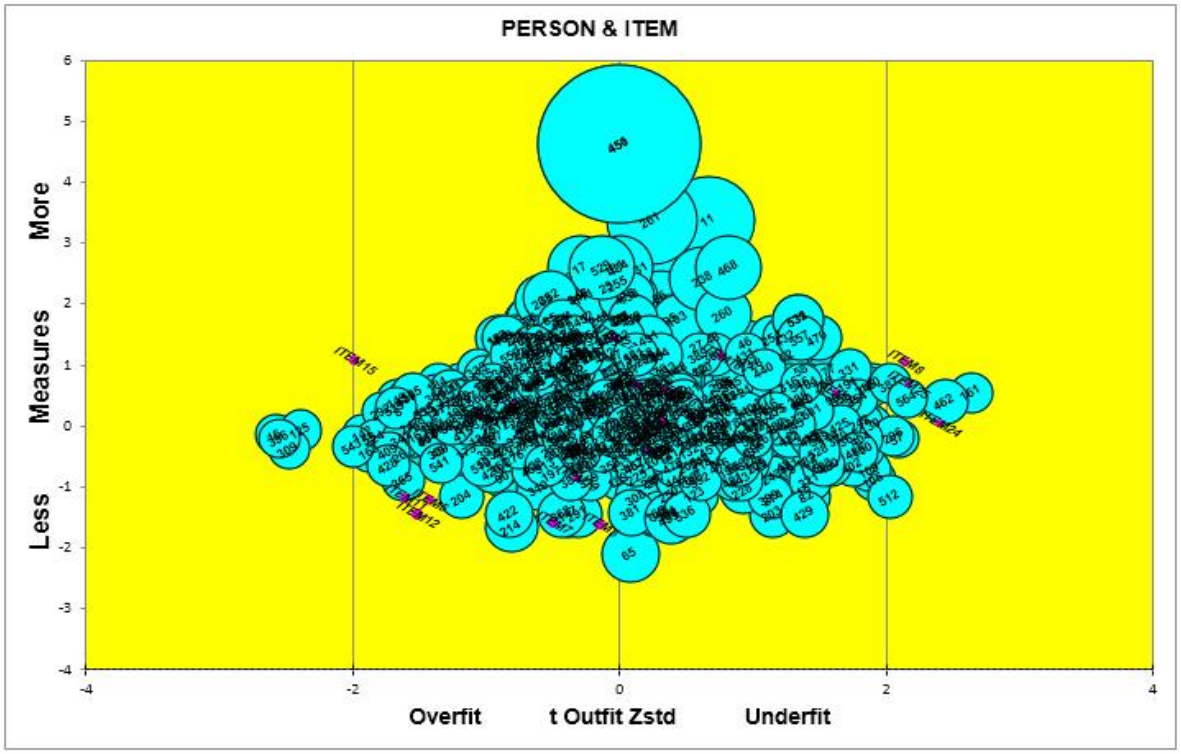


الشكل (31) يوضح علاقة دالة معلومات الاختبار مع الخطأ المعياري للقدرة الكامنة.

يتضح من الشكل رقم (31)، أن منحنى دالة معلومات الاختبار متماثلاً بطريقة مقبولة، وذا إرتفاع واضح جداً في الوسط، وبالتالي فهو يأخذ الشكل الاعتدالي المتوقع لمثل هذه المنحنيات، ويشير شكل المنحنى إلى أنه يتم تقدير القدرة بأعظم دقة بمصاحبة متوسط مقياس القدرة، وبما أن القدرة متغير متصل تكون المعلومات متغيراً مستمراً أيضاً، لهذا تحدث ذروة دالة معلومات الاختبار في نقطة أعلى بدرجة طفيفة

من النقطة المتوسط لمقياس القدرة، حيث أن المحور الصادي (y) يمثل المعلومات لـ"فيشر ،fisher"، والمحور السيني (x) يمثل طول المتغير الكامن باللوغريتم، كما يلاحظ إنخفاض في الخطأ المعياري (الممثل بالخط الأحمر المتقطع) أين كانت أقصى قيم للمعلومات عن مستوى قدرات التلامذة التي يمدنا بها اختبار مستويات التفكير الهندسي، بدقة مرتفعة إلى حد ما داخل مدى القدرة بين 1.17 و−1.16 لوجيت وهو ما كان متوقعاً، لإن معظم مفردات الاختبار غطت هذا المدى، وخارج هذا المدى ينخفض مقدار المعلومات بسرعة ولا يتم تقدير مستويات القدرة بدرجة جيدة، مما يعني أن الخطأ المعياري في التقدير يكون أقل ما يمكن عند مستويات القدرة التي تناظر أقصى المعلومات، فأقل قيمة للخطأ المعياري، كما نلاحظ في الجدول أعلاه يساوي 0.47 والذي يناظر أقصى قيمة للمعلومات التي يمنحها الاختبار، وهي قيمة للقدرة تساوي 4.58 لوجيت عند الدرجة الخام 11 ومستوى القدرة 0.02. وتنخفض قيمة المعلومات التي يمكن الحصول عليها من الاختبار بشكل واضح مع زيادة القدرة عن 1.17 أو تنخفض عن مستوى القدرة −1.16، وهو ما يؤكد حاجة الاختبار إلى إضافة مفردات تغطي تلك المستويات من القدرة، فمقدار المعلومات التي تسهم بها مجموعة من المفردات عند مستوى قدرة معين يتناسب تناسباً عكسياً مع مربع سعة فترة الثقة حول تقديرات القدرة عند مستوى القدرة (θ) معين، أي أن الخطأ المعياري للقياس يساوي $SE(\theta)=\frac{1}{\sqrt{I(\theta)}}$ فبالتعويض في هذه المعادلة بالقيم، فإن: $SE(\theta)=\frac{1}{\sqrt{4.58}}=\frac{1}{2.14}=0.467$، وبتقريب القيمة للخطأ المعياري للقياس فإنها تساوي القيمة 0.47 والتي تناظر أقصى قيمة لدالة معلومات اختبار مستويات التفكير الهندسي.

وبالعودة إلى الجدولين رقم (38 و37)، يتضح أن قيم الأخطاء المعيارية للقياس بالنسبة لصعوبة مفردات اختبار مستويات التفكير الهندسي، تراوحت ما بين 0.09 و 0.12 لوجيت، بمتوسط حسابي قدره 0.10، وانحراف معياري 0.01 وهي قيم صغيرة جداً وقريبة من الصفر. في حين أن قيم الأخطاء المعيارية للقياس بالنسبة لتقديرات قدرات الأفراد المقابلة لكل درجة خام محتملة على اختبار مستويات التفكير الهندسي باستخدام نموذج راش، فتراوحت ما بين 1.04 و 1.04 لوجيت، بمتوسط حسابي 0.60 وانحراف معياري 0.17. ويقدم الشكل رقم (32) ملخص تصوري يقدم معلومات حول حجم الخطأ المعياري للقياس لكل من المفردات والأفراد من تطوير بوند وفوكس و ليناكر (Bond and Fox 2001, and Linacre) وهو بمثابة الدقة في القياس.



الشكل (32) يوضح خريطة المسار(The Pathway Map) أو المخطط الفقاعي، للأخطاء المعيارية للقياس لكل من صعوبة مفردات اختبار مستويات التفكير الهندسي، وقدرات الأفراد.

يظهر في المحور الأفقي للشكل رقم (32) المؤشر الاحصائي للملائمة التباعدي (الخارجي) المعياري (The Standardized Information Weighted Fit Statistics t Outfit Zstd)، والذي يشير إلى أن معظم مفردات الاختبار والأفراد تقريباً تتجمع بالقرب من الوسط، والتي تشير إلى أن المفردات تقيس بشكل أكثر دقة القدرة على التفكير الهندسي لدى التلامذة، والتي وقعت في المجال مابين –0.2 و 0.2 تقريباً، كما يمثل المحور العمودي قيم صعوبة المفردات وقدرات الأفراد والتي تراوحت ما بين –1.61 و 1.46 لوجيت بالنسبة لصعوبة المفردات، ومابين –3.44 و 3.38 لوجيت بالنسبة لقدرات الأفراد. كما يتضح من الشكل رقم (32)، تناسب حجم الفقاعات الممثلة لكل من المفردات والأفراد مع حجم الأخطاء المعيارية للقياس لكل منهما، حسب مدى تقارب البيانات الخاصة بكل فرد من النموذج على اختبار مستويات التفكير الهندسي، وتدل قيم الأخطاء المعيارية الصغيرة نسبياً على دقة وثبات تقديرات قدرات الأفراد على الاختبار.

1-5-عرض وتحليل نتائج الفرضية الخامسة:

والتي تنص على أن:" يحقق تدريج اختبار مستويات التفكير الهندسي باستخدام نموذج راش استقلالية القياس باختلاف عينة التدريج و الاختبار المستخدم".

وللتحقق من صحة هذه الفرضية، يجب التحقق من استقلالية القياس من ناحيتين:

أولاً. استقلال القياس عن قدرة العينة التي تؤدي الاختبار أي:

أ. لا يعتمد تقدير قدرة الفرد على قدرة باقي الأفراد الذين يجيبون على الاختبار.

ب. لا يعتمد تقدير صعوبة المفردة على قدرة الأفراد الذين يجيبون على الاختبار.

ولتحقيق هاتين النقطتين تطلب من الباحث وجود عينتين من الأفراد الملائمين، وذلك لتأدية نفس مجموعة المفردات المدرجة بواسطة نموذج (راش).

1. تجزئة عينة التحليل الكلية إلى عينتين إحداهما "مرتفعة المستوى"، والأخرى "منخفضة المستوى" بعد ترتيب أفراد عينة التحليل الكلية البالغة 566 فرداً، وتتكون العينة المرتفعة المستوى من الفرد الأول وحتى الفرد 283، والعينة المنخفضة المستوى من الفرد 284 وحتى الفرد 566.

2. باستخدام برنامج وينستبس حللت نتائج استجابات كل عينة على حده على مفردات الاختبار في صورته النهائية 22 مفردة، وذلك وفقاً لنموذج (راش).

3. حددت العلاقة التقيسية بين الدرجة الكلية المحتملة على الاختبار من الدرجة الكلية 1 إلى الدرجة الكلية 22، وتقديرات القدرة المشتقة من تحليل أداء كل من العينة المرتفعة المستوى، والعينة المنخفضة المستوى، وكذا أخطائهما المعيارية.

4. حددت العلاقة التقيسية بين مفردات الاختبار وتقديرات صعوبتهما المشتقة من كل عينة من العينتين، المرتفعة المستوى، والمنخفضة المستوى، وكذا أخطائهما المعيارية.

5. عملت المقارنة بين تقديرات القدرة المقابلة لكل درجة كلية محتملة على الاختبار، كما تشتق من تحليل أداء أفراد العينة الكلية، وكل من العينة المرتفعة المستوى، والعينة المنخفضة المستوى، وذلك للتحقق من تكافئهما إحصائياً.

6. عملت المقارنة بين التقديرات المتناظرة لصعوبة مفردات الاختبار، المشتقة من تحليل أداء أفراد العينة الكلية، وكل من العينة المرتفعة المستوى، والعينة المنخفضة المستوى، وذلك للتحقق من تكافئهما إحصائياً.

- **المقارنـة بين تقديرات القدرة المشتقة من كل عينة من العينات الثلاث:**

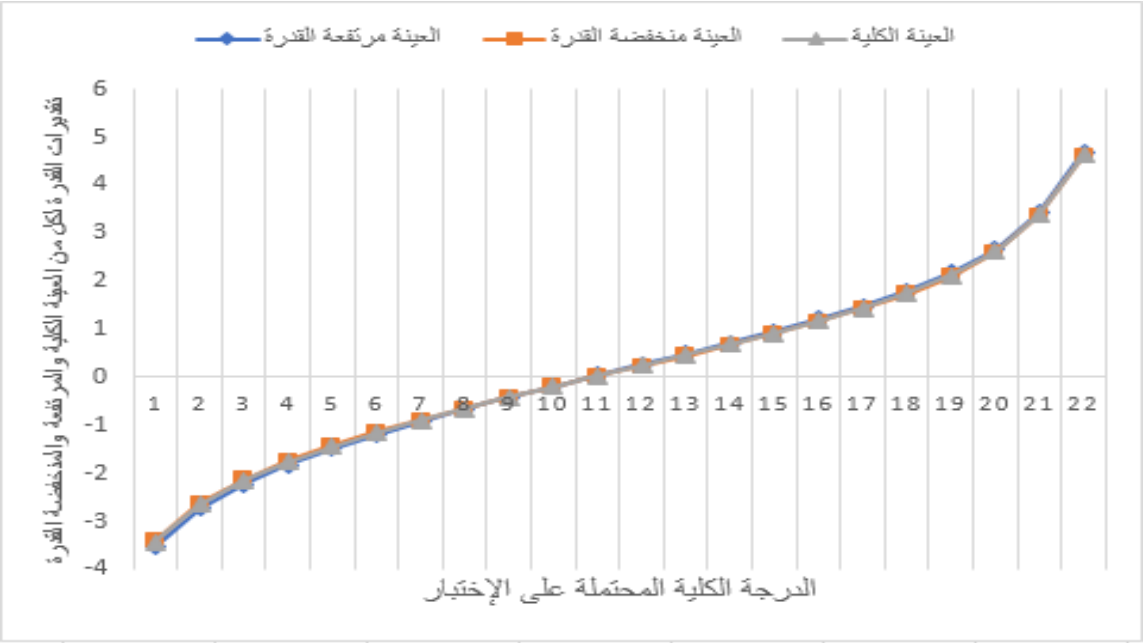
تم إجراء تقديرات القدرة المقابلة لكل درجة كلية محتملة على الاختبار والمشتقة من تحليل أداء العينة الكلية 566 فرداً، والعينة المرتفعة المستوى 283 فرداً، والعينة المنخفضة المستوى 283 فرداً، وكذا أخطائهما المعيارية، وهوما يوضحه الجدول رقم (39):

الجدول (39)

تقديرات القدرة (باللوجيت) المقابلة لكل درجة محتملة على الاختبار المشتقة من كل من العينة الكلية – العينة المرتفعة – العينة المنخفضة – وكذا الإخطاء المعيارية.

الدرجة الكلية	العينة الكلية		العينة منخفضة المستوى			العينة المرتفعة المستوى		
	تقدير القدرة	الخطأ المعياري	تقدير القدرة	الفرق بين التقديرين	الخطأ المعياري	تقدير القدرة	الفرق بين التقديرين	الخطأ المعياري
1	3.44–	1.04	3.4–	0.04–	1.04	3.53–	0.09	1.05
2	2.65–	0.77	2.62–	0.03–	0.77	2.73–	0.08	0.78
3	2.15–	0.66	2.13–	0.02–	0.65	2.22–	0.07	0.66
4	1.77–	0.59	1.75–	0.02–	0.59	1.82–	0.05	0.6
5	1.44–	0.55	1.43–	0.01–	0.55	1.48–	0.04	0.56
6	1.16–	0.52	1.14–	0.02–	0.52	1.19–	0.03	0.53
7	0.89–	0.5	0.89–	0	0.5	0.91–	0.02	0.51
8	0.65–	0.49	0.65–	0	0.48	0.66–	0.01	0.5
9	0.42–	0.48	0.42–	0	0.47	0.42–	0	0.49
10	0.2–	0.47	0.2–	0	0.47	0.19–	0.01–	0.48
11	0.02	0.47	0.02	0	0.46	0.04	0.02	0.47
12	0.24	0.47	0.23	0.01–	0.46	0.26	0.02	0.47
13	0.46	0.47	0.45	0.01–	0.47	0.49	0.03	0.48
14	0.69	0.48	0.67	0.02–	0.48	0.72	0.03	0.48
15	0.92	0.49	0.91	0.01–	0.49	0.95	0.03	0.5
16	1.17	0.51	1.16	0.01–	0.51	1.21	0.04	0.51
17	1.44	0.54	1.43	0.01–	0.54	1.48	0.04	0.54
18	1.75	0.58	1.74	0.01–	0.58	1.8	0.05	0.58
19	2.12	0.64	2.1	0.02–	0.64	2.17	0.05	0.65
20	2.61	0.76	2.59	0.02–	0.76	2.66	0.05	0.76
21	3.38	1.04	3.36	0.02–	1.04	3.43	0.05	1.04
22	4.63	1.85	4.61	0.02–	1.85	4.69	0.06	1.85

وباعتبار الخطأ المعياري لتقديرات القدرة الموضحة بالجدول رقم (39)، نلاحظ تكافؤ تلك التقديرات المتناظرة المشتقة من تحليل أداء أفراد العينة الكلية، وتلك المشتقة من كل من العينة المرتفعة المستوى، والعينة المنخفضة المستوى، وذلك لقدرة الأفراد الحاصلين على كل درجة كلية محتملة (حقيقية) على الاختبار، وهذا يعنى عدم تأثر تلك التقديرات باختلاف مستوى عينة التحليل. والشكل رقم (33) يوضح العلاقة التقيسية بين الدرجة الكلية المحتملة على الاختبار، وتقديرات القدرة (المنحنى المحدد للاختبار)، كما تشتق كل من العينة المرتفعة المستوى، والعينة المنخفضة المستوى، وأفراد العينة الكلية.



الشكل (33) يوضح منحنى العلاقة التقيسية بين الدرجة الكلية المحتملة على الاختبار وتقديرات القدرة المشتقة من كل من العينة المرتفعة والعينة المنخفضة.

يتضح من الشكل رقم (33)، تطابق العلاقة التقيسية كما تشتق من كل من العينتين، والعينة الكلية، بما يؤكد تحرر تقديرات القدرة من أداء العينة التي تجرى الاختبار. **وبهذا يتحقق الشق الأول من استقلالية القياس عن العينة المستخدمة، وهو تحرر تقديرات القدرة من أداء العينة التي تجري الاختبار.**

- **المقارنة بين تقديرات الصعوبة المشتقة من كل عينة من العينات الثلاث:**

قام الباحث بتحليل أداء وتقديرات الصعوبة لكل مفردة من مفردات الاختبار المشتقة لكل من أفراد العينة الكلية، والعينة المرتفعة المستوى، والعينة المنخفضة المستوى، وكذا أخطائهما المعيارية. وهو ما يوضحه

الجدول رقم (40):

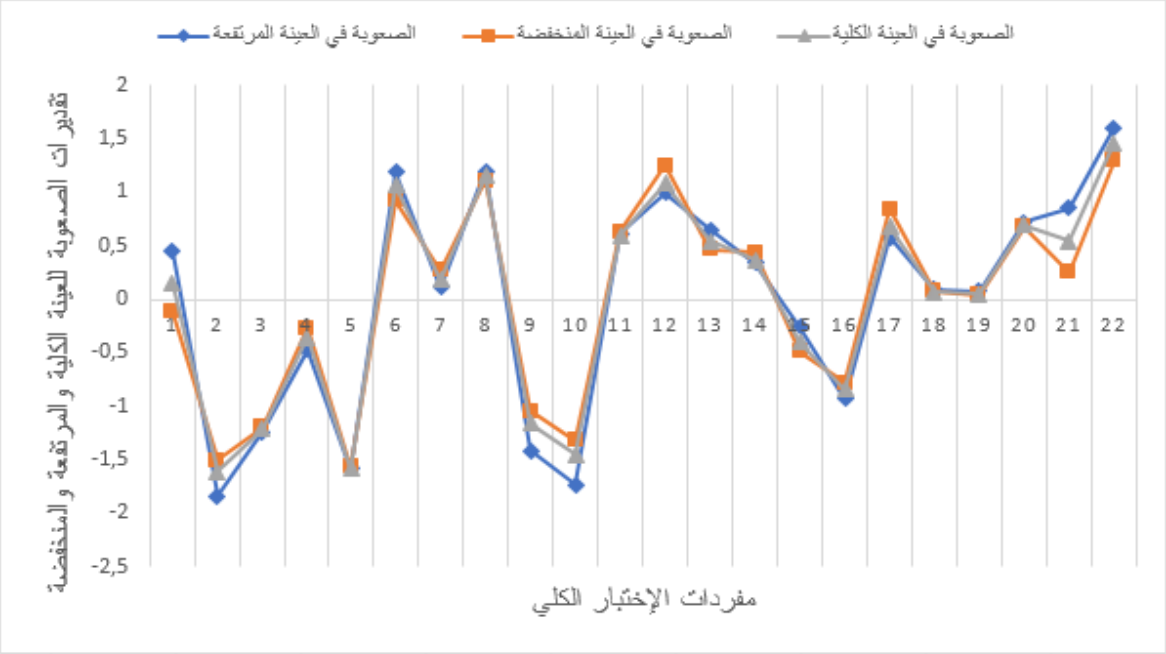
الجدول (40)

تقديرات الصعوبة (باللوجيت) المقابلة لكل مفردة من مفردات الاختبار المشتقة من كل من العينة الكلية، والعينة المرتفعة، والعينة المنخفضة، وكذا أخطائها المعيارية.

	العينة الكلية		العينة المنخفضة المستوى			العينة المرتفعة المستوى		
رقم المفردة	تقدير الصعوبة	الخطأ المعياري	تقدير الصعوبة	الفرق بين التقديرين	الخطأ المعياري	تقدير الصعوبة	الفرق بين التقديرين	الخطأ المعياري
1	0.15	0.09	0.12–	0.03–	0.12	0.45	0.3	0.13
2	1.61–	0.12	1.5–	0.11–	0.14	1.84–	0.23–	0.24
3	1.21–	0.11	1.19–	0.02–	0.13	1.24–	0.03	0.19
4	0.36–	0.09	0.27–	0.09–	0.12	0.47–	0.11	0.15
5	1.58–	0.12	1.56–	0.02–	0.14	1.58–	0	0.22
6	1.06	0.1	0.92	0.14–	0.15	1.19	0.13	0.13
7	0.18	0.09	0.27	0.09	0.13	0.11	0.07–	0.13
8	1.15	0.1	1.11	0.04–	0.16	1.2	0.05	0.13
9	1.16–	0.11	1.04–	0.12–	0.13	1.41–	0.25	0.2
10	1.45–	0.12	1.32–	0.13–	0.14	1.73–	0.28	0.23
11	0.6	0.09	0.63	0.03	0.14	0.61	0.01	0.13
12	1.08	0.1	1.25	0.17	0.16	0.99	0.09–	0.13
13	0.54	0.09	0.46	0.08–	0.13	0.65	0.11	0.13
14	0.37	0.09	0.44	0.07	0.13	0.34	0.03–	0.13
15	0.4–	0.1	0.49–	0.09	0.13	0.26–	0.14–	0.14
16	0.84–	0.1	0.78–	0.06–	0.13	0.93–	0.09	0.17
17	0.68	0.09	0.84	0.16	0.14	0.58	0.1–	0.13
18	0.07	0.09	0.08	0.01	0.13	0.09	0.02	0.13
19	0.05	0.09	0.04	0.01–	0.13	0.08	0.03	0.13
20	0.69	0.09	0.68	0.01–	0.14	0.72	0.03	0.13
21	0.54	0.09	0.25	0.29–	0.13	0.85	0.31	0.13
22	1.46	0.1	1.29	0.17–	0.16	1.6	0.14	0.13

باعتبار الخطأ المعياري لتقديرات الصعوبة، نلاحظ تكافؤ تلك التقديرات المتناظرة المشتقة من تحليل أداء أفراد العينة الكلية، كتقديرات مرجعية، وتلك المشتقة من أداء كل من العينة المرتفعة المستوى، والعينة المنخفضة المستوى، وذلك لصعوبة مفردات الاختبار بعد التدريج، حيث أن الفرق بين أي تقديرين متناظرين

من تقديرات الصعوبة أقل من مجموع الخطأ المعياري لهما. باستثناء المفردة 21 والتي كان فيها الفرق بين تقديري صعوبة العينة الكلية و العينة المنخفضة المستوى يساوي 0.29- وهو أكبر بقليل من مجموع الخطأ المعياري للتقديرين المتناظرين والذي يساوي 0.22. وهو ما يوضحه الشكل رقم(34) لتقديرات الصعوبة لكل مفردة المشتقة من مفردات الاختبار الكلي، و أفراد العينتين المرتفعة والمنخفضة المستوى.



الشكل (34) يوضح تقديرات الصعوبة لكل مفردة المشتقة من مفردات الاختبار الكلي و أفراد العينتين المرتفعة والمنخفضة المستوى.

يتضح من الشكل رقم (34) التكافؤ الاحصائي بين تقديرات الصعوبة لكل مفردة من مفردات اختبار مستويات التفكير الهندسي المشتقة من العينة المرتفعة، والمنخفضة المستوى، والعينة الكلية.

وبهذا يتحقق بوجه عام الشق الثاني من استقلالية القياس عن العينة المستخدمة لهذه المفردات، وهو تحرر تقديرات صعوبة المفردة عن قدرة الأفراد الذين يجيبون على الاختبار.

ثانياً. استقلال القياس عن مجموعة المفردات التي يجيب عليها الأفراد أي:

أ. لا يعتمد تقدير صعوبة المفردة على باقي مفردات الاختبار.

ب. لا يعتمد تقدير قدرة الفرد على المفردات المعينة التي يجيب عليها.

ويتطلب التحقق من هذا الفرض، تحليل استجابات مجموعة واحدة من الأفراد لاختبارين مكونين من مجموعتين من المفردات الملائمة للنموذج بشرط:

- استخدام نموذج (راش) في تدريج مفردات الاختبارين في تدريج واحد مشترك.

- تقارب مستوى قدرة الأفراد مع صعوبة المفردات. (Elliott, 1983) كما ورد عن (كاظم، 1988ب، ص.105).

ولتحقيق هاتين النقطتين قام الباحث بما يأتي:

1. تجزئة الاختبار الكلي المكون من 22 مفردة إلى اختبارين فرعيين، أحدهما يمثل الاختبار الصعب ويضم 11 مفردة الأصعب من مفردات الاختبار الكلي، والآخر يمثل الاختبار السهل ويضم 11 مفردة الأسهل من مفردات الاختبار الكلي.

2. وباستخدام برنامج وينستبس، حللت بيانات استجابات عينة الأفراد على كل من مفردات الاختبار الصعب، وكذا مفردات الاختبار السهل، وذلك وفقاً لنموذج (راش).

3. حددت تقديرات الصعوبة المقابلة لكل مفردة من مفردات الاختبار الصعب، كما حددت كذلك بالنسبة لمفردات الاختبار السهل، وكذا أخطاؤهما المعيارية.

4. حدد تقديرات القدرة للأفراد المقابلة لكل درجة محتملة على الاختبار الصعب، كما حددت كذلك بالنسبة للاختبار السهل، وكذا أخطاؤهما المعيارية.

5. عمل الباحث إجراءات التعادل الرأسي لكل من الاختبارين السهل والصعب وذلك بوضعهما على تدريج واحد مشترك باستخدام مجموعة مشتركة من الأفراد هي عينة التحليل الكلية 566 فرداً. (Wright and Stone,1979) كما ورد عن (كاظم، 1988ب، ص.106).

6. قورنت تقديرات صعوبة المفردات المشتقة من أي من الاختبارين الصعب أو السهل بعد إجراء التعديل المطلوب، بتلك المشتقة من الاختبار الكلي حيث تتدرج المفردات باعتبارها اختبار كلي واحد.

7. قورنت تقديرات القدرة المشتقة من الاختبار الصعب بعد التعديل، بتلك المشتقة من الاختبار السهل بعد التعديل، بتلك المشتقة من الاختبار الكلي.

- قام الباحث بإجراء عملية التعادل الرأسي بين الاختبارين الصعب والسهل، وفق الخطوات التالية:

1. تقدير الفرق بين صعوبة الاختبارين الصعب والسهل، وذلك بواسطة الفرق الملاحظ بين متوسطي قدرة الأفراد الذين استجابوا للاختبارين.

2. تقسيم الفرق الملاحظ على مجموع المفردات الصعبة والسهلة، ليصبح متوسط صعوبة المفردات الكلي= صفر.

3. إزاحة مفردات الاختبار الصعب= $\frac{\text{عدد المفردات الكلية+عدد المفردات الصعبة}}{\text{عدد المفردات الكلية}}$ X الفرق بين متوسطي القدرة

4. إزاحة مفردات الاختبار السهل= $\frac{\text{عدد المفردات الكلية+عدد المفردات السهلة}}{\text{عدد المفردات الكلية}}$ X الفرق بين متوسطي القدرة

الجدول (41)

يوضح تقسيم مفردات الاختبارين الصعب والسهل ومتوسط القدرة ومقدار الإزاحة للاختبار.

الاختبار	عدد المفردات	عدد مفردات للاختبار		متوسط قدرة الاختبار		الفرق بين القدرتين	مقدار الإزاحة للاختبار	
		السهل	الصعب	السهل	الصعب		السهل	الصعب
	22	11	11	1.15	0.48–	1.63	0.82–	0.82+

5. ليصبح التدريج مشتركا تطرح الإزاحة لمفردات الاختبار السهل من كل مفردة في الاختبار السهل، وتضاف الإزاحة لمفردات الاختبار الصعب لكل مفردة من مفردات الاختبار الصعب، و تحسب بالأسلوب نفسه لتقديرات القدرة.

6. تقارن تقديرات صعوبة الاختبار السهل والصعب بعد تعديل تدريجهما إلى التدريج المرجعي، وتعد هذه التقديرات متكافئة إحصائياً إذا لم يتجاوز الفرق بين أي تقديرين متناظرين مجموع الخطأ المعياري لهما والشيء نفسه لتقديرات القدرة، والجدول رقم (42) يوضح المقارنة بين صعوبة مفردات الاختبارين السهل والصعب والتدريج المرجعي والتدريج المشترك بعد إجراء عملية التعادل الرأسي للاختبارين.

الجدول (42)

يوضح عملية إجراء الموازنة بين صعوبة مفردات كل من الاختبارين الصعب والسهل والتدريج المرجعي بعد إجراء التعادل الرأسي.

رقم المفردة	التدريج المستقل لكل من الاختبارين		الخطأ المعياري للاختبار السهل أو الصعب	التدريج المشترك للاختبارين (بعد التعديل)		التدريج المرجعي الكلي	الفرق بين التدريجين	الخطأ المعياري للتدريج المرجعي
	السهل	الصعب		السهل	الصعب			
				0.82–	0.82+			
1	0.97		0.10	0.15–		0.15	0.00	0.09
3	0.90–		0.12	1.72		1.61–	0.11	0.12
5	0.49–		0.11	1.31		1.21–	0.10	0.11
6	0.42		0.10	0.40		0.36–	0.04	0.09
7	0.87–		0.12	1.69		1.58–	0.11	0.12
8		0.31	0.10		1.13	1.06	0.07	0.10
9		0.60–	0.09		0.22	0.18	0.04	0.09
10		0.40	0.10		1.22	1.15	0.07	0.10
11	0.44–		0.11	1.26		1.16–	0.10	0.11
12	0.74–		0.12	1.56		1.45–	0.11	0.12
14		0.16–	0.09		0.66	0.60	0.06	0.09
15		0.33	0.10		1.15	1.08	0.07	0.10
16		0.23–	0.09		0.59	0.54	0.05	0.09
17		0.40–	0.09		0.42	0.37	0.05	0.09
18	0.38		0.10	0.47–		0.40–	0.07	0.10
19	-0.09		0.10	0.91		0.84–	0.07–	0.10
22		0.08–	0.09		0.74	0.68	0.06	0.09
23	0.90		0.10	0.08–		0.07	0.01	0.09
24	0.86		0.10	0.04–		0.05	0.01–	0.09
25		0.07–	0.10		0.64	0.69	0.05	0.09
26		0.23–	0.09		0.59	0.54	0.05	0.09
27		0.72	0.11		1.54	1.46	0.09	0.10
المتوسط	0	0		0.54	0.43–	0.00		

يتضح من الجدول رقم (42)، أن الفرق بين تقدير صعوبة مفردات الاختبار الصعب بعد تعديل تدريجه، والتقديرات المتناظرة لصعوبة تلك المفردات كما تشتق من تحليل أداء الأفراد على الاختبار الكلي، كانت متكافئة إحصائياً ولم يتجاوز الفرق بين أي تقديرين متناظرين مجموع الخطأ المعياري لهما، كما يلاحظ أن الفرق بين تقدير صعوبة مفردات الاختبار السهل بعد تعديل تدريجه، والتقديرات المتناظرة لصعوبة تلك المفردات كما تشتق من تحليل أداء الأفراد على الاختبار الكلي، كانت متكافئة إحصائياً ولم يتجاوز الفرق بين أي تقديرين متناظرين مجموع الخطأ المعياري لهما، أي أن تقديرات صعوبة المفردات المشتقة من تحليل أداء الأفراد على الاختبار الكلي تتكافأ إحصائياً مع التقديرات المتناظرة لصعوبة المفردات المشتقة من تحليل أداء نفس الأفراد على الاختبار السهل أو الاختبار الصعب، مما يعني عدم تأثر تقدير صعوبة المفردات بإختلاف مجموعة المفردات المستخدمة في التحليل. **وهذا يعني تحرر تقدير صعوبة المفردة من صعوبة باقي المفردات المستخدمة في الاختبار**. وبهذا يتحقق الشطر الأول من استقلالية القياس عن مجموعة المفردات المستخدمة في الاختبار.

▪المقارنة بين تقدير القدرة المشتقة من كل اختبار على حده:

لا يتوقع أنه اذا تساوت الدرجة الكلية على الاختبار أن يتساوى تقدير قدرة الفرد، ولكن ما يتوقع هو أن يكون للفرد الواحد درجة كلية مختلفة على الاختبار. فإذا قابلت كل درجة تقديراً متكافئاً من تقديرات القدرة فهذا يعني تحرر قدرة الفرد من مجموعة المفردات المستخدمة. ولما كانت كل من صعوبة المفردة وقدرة الفرد تتدرجان على نفس التدريج، فينبغي عند تقديرنا لقدرة الفرد المقابلة للدرجة الكلية على الاختبار الصعب، وتلك المقابلة للدرجة على الاختبار السهل، أن نراعي مقدار الإزاحة الذي سبق أن راعيناه عند تعديل تدرج صعوبة المفردات لكل من الاختبارين، وعلى هذا عند إضافة المقدار +0.82 إلى تقدير قدرة الفرد المشتق من الاختبار الصعب، وطرح المقدار −0.82 من تقدير قدرته المشتق من الاختبار السهل نحصل على تقديرين لقدرة هذا الفرد، أي وضع متوسط صعوبة المفردات عند الصفر أو بما يسمى (إجراء الاعتماد)، مما يؤدي إلى متوسط قدرة مقدرة للمجموعة تعكس متوسط صعوبات المفردات قبل إعادة القياس. ولذلك ما كان في الأصل اختلافات في متوسط صعوبة الاختبارات الثلاثة يصير الأن فروق في متوسط قدرة المجموعة المشتركة للمفحوصين. فإذا تكافأ التقديران مع بعضهما ومع تقدير قدرة هذا الفرد المشتق من الاختبار الكلي كان هذا دليلاً على عدم تأثر تقدير قدرة الفرد بمفردات الاختبار المستخدم. والجدول رقم (43) يوضح نتائج

الجدول (43)

يوضح تقديرات قدرات عينة من الأفراد المشتقة من الاختبار الصعب والسهل والكلي وأخطائهما المعيارية

الفرد	الدرجة الكلية على الاختبار			تقديرات القدرة قبل الإزاحة على الاختبار		تقديرات القدرة بعد الإزاحة على الاختبار		تقدير القدرة من الاختبار	الفرق بعد الإزاحة بين الاختبار الكلي والاختبار		الخطأ المعياري لتقديرات القدرة للاختبارات		
	الصعب	السهل	الكلي	الصعب	السهل	الصعب	السهل		الصعب	السهل	الصعب	السهل	الكلي
1	8	10	18	1.01	2.5	1.83	1.68–	1.75	0.08	0.07	0.69	1.07	0.58
2	4	6	10	0.58–	0.2	0.24	0.62	0.2–	0.04	0.42	0.64	0.64	0.47
3	6	10	16	0.38	2.5	1.2	1.68–	1.31	0.11	0.37–	0.66	1.07	0.54
4	2	7	9	0.84–	1.9	0.02–	1.08–	0.45	0.43	0.63–	0.88	1.1	0.61
5	3	8	11	1.01–	1.55	0.19–	0.73–	0.18	001–	0.55–	0.69	0.82	0.48
6	6	8	14	0.34	1.09	1.16	0.27–	0.76	0.4	0.49	0.65	0.71	0.49
7	7	10	17	0.58	2.5	1.4	1.68–	1.44	0.04	0.24–	0.64	1.07	0.54
8	2	10	12	1.55–	2.5	0.73–	1.68–	0.24	0.49–	1.44–	0.79	1.07	0.47
9	4	7	11	0.31–	0.62	0.51	0.2	0.15	0.36	0.05	0.68	0.66	0.49
10	7	7	14	0.58	0.62	1.4	0.2	0.69	0.71	0.49	0.64	0.66	0.48

يتضح من الجدول رقم (43)، باعتبار الخطأ المعياري للقياس، تكافؤ قدرة الفرد المشتقة من الاختبار السهل، ومن الاختبار الصعب بعد تعديل تدريجهما، وكذلك المشتقة من الاختبار الكلي. وهذا ما يعني عدم تأثر تقدير قدرة الفرد بمجموعة المفردات المستخدمة في الاختبار، وهذا ما يدل على تحرر تقدير قدرة الفرد عن مجموعة المفردات المستخدمة في الاختبار، بما يعني التحقق من الشطر الثاني من استقلالية القياس عن مجموعة المفردات المستخدمة في الاختبار. **وبذلك تحقق ثاني فرضيات النموذج، وهي استقلالية القياس عن كل من عينة الأفراد و مجموعة المفردات المستخدمة.**

1-6- عرض وتحليل نتائج الفرضية السادسة:

وتنص على أن:" يمكن اشتقاق معايير كمية تفسر تقديرات الأفراد على اختبار مستويات التفكير الهندسي في صورته النهائية بعد التدريج على تلامذة المرحلة الثانوية".

لاختبار صحة الفرضية، تم تحويل تدريج وحدات اللوجيت لصعوبة مفردات الاختبار في صورته النهائية (بعد التدريج)، إلى تدرج وحدات مئوية بوحدة الواط، حيث تشتق وحدة اللوجيت من نموذج راش اللوغاريتمي، والذي يتناول التقدير الاحتمالي للاستجابة الصحيحة للفرد على المفردة ضمن التدريج الذي نقطة صفره متوسط صعوبات المفردات، ونظراً لوجود مشكلة الكسور، وعدم ألفة الباحثين والمدرسين لهذا النوع من التدريج، ولمعالجة هذه العيوب، وبما أن نظام التدريج المئوي هو أكثر التدريجات ألفة في أغلب مجالات القياس، أوجد بعض المتخصصين في القياس عدداً من الوحدات الجديدة مثل وحدة التدريج المعتمد على محك مستقل سيت، ووحدات التدريج الجماعية نيت، ووحدات التدريج الخاصة باحتمال الاستجابة الصواب شيبس، ووحدة التدريج المستخدمة في المقاييس البريطانية للقدرات باص، ولهذا وقع اختيار الباحث من تلك الوحدات على وحدة قياس الواط التي قدمها (Masters,1984)، لتحويل تقديرات كل من صعوبات المفردات وقدرات الأفراد من وحدة القياس باللوجيت إلى هذه الوحدة الجديدة وهي وحدة الواط، وذلك باستخدام المعادلتين التاليتين:

$$B= 50 + (15 / \log 4)b$$

$$D= 50 + (15 / \log 4)d$$

حيث أن B، D هما تقدير كل من القدرة والصعوبة على الترتيب مقدرتين بالواط، و b، d هما تقدير كل من القدرة والصعوبة على الترتيب مقدرتين باللوجيت، وحيث log هو اللوغاريتم الطبيعي، ولذا يكون متوسط صعوبة المفردات 50، كما تتدرج كل من B، D من الصفر إلى المائة.

إضافة إلى ذلك استخرج الباحث معايير الدرجات التائية، وكذلك معايير الرتب المئينية المقابلة لكل درجة من الدرجات الخام التى حصل عليها الأفراد في عينة التقنين، وهذه المعايير هي الطريقة الكلاسيكية المألوفة، وقد قدمها الباحث لإجراء مقارنة بين المعايير المستخرجة وفق النظريتين الكلاسيكية في القياس، ونظرية الاستجابة للمفردة (نموذج راش). وفي ما يلي عرض لهذه المعايير التي على أساسها يمكن تحديد وتفسير قدرة التلامذة على هذا الاختبار سواء بالتفسير المحكي أو المعياري المرجع. والجدول رقم (44) يوضح تقدير صعوبة مفردات اختبار مستويات التفكير الهندسي في صورته النهائية مقدرة بوحدة اللوجيت، ووحدة الواط.

الجدول (44)

لتقدير صعوبة المفردات مقدرة باللوجيت والواط مرتبة تصاعدياً تبعاً للصعوبة 22 مفردة.

المفردة	الدرجة	تقديرات الصعوبة		أخطائها المعيارية	
	الخام	باللوجيت	بالواط	باللوجيت	بالواط
27	146	1.46	72	0.10	1.56
10	172	1.15	67	0.10	1.50
15	186	1.08	66	0.10	1.46
8	173	1.06	65	0.10	1.51
25	226	0.69	60	0.09	1.41
22	229	0.68	60	0.09	1.40
14	241	0.60	59	0.09	1.38
16	247	0.54	58	0.09	1.38
26	246	0.54	58	0.09	1.38
17	266	0.37	56	0.09	1.38
9	294	0.18	53	0.09	1.36
1	295	0.15	52	0.09	1.37
23	303	0.07	51	0.09	1.37
24	310	0.05	51	0.09	1.37
6	361	0.36–	45	0.09	1.40
18	355	0.40–	44	0.10	1.43
19	404	0.84–	37	0.10	1.52

1.62	0.11	33	1.16–	441	11
1.65	0.11	32	1.21–	442	5
1.74	0.12	28	1.45–	465	12
1.80	0.12	26	1.58–	475	7
1.82	0.12	26	1.61–	475	3

يتضح من الجدول رقم (44) تقدير الصعوبة لمفردات الاختبار مقدرة بوحدة اللوجيت، وكذا بوحدة الواط، مرتبة تصاعدياً تبعاً لصعوبتها، والتي انحصرت فيها الدرجات الخام ما بين 146 و475، و كانت أقل قيمة من نصيب المفردة 27 وهي مفردة من المستوى الاستدلال الشكلي لمستويات فان هيل للتفكير الهندسي وهو أرقى المستويات في التدرج الهرمي لنموذج (فان هيل)، والذي تقابله قيمة تقديرية للصعوبة بوحدة اللوجيت 1.46 وخطأ معياري قدره 0.10 والذي يقترب من الصفر، و الدرجة 72 بوحدة الواط وبخطأ معياري قدره 1.56. في حين كانت أكبر قيمة من نصيب المفردة 3 وهي مفردة في المستوى الإدراكي أو البصري لمستويات فان هيل للتفكير الهندسي وهو أدنى المستويات في التدرج الهرمي لنموذج (فان هيل)، والذي تقابله قيمة تقديرية للصعوبة بوحدة اللوجيت –1.61 وخطأ معياري قدره 0.12 والذي يقترب من الصفر هو كذلك، والدرجة 26 بوحدة الواط وبخطأ معياري قدره 1.82. فالطريقة الملائمة لتفسير أي قيمة رقمية لمعلم الصعوبة لمفردات اختبار مستويات التفكير الهندسي، تقوم على تحديد أين تعمل المفردة على مقياس القدرة. والجدول رقم (45) يوضح معايير القدرة لأفراد عينة البحث، بإستخدام نموذج راش (وحدة اللوجيت – وحدة الواط)، و النظرية الكلاسيكية في القياس (المعايير التائية – الرتب المئينية).

الجدول (45)

يوضح معايير القدرة لعينة التقنين 566 تلميذًا وتلميذة، باستخدام نموذج راش (وحدة اللوجيت – وحدة الواط)، و باستخدام النظرية الكلاسيكية في القياس (المعايير التائية – الرتب المئينية).

الدرجة	معايير القدرة المقابلة			
الخام	وحدة اللوجيت	وحدة الواط	الدرجة التائية	الرتب المئينية
3	2.15–	18	34	1
4	1.77–	23	36	1
5	1.44–	28	37	2
6	1.16–	33	39	4
7	0.89–	37	41	6

11	42	40	0.65–	8
18	44	44	0.42–	9
29	46	47	0.2–	10
40	**47**	**50**	**0.02**	**11**
51	49	54	0.24	12
62	51	57	0.46	13
71	53	60	0.69	14
78	54	64	0.92	15
83	56	68	1.17	16
89	58	72	1.44	17
94	59	76	1.75	18
97	61	82	2.12	19
99	63	89	2.61	20
99	64	101	3.38	21
99	66	119	4.63	22

يتضح من الجدول رقم (45) تقدير القدرة المقابل لكل درجة من درجات عينة التقنين مقدرة بوحدة اللوجيت، وتقابلها كذلك نفس هذه التقديرات مقدرة بوحدة الواط. وفي نفس هذا الجدول تقدم الرتب المئينية والدرجات التائية المقابلة لكل درجة من هذه الدرجات الخام التي سجلها أفراد عينة التقنين.

- **المقارنة بين قدرات الأفراد بمفردات الاختبار ومجموعات المعايرة:**

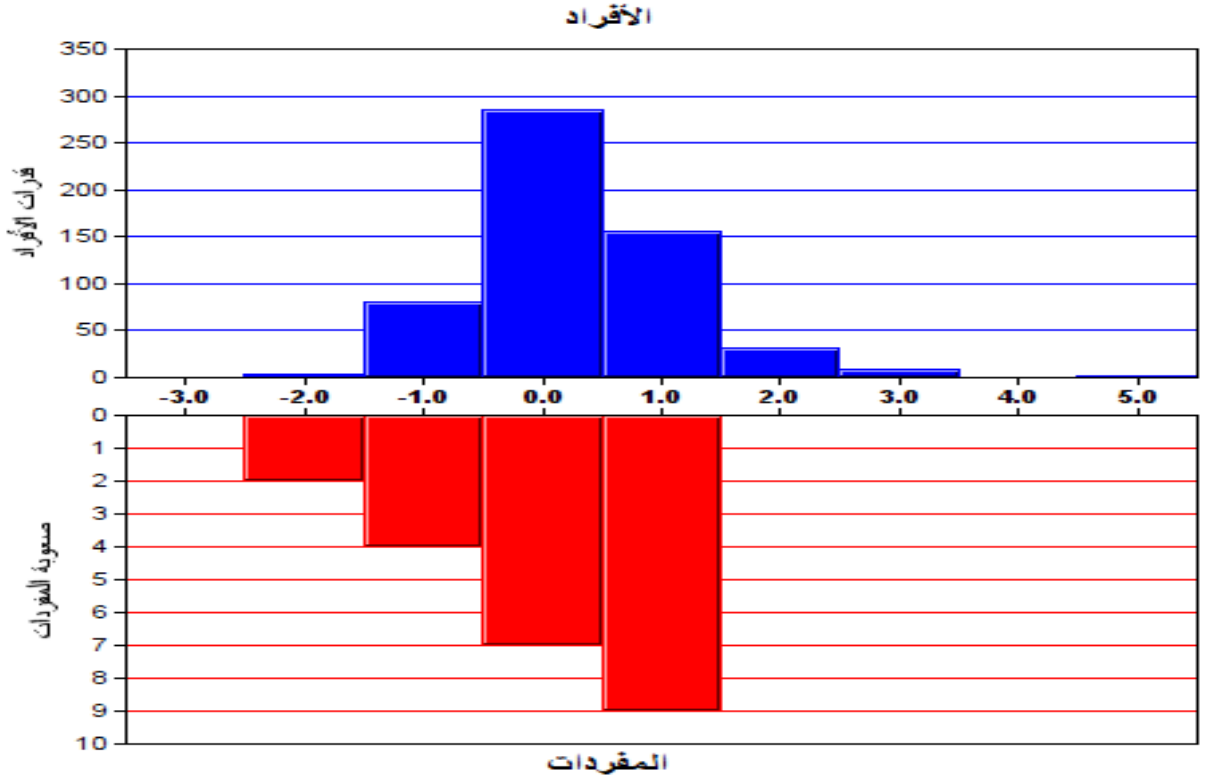
وفي هذه العملية قام الباحث بتمثيل توزيع قدرات مجموعة المعايرة – باستخدام برنامج وينستبس – بمدرجات تكرارية في أعلى المحور الذي يدل على القدرات للأفراد، كما تم تمثيل توزيع درجات الصعوبة للمفردات بمدرجات تكرارية في أسفل المحور ذاته، كما يوضحه الشكل رقم (35)، أي أنه إذا تم التعرف على قدرة فرد معين فيمكن وصف مكانته بالنسبة لكل من توزيع القدرة في مجموعة المعايرة وتوزيع درجة الصعوبة للمفردات، من خلال تمثيل توزيع متغير وحدة اللوجيت على مجموعة المعايرة التي عددها 566 تلميذاً وتلميذة، إذ أن توزيع متغير درجة الصعوبة للمفردات الذي هو على نفس مقياس قيم القدرة محسوبة من خلال النموذج ذو المعلمة الواحدة (نموذج راش) يتبع التوزيع المبين في الجدول رقم (46):

الجدول (46)

يمثل توزيع متغير القدرة باللوجيت على مجموعة المعايرة 566 تلميذاً وتلميذة، وتوزيع متغير درجة صعوبة مفردات اختبار مستويات التفكير الهندسي باللوجيت 22 مفردة.

فئات القدرة للأفراد	تكراراتها	فئات صعوبة مفردات الاختبار	تكراراتها
(2.5-)-(2.0-)	1	(2.0-)-(1.5-)	2
(2.0-)-(1.5-)	1	(1.5-)-(1.0-)	3
(1.5-)-(1.0-)	23	(1.0-)-(0.5-)	1
(1.0-)-(0.5-)	57	(0.5-)-(0.0)	2
(0.5-)-(0.0)	115	(0.0)-(0.5)	5
(0.0)-(0.5)	181	(0.5)-(1.0)	5
(0.5)-(1.0)	77	(1.0)-(1.5)	4
(1.0)-(1.5)	69	المجموع	22
(1.5)-(2.0)	21		
(2.0)-(2.5)	10		
(2.5)-(3.0)	7		
(3.0)-(3.5)	2		
(3.5)-(4.5)	2		
المجموع	566		

ومن خلال الجدول (46)، يمكننا أن نقوم بتمثيل وبربط التوزيعين معاً، بحيث تتم مقارنة قدرة أي من الأفراد بالنسبة للمفردات التي تمت بموجبها عملية المعايرة على عينة التقنين، والتي لا يوضع في الاعتبار فيها توزيع المفحوصين على مدى القدرات للمجموعة، لكن يهتم فقط بمستويات القدرة. ذلك لأن عدد الأفراد المفحوصين عند كل مستوى لا يؤثر على خاصية الثبات للمجموعة، كما يتضح من الشكل رقم (35):



الشكل (35) توزيع العتبات الفارقة للمفردات والأفراد على عينة التدريج 566 تلميذاً وتلميذة.

يتبين من الشكل رقم (35)، أن توزيع القدرات لمجموعة (التدريج) المعايرة، بأنه متماثلاً بطريقة مقبولة وذا إرتفاع واضح جداً في الوسط، والذي يشير إلى أنه يتم تقدير الصعوبة والقدرة بأعظم دقة بمصاحبة متوسط مقياس القدرة، والذي يستطيع أن يأخذ أياً من الأشكال، حيث لا يفترض فيه أن يأخذ شكل توزيع المنحنى الطبيعي، والذي يبين العتبات الفارقة للمفردات بدلاً من الصعوبة، والذي يظهر كذلك توزع عتبات مفردات اختبار مستويات التفكير الهندسي، من أسهل المفردات –1.61 لوجيت إلى الأصعب نوعاً ما 1.46 لوجيت. وتتدرج قدرات الأفراد على الاختبار من أقل قدرة حوالي –2.15 لوجيت إلى أكبر قدرة حوالي 4.63 لوجيت.

1-7- عرض وتحليل نتائج الفرضية السابعة:

وتنص على أن:" يتحقق صدق وثبات اختبار مستويات التفكير الهندسي في صورته النهائية بعد التدريج باستخدام نموذج راش".

ولاختبار صحة هذا الفرض، اعتمد الباحث برنامج وينستبس لتقدير مؤشرات صدق وثبات درجات اختبار مستويات التفكير الهندسي بصورته النهائية (بعد التدريج).

أولاً: تقويم الصدق:

تبين للباحث من خلال تحليل النتائج الموضحة في الجدول رقم (37)، والشكل رقم (27)، إلى أن قيم مؤشرات الملاءمة الإحصائية لمتوسط المربعات التباعدية والتقاربية لمفردات اختبار مستويات التفكير الهندسي كانت قريبة من الوضع المثالي الذي يفترضه النموذج وهو القيمة واحد، والتي تراوحت ما بين 0.84 و 1.13، بمتوسط حسابي 1 وانحراف معياري 0.04 بالنسبة لمؤشرات الملاءمة الإحصائية لمتوسط المربعات التقاربية (الداخلية)، وبمتوسط حسابي 0.99 وانحراف معياري 0.08 بالنسبة لمؤشرات الملاءمة الإحصائية لمتوسط المربعات التباعدية (الخارجية). و بالتالي فهي في مجال القياس المنتج الذي حدده كلٍ من رايت ولينكر (Wright and Linacre, 1994) بالقيم ما بين 0.70 و1.30 أو ما بين 0.50 و1.5 عموماً والقريبة من الواحد. وتتعلق هذه المؤشرات بصدق تدرج مفردات الاختبار في تعريفها للمتغير موضوع القياس. كما تتعلق أيضاً بصدق تدرج قدرات الأفراد على متصل هذا المتغير، الذي يقوم على صدق استجابات الأفراد على الاختبار. ولإضفاء المزيد من المصداقية على الاختبار الحالي، قام الباحث باستخدام الأساليب التالية للتحقق من الصدق:

1. التحقق من قطبية المفردات (Polarity):

أظهرت نتائج التحليل الموضحة في الجدول رقم (37)، أن نقطة قياس معاملات الارتباط لكل مفردات اختبار مستويات التفكير الهندسي كانت قيمها موجبة (point measure correlation PT-MEA CORR)، والتي تراوحت ما بين 0.27 للمفردة 7 و8، والقيمة 0.47 للمفردة 15، مما يدل على أن مفردات اختبار مستويات التفكير الهندسي تقيس التكوين الفرضي الذي ينوي الباحث قياسه.

2. التحقق من أحادية البعد:

استخدم الباحث طريقة التقدير وفقاً لنموذج (راش) للمكونات الأساسية المعتمدة على البواقي، والذي تظهر الاختلافات بين الأبعاد، و النتائج يوضحها الجدول رقم (47):

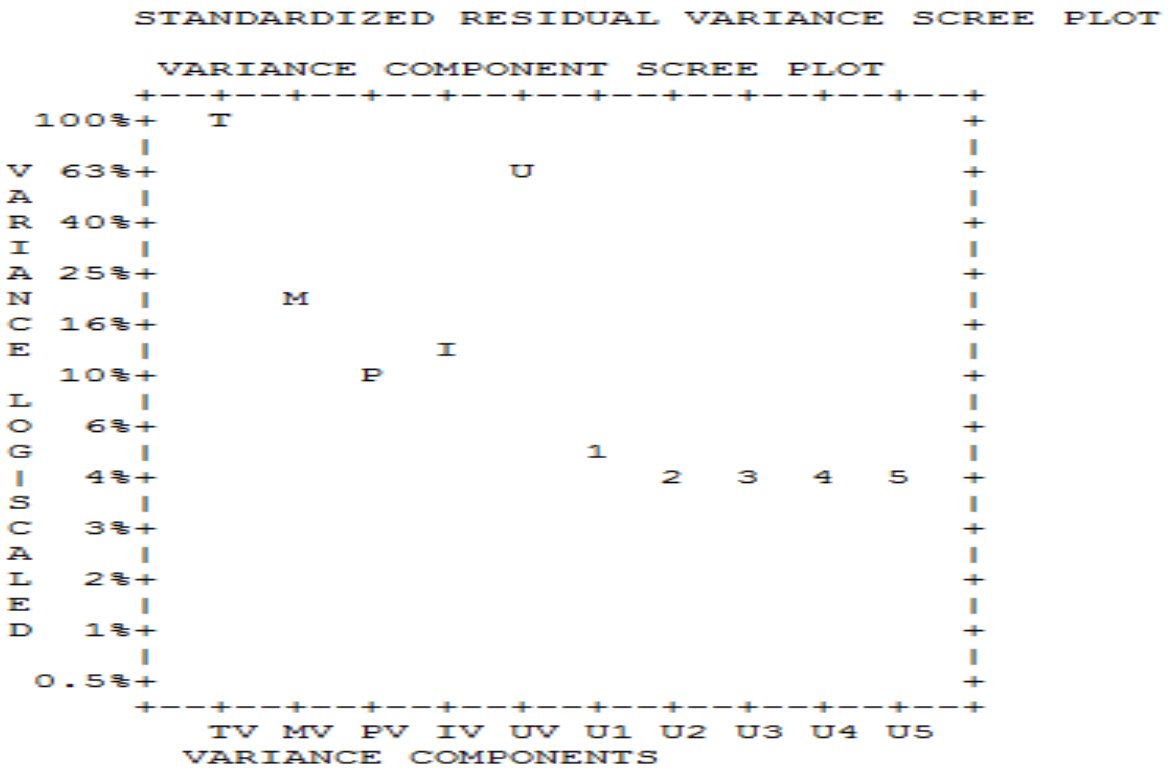
الجدول (47)

نتائج التحليل العاملي للمكونات الأساسية للبواقي (PCAR) تقديرات نموذج (راش) بعد التدريج.

حجم تباين البواقي المعيارية مقدرا بوحدات القيم *Eigen*.	الجذر الكامن	الملاحظ	المتوقع
التباين الكلي في الاستجابات	28.7690	%100	%100
التباين الذي فسره العامل الرئيسي	6.7690	%23.5	%23.6

التباين المفسر بواسطة الأفراد	3.0125	%10.5	%10.5
التباين المفسر بواسطة المفردات	3.7565	%13.1	%13.1
مجموع التباين غير المفسر	22.0000	%76.5	%76.4
التباين الذي فسره العامل الثاني (الأول في البواقي)	1.4738	%5.1	–

يتضح من نتائج الجدول رقم (47) تحقق المحكات الثلاثة المعتمدة في تحديد أحادية البعد، حيث يتحقق المحك الأول وكقاعدة عامة إذا كانت قيمة التباين المفسر بواسطة التقديرات (Raw variance explained by measures) أكبر من أو يساوي 50 %، واسترشاداً بما أشار إليه لينكر (Linacre) مطور برنامج (winsteps)، الذي حدد المجال من 20 % إلى 80 % بأنه مجال مقبول و يعد مؤشراً قوياً لأحادية البعد– (إتصال شخصي في 18 أفريل2017)– أنظر الملحق رقم(10)، وهو ما تحقق في بحثنا الحالي، حيث إرتفع مؤشر القيمة الملاحظة للتباين الذي فسره العامل الرئيسي إلى 23.5 % بعد حذف ستة 6 مفردات من الاختبار، عن قيمته السابقة والتي كانت 21.8 % قبل الحذف، كما أن الفرق بين القيمة الملاحظة و القيمة المتوقعة للنموذج هي 0.1 % فقط. أما المحك الثاني، هو أن نسبة التباين الذي يفسره العامل الثاني (Unexplained Variance in 1st contrast)، يجب أن تكون قيمته أقل من 5 % وهذا ما تحقق في البحث الحالي تقريباً بالقيمة 5.1 %، وهذا دليل آخر على أحادية البعد. في حين يعتمد المحك الثالث، على قيمة الجذر الكامن لنسبة التباين الذي يفسره العامل الثاني (Unexplained Variance in 1st contrast) والذي يجب أن لا يتعدى أو يكون أقل قيمة له 3، وهذا ما تحقق في البحث الحالي بتسجيل القيمة 1.4738 وهي أقل من القيمة 3 المنصوح بها من ليناكر (Linacre, 2008, p.272). والشكل رقم (36) يوضح نتائج التحليل العاملي بالمكونات الأساسية للبواقي باستخدام نموذج راش أحادي البارامتر.



الشكل (36) يوضح اختبار (Scree Plot) لتباين البواقي المعيارية (أخطاء القياس)، مستخرج بواسطة برنامج وينستبس.

ثانياً: تقويم الثبات:

يبدو ثبات القياس بتحقق ما تبقى من مطالب الموضوعية في القياس، عندما تستخدم أداة القياس التي تبنى باستخدام نموذج (راش)، حيث أمكن التحقق من استقلالية القياس – في ما سبق من هذا البحث – من ناحيتين:(أ) تحرر القياس من قدرة العينة التي تؤدي الاختبار؛ (ب) تحرر القياس من المفردات التي يجيب عليها الأفراد، ولإضافة المزيد من المعلومات حول ثبات الاختبار استخدم الباحث الأساليب التالية:

1. تقويم اللاتغاير (Invariance) بالتقسيم العشوائي لعينة التدريج:

لتقويم نوعية المطابقة بين البيانات والنموذج، تم استخدام طريقة التقسيم العشوائي للعينة المدرجة باستخدام نموذ راش، إلى عينتين فرعيتين الأولى مكونة من 282 تلميذاً وتلميذة، بمتوسط عمري 18 سنة، وانحراف معياري قدره 1.62، والثانية مكونة من 284 تلميذاً وتلميذة، بمتوسط عمري 18 سنة، وانحراف معياري قدره 1.55. وبعد تقسيم العينة الكلية تمت المقارنة بين تقديرات معالم المفردة بينهما لتحديد درجة الخطية (Linearity) فيها، وذلك بحساب معامل الارتباط بيرسون بين تقديرات معالم المفردات للمجموعتين والذي بلغ القيمة 0.99.

2. معامل ثبات الأفراد ومعامل ثبات المفردات:

باستخدام نموذج (راش) أمكن الحصول على القيم المتحررة لكل من صعوبة المفردات وقدرات الأفراد، ثم الحصول على نوعين من المعاملات: معامل الثبات الخاص بالأفراد، ومعامل الثبات الخاص بالمفردات، فالثبات وفق نظرية الاستجابة للمفردة، يعني الدقة في تقدير موقع كل من الأفراد والمفردات على متصل السمة، ويمكن تحديد دقة المفردات في تعريف هذا المتصل بحساب معامل الفصل للمفردات (GI). والذي يعرف بأنه النسبة بين الانحراف المعياري للقيم التدريجية المتحررة من المفردات، ومتوسط الخطأ المعياري لهذه القيم، والتي يوضحها الجدول رقم (48):

الجدول (48)

ملخص لنتائج معاملات الثبات والفصل وأخطاءهما المعيارية للأفراد والمفردات وفق نموذج راش بعد التدريج النهائي لاختبار مستويات التفكير الهندسي في صورته النهائية 22 مفردة .

معامل الثبات		معامل الفصل		الإنحراف المعياري SD		الجذر التربيعي لمتوسط مربعات الأخطاء المعيارية للقياس RMSE	
الأفراد	المفردات	الأفراد	المفردات	الأفراد	المفردات	الحقيقي للأفراد REAL RMSE	الحقيقي للمفردات REAL RMSE
0.63	0.99	1.32	9.13	0.70	0.92	0.53	0.10
						لنموذج الأفراد MODEL RMSE	لنموذج المفردات MODEL RMSE
0.65	0.99	1.37	9.20	0.71	0.92	0.52	0.10

معامل الثبات ألفا كرونباخ أو(*KR-20*) يساوي 0.64 . و الخطأ المعياري للقياس يساوي SEM=2.09

يتضح من الجدول رقم (48)، أن قيمة معامل ثبات المفردات بلغت 0.99، ومعامل الفصل لمفردات الاختبار بلغ 9.20 وهي قيمة أكبر من 2. كما بلغت قيمة معامل الثبات الأفراد 0.65، والقيمة 1.32 لمعامل الفصل للأفراد (GP)، وهي قيمة قريبة من 2 نوعاً ما.

كما حدد الباحث عدد الطبقات الإحصائية المتمايزة لكل من المفردات والأفراد باستخدام الصيغة الرياضية التالية: H=(4*G+1) /3، حيث يرمز H إلى عدد الطبقات الإحصائية، و يرمز G لمعامل الفصل وبالتعويض في المعادلة السابقة بالنسبة للأفراد H=(4*1.32+1) /3؛ وكذلك بالنسبة للمفردات H=(4*9.2+1) /3، وبالتالي فإن عدد الطبقات الإحصائية للأفراد بلغ 2 ، و 13 للمفردات على التوالي.

كما يتضح من الجدول رقم (48)، أن قيم جذر مربع متوسط الخطأ المعياري للثبات الحقيقي للأفراد والمفردات لمعامل الفصل للأفراد والمفردات (REAL RMSE)، وهو تباين الخطأ الحقيقي، وهو يمثل أسوأ

حالة للثبات أو أدنى قيمة ثبات للاختبار، حيث سجل الباحث القيم 0.53؛ 0.52 على التوالي، فقيمة جذر مربع متوسط الخطأ المعياري للنموذج بالنسبة للأفراد، تكافئ قيمة الخطأ المعياري للقياس في النظرية الكلاسيكية للقياس والتي بلغت 2.09. (Linacre, 2012, p.316)

1-8- عرض وتحليل نتائج الفرضية الثامنة:

وتنص على أن:" لا يظهر تدريج صعوبة مفردات اختبار مستويات التفكير الهندسي في صورته النهائية أداءً تفاضلياً تبعاً لمتغير الجنس".

الخطوة الأولى التي استخدمها الباحث للكشف عن الأداء التفاضلي للمفردة تبعاً لمتغير الجنس، هي فحص الأبعاد الكامنة لمجموعة من المفردات للتحقق من شرط أحادية البعد وتكافؤ البنية العاملية للاختبار لكل مجموعة. اعتمد الباحث في هذا على التحليل العاملي باستخدام تحليل نموذج (راش) للمكونات الأساسية المعتمدة على تحليل تباين البواقي، والذي يظهر الاختلافات بين الأبعاد، والجدول رقم(49) يوضح مؤشرات هذا التحليل:

الجدول (49)

يوضح نتائج التحليل العاملي المكونات الأساسية للبواقي(PCAR) وفق تقديرات راش بعد التدريج وفقاً للجنس (الذكور، والإناث).

العينة	حجم تباين البواقي المعيارية مقدرا بوحدات القيم *Eigen*.		
	الجذر الكامن للتباين الذي فسره العامل الرئيسي	القيمة الملاحظة للتباين الذي فسره العامل الرئيسي	القيمة المتوقعة للتباين الذي فسره العامل الرئيسي
ذكور ن=(240)	6.161	%22.6	%22.6
إناث ن=(326)	7.1227	%24.5	%24.6
العينة	حجم التباين البواقي الذي فسره العامل الثاني (الأول في البواقي)		
	الجذر الكامن للعامل الثاني	القيمة الملاحظة	القيمة المتوقعة
ذكور ن=(240)	1.7063	%6.0	%7.8
إناث ن=(326)	1.6373	%5.6	%7.4

ولكي نحكم على أحادية البعد للاختبار بالتحليل العاملي بالمكونات الأساسية للبواقي باستخدام نموذج (راش)، تحقق الباحث من بعض المحكات من خلال ما يوضحه الجدول رقم (49):

1. تحقق المحك الأول، حيث بلغت قيمة التباين المفسر بواسطة التقديرات 22.6% عند الذكور و24.5% عند الإناث، واسترشاداً بما أشار إليه لينكر، والذي حدد المجال من 20% إلى 80% بأنه مجال مقبول، فإن هذا يعد مؤشراً قويا لأحادية البعد.

2. تحقق المحك الثاني، حيث أن قيمة الجذر الكامن لنسبة التباين الذي يفسره العامل الثاني بلغت 1.7063 عند الذكور، و1.6373 عند الإناث، وهي قيم أقل من القيمة 3 المحددة لهذا المحك. وبالتالي فهي دليل آخر على أحادية البعد. (Linacre, 2008, p.272)

وبعد أن تحقق الباحث من توفر شرط أحادية البعد للاختبار لكلا الجنسين، تم تحليل استجابات التلامذة عن مفردات اختبار مستويات التفكير الهندسي في صورته النهائية والمتكون من 22 مفردة، باستخدام برنامج وينستبس للكشف عن الأداء التفاضلي لمفردات الاختبار، واعتمادا على عدة مؤشرات أو اختبارات إحصائية، فإن المفردة تبدي أداءً تفاضلياً عندما تكون خصائصها الإحصائية مختلفة بين المجموعتين المرجعية والمستهدفة، بعد معادلة المجموعتين على مقياس القدرة، والجدول رقم (50) يوضح مؤشرات هذا التحليل:

الجدول (50)

يوضح نتائج الأداء التفاضلي لمفردات اختبار مستويات التفكير الهندسي في صورته النهائية حسب متغير الجنس (ذكور، إناث)، وباستخدام نموذج راش.

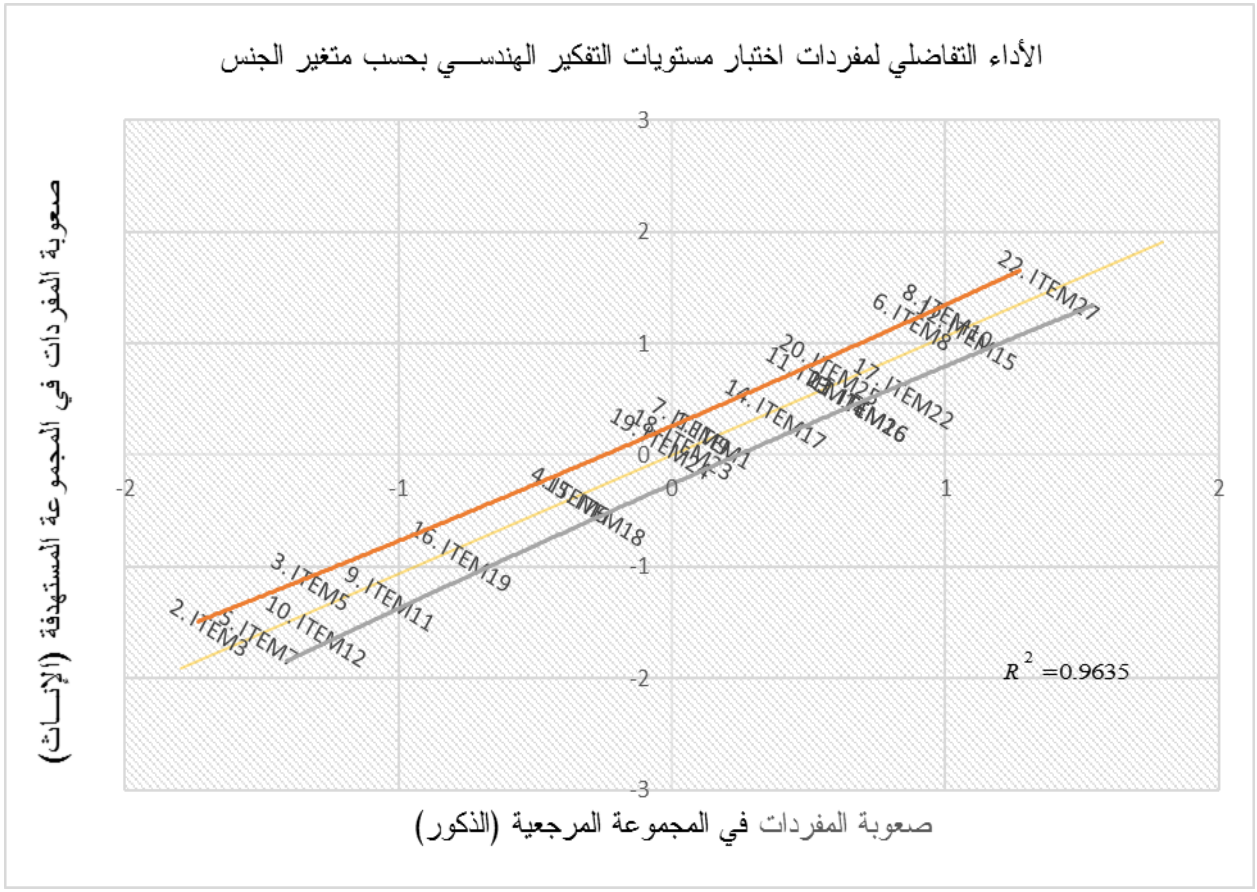
المفردات	صعوبة المفردة للمجموعة المرجعية الذكور	صعوبة المفردة للمجموعة المستهدفة الإناث	الفرق بين الصعوبتين حجم DIF	راش – ويلش Rasch-welch			مانتل – هانزل Mantel-Haensel		
				قيم إختبار "ت"	درجة الحرية d.f	مستوى الدلالة P	قيم اختبار كا^2	مستوى الدلالة P	النسبة الأرجحية لمانتل هانزل
1	0.15	0.15	0.00	0.00	508	1.00	.011	.92	0.00
3	1.70–	1.54–	0.16–	0.67–	509	.50	.143	.71	0.14–
5	1.33–	1.13–	0.20–	0.88–	492	.38	1.30	.25	0.33–
6	0.36–	0.36–	0.00	0.00	518	1.00	.00	.99	0.03–
7	1.51–	1.63–	0.12	0.48	527	.63	.01	.92	0.01
8	0.89	1.19	0.31–	1.52–	481	.13	.33	.57	0.15–
9	0.07	0.26	0.19–	1.01–	507	.31	.28	.60	0.13–

10	1.02	1.25	0.23–	1.14–	487	.26	.16	.69	0.11–
11	1.03–	1.28–	0.24	1.13	535	.26	1.26	.26	0.33
12	1.31–	1.57–	0.26	1.13	536	.26	.90	.34	0.29
14	0.53	0.66	0.12–	0.67–	510	.51	.00	.97	0.03
15	1.08	1.08	0.00	0.00	503	1.00	.02	.90	0.06–
16	0.68	0.44	0.24	1.28	503	.20	1.82	.18	0.31
17	0.37	0.37	0.00	0.00	498	1.00	.00	.99	0.02
18	0.29–	0.49–	0.21	1.08	513	.28	.68	.41	0.20
19	0.77–	0.90–	0.13	0.65	520	.52	.12	.73	0.11
22	0.85	0.56	0.29	1.54	494	.12	.00	.96	0.03
23	0.04	0.10	0.06–	0.3–	506	.76	.06	.80	0.07–
24	0.04–	0.11	0.15–	0.83–	512	.41	.00	.99	0.02–
25	0.58	0.77	0.19–	1.01–	503	.31	3.45	.06	0.4–
26	0.69	0.43	0.25	1.35	498	.18	.94	.33	0.22
27	1.38	1.52	0.14–	0.66–	499	.51	.00	.99	0.02–

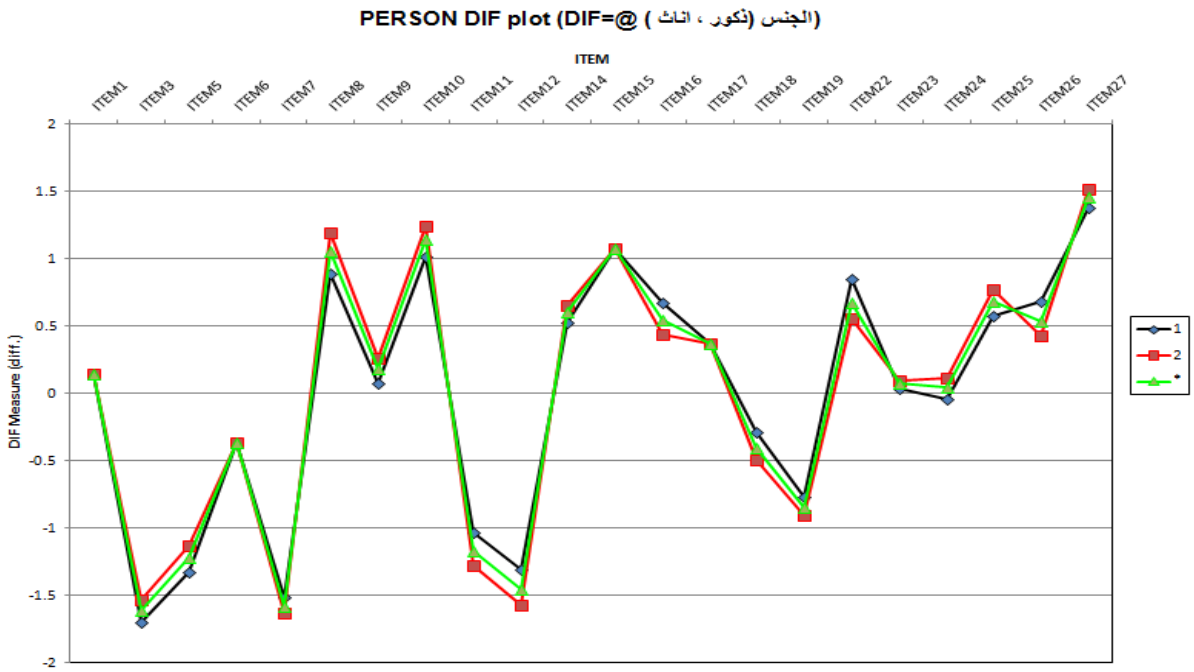
يتضح من نتائج التحليل في الجدول رقم (50)، المؤشرات الإحصائية للاداء التفاضلي للمفرادت التي تعتمد على النظرية الكلاسيكية في القياس من خلال طريقتي قيم الإحصائي اللامعلمي مربع كاي لـ مانتل– هانزل، ومؤشر لوغريتم نسبة الارجحية التركمية لـ مانتل – هانزل (MH-LOR) باستخدام طريقة الانحدار اللوجستي، والمؤشرات الإحصائية التي تعتمد على نظرية الاستجابة للمفردة الاختبارية، من خلال قيم الإحصائي المعلمي اختبار (ت) لـ راش– ويلش (Rasch-welch). وكذلك قيم حجم التباين (DIF contrast) والذي يعتمد على الفرق بين صعوبة المفردة في المجموعتين، أي المجموعة المرجعية (الذكور)، والمجموعة المستهدفة أو البؤرية (الإناث).

وللحكم على أن المفردة تبدي أداءً تفاضلياً لمتغير الجنس، إذا كان مستوى الدلالة الإحصائية المقابل لها أقل من أو يساوي 0.05 بالنسبة لـ اختبار (كا^2) كاي مربع لـ(مانتل– هانزل)، كما يمكن الحكم كذلك على الأداء التفاضلي للمفردة من خلال اختبار (ت) لـ(راش– ويلش) بمستوى الدلالة الإحصائية المقابل لها أقل من أو يساوي 0.05. ويتضح من الجدول (50)، بأن قيم حجم الفرق بين صعوبة المفردة في المجموعتين (المرجعية و المستهدفة) لمفردات الاختبار في صورته النهائية بعد التدريج 22 مفردة، بأنها قيم أقل من القيمة المحددة لـ (DIF) و المحصورة ما بين –0.5 و 0.5 لوغريتم. كما أن قيم اختبار (ت)

محصورة ما بين –2 و2 لوغاريتم. و قيمة مستوى الدلالة المحصورة ما بين –0.05 و0.05. كما أن قيم اختبار (كا^2) كاي مربع لـ (مانتل– هانزل) ليست دالة إحصائياً عند مستوى الدلالة الإحصائية 0.05 بالنسبة لمفردات الاختبار ككل. واعتماداً كذلك على اختبار (ت) لـ(راش– ويلش)، فإن قيم مفردات الاختبار غير دالة إحصائياً عند مستوى الدلالة الاحصائية 0.05 بالنسبة لمفردات الاختبار ككل. ووفقاً للمؤشرات الإحصائية لمفردات اختبار مستويات التفكير الهندسي بعد تدريجه، والخاصة بأداء المجموعة المرجعية والمجموعة المستهدفة، تبين أن مفردات الاختبار تعمل في الإتجاه المفترض منها، وبالتالي لم يستبعد الباحث أي مفردة في هذه المرحلة من مراحل تصميم الاختبار وفقاً لنموذج راش، وهو ما يوضحه الشكلان رقم (37 و 38).



الشكل (37) لمنحنى بياني يوضح الأداء التفاضلي لاختبار مستويات التفكير الهندسي في صورته النهائية، بالنسبة للاختبار ككل (DTF) للمجموعة المرجعية(الذكور)، والمجموعة المستهدفة (الإناث).



الشكل (38) لمنحنى بياني يوضح الأداء التفاضلي لمفردات اختبار مستويات التفكير الهندسي في صورته النهائية، بحسب المفردات(DIF) للمجموعة المرجعية ذكور(1)، والمجموعة المستهدفة الإناث(2).

يتضح من خلال الشكلين رقم (37 و 38)، أن مفردات اختبار مستويات التفكير الهندسي في صورته النهائية لم تبدي أي شذوذاً غير محتمل باعتبار أنها تقيس نفس التكوين الفرضي بين المجموعتين المرجعية والمستهدفة، وبالتالي فهي تدل على تساوي القدرة فيما يتعلق بالتكوين الفرضي الذي صمم الاختبار لتقييمه. وبناءً على ذلك يمكننا الاحتفاظ بجميع مفردات الاختبار وبكل إطمئنان لأن المفردات لم تبدي أداءً تفاضلياً يعزى لمتغير الجنس (ذكور، إناث).

1-9- عرض وتحليل نتائج الفرضية التاسعة:

والتي تنص على أن:" يبلغ تلامذة المرحلة الثانوية مستويات متقدمة في مستويات (فان هيل) للتفكير الهندسي بعد تطبيق اختبار مستويات التفكير الهندسي في صورته النهائية".

ومن أجل اختبار صحة هذه الفرضية، رصد الباحث توزيع تلامذة المرحلة الثانوية الشعب العلمية على مستويات التفكير الهندسي وفقاً لنموذج (فان هيل)، من بيانات اختبار مستويات التفكير الهندسي المطبق في صورته النهائية بعد تدريجه، تم جمع استجابات كل تلميذ على الخمسة أو الستة مفردات بحسب مكونات كل مستوى من مستويات الاختبار الأربعة، وتصحيحها باستخدام المعايير التالية:

1. يحقق التلميذ المستوى الأول والثالث إذا حصل على 3 إجابات صحيحة من 5 إجابات كحد أدنى، مع الأخذ بعين الاعتبار أن احتمال الحصول على إجابة صحيحة من أربعة خيارات هو 0.25، ويكون احتمال

تخمين ثلاثة إجابات صحيحة –على الأقل– من خمسة مفردات يساوي 0.103[1]، أي 0.103 x 566 وبالتالي هناك ما يقارب 58 تلميذاً، يمكنهم تحقيق أي مستوى بالتخمين الصرف.

2. يحقق التلميذ المستوى الثاني والرابع إذا حصل على 4 إجابات صحيحة من 6 إجابات كحد أدنى، مع الأخذ بعين الاعتبار أن احتمال الحصول على إجابة صحيحة من أربعة خيارات هو 0.25، ويكون احتمال تخمين أربعة إجابات صحيحة –على الأقل– من ستة مفردات يساوي 0.037[2]، أي 0.037 x 566 وبالتالي هناك ما يقارب 21 تلميذاً، يمكنهم تحقيق أي مستوى بالتخمين الصرف.

3. يجب تحقيق المستوى الأول كحد أدنى كي يتم تصنيف التلميذ على مستويات (فان هيل)، وغير ذلك اعتبر التلميذ أنه غير مصنف.

4. يجب تحقيق المستوى السابق للمستوى الذي يجري فحصه، فمثلاً، إذا تحقق المستوى الثالث عند تلميذ ما ولم يحقق المستوى الثاني لا يؤخذ بهذه النتيجة.

ولابد من التنويه إلى أن التصنيف لكل مستوى يعني نسبة التلامذة الذين حققوا هذا المستوى على الأكثر، بمعنى أنهم قد يحققون الأدنى منه (إن وجد) ولكنهم لم يحققوا الأعلى منه. مثلا، التلامذة الذين حققوا المستوى الثاني، يكونون قد حققوا المستوى الأول والمستوى الثاني، ولكنهم لم يحققوا المستوى الثاني أو أكثر. وبعد تصحيح إجابات التلامذة وفقا للمعايير السابقة، تم تصنيف التلامذة على أربعة مستويات للتفكير الهندسي وفقاً لنموذج (فان هيل)، والنتائج يوضحها الجدول رقم (51):

الجدول (51)

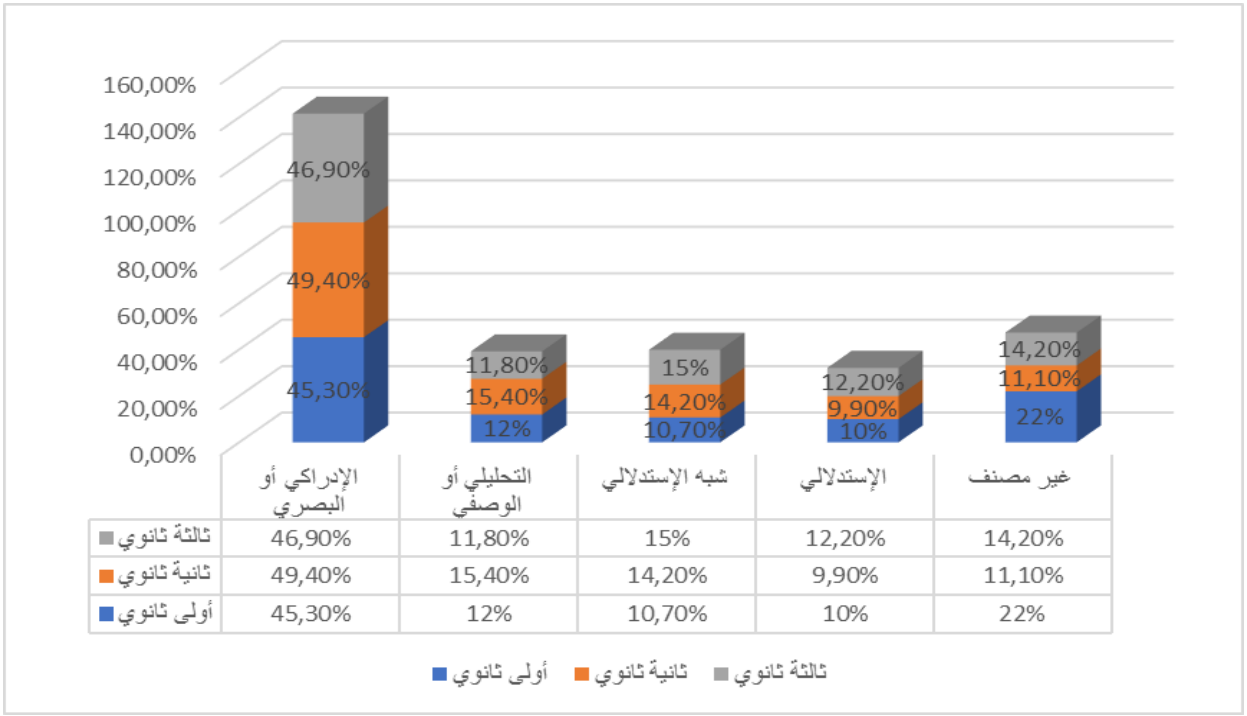
يوضح النسب المئوية لتوزيع تلامذة المرحلة الثانوية (الشعب العلمية) على مستويات (فان هيل) بحسب المستوى الدراسي.

المستوى الدراسي	التلامذة المصنفون على مستويات (فان هيل)				تلامذة لم يصنفوا
	الإدراكي أو البصري	التحليلي أو التصنيفي	شبه الاستدلالي	الاستدلالي	
أولى ثانوي	%45.3	%12	%10.7	%10	%22
ثانية ثانوي	%49.4	%15.4	%14.2	%9.9	%11.1
ثالثة ثانوي	%46.9	%11.8	%15	%12.2	%14.2
المجموع	%47.17	%12.89	%13.6	%10.95	%15.37

[1] - $C_5^3 (0.25)^3(0.75)^2+C_5^4(0.25)^4(0.75)^1+C_5^5(0.25)^5(0.75)^0 = 0.103$

[2] - $C_6^4 (0.25)^4(0.75)^2+C_6^5(0.25)^5(0.75)^1+C_6^6(0.25)^6(0.75)^0 = 0.037$

يتضح من نتائج الجدول رقم (51)، أن ما نسبته 78% من تلامذة السنة الأولى ثانوي؛ و88.9% من تلامذة السنة الثانية ثانوي؛ و85.8% من تلامذة السنة الثالثة ثانوي، أي ما نسبته 84.61% من العينة الكلية أمكن تصنيفهم على مستويات (فان هيل)، بمعنى أن النظرية نجحت في تصنيفهم، وما نسبته 22% من تلامذة السنة الأولى ثانوي؛ و11.10% من تلامذة السنة الثانية ثانوي؛ و14.20% من تلامذة السنة الثالثة ثانوي، لم يتم تصنيفهم بحسب نموذج (فان هيل) لمستويات التفكير الهندسي، وفقاً للمحكات المستخدمة في البحث الحالي، حيث توزع تلامذة السنة الأولى ثانوي على مستويات (فان هيل) الأربعة الأولى، كالتالي: ما نسبته 45.3% في المستوى الأول؛ و12% في المستوى الثاني؛ و10.7% في المستوى الثالث؛ و10% في المستوى الرابع. أما بالنسبة لتلامذة السنة الثانية الثانوي فكان توزيعهم، كالتالي: ما نسبته 49.4% في المستوى الأول؛ و15.4% في المستوى الثاني؛ و14.2% في المستوى الثالث؛ و 9.9% في المستوى الرابع. أما بالنسبة لتلامذة السنة الثالثة الثانوي فكان توزيعهم، كالتالي: ما نسبته 46.9% في المستوى الأول؛ و11.8% في المستوى الثاني؛ و15% في المستوى الثالث؛ و12.2% في المستوى الرابع. و الشكل رقم (39) يوضح توزيع تلامذة المرحلة الثانوية على مستويات (فان هيل) للتفكير الهندسي:



	الإدراكي أو البصري	التحليلي أو الوصفي	شبه الإستدلالي	الإستدلالي	غير مصنف
ثالثة ثانوي	46,90%	11,80%	15%	12,20%	14,20%
ثانية ثانوي	49,40%	15,40%	14,20%	9,90%	11,10%
أولى ثانوي	45,30%	12%	10,70%	10%	22%

الشكل (39) يوضح توزيع التلامذة على مستويات التفكير الهندسي لفان هيل وفقاً للمستوى الدراسي.

الفصـــل التاسع

تفسير و مناقشـــة نتائج البحث

تمهيد:

بعد ما تم عرض وتحليل نتائج فرضيات البحث، قام الباحث من خلال هذا الفصل بتفسير ومناقشة نتائج هذه الفرضيات، بناءً على الجانب النظري والدراسات السابقة في هذا الموضوع، كالتالي:

1- تفسير ومناقشـــة نتائج فرضيـــات البحث:

1-1- تفسير ومناقشـــة نتائج الفرضية الأولـــى:

والتي تنص علـــى أن:" يتوفر اختبار مستويات التفكير الهندســـي على خصائص سيكومترية مقبولة باستخدام النظرية الكلاسيكية في القياس".

تبين من خلال النتائج الموضحة في الجدول رقم (22)، أن قيم مؤشرات صعوبة مفردات اختبار مستويات التفكير الهندسي تراوحت ما بين 0.21 و0.85، وبمتوسط حسابي بلغ 0.53، والتي تدل على أن مفردات الاختبار لها مؤشرات صعوبة مقبولة، وعلى الرغم من أن خبراء القيـــاس يرون بأنه لا يجوز بأي حال من الأحوال محاولة إجراء أي تعديل في مستوى صعوبـــة المفردة في الاختبـــارات مرجعية المحك، أوحذف المفردات السهلة من الاختبار لتوسيع مدى درجات الاختبار، لأن مستوى صعوبة المفردة في الاختبارات مرجعية المحك لا يتأسس على قدرة المفردة على التمييز بين المستوى المرتفع والمستوى المنخفض للمفحوصين – كما في حالة الاختبارات معيارية المرجع – وإنما تتحدد صعوبة المفردة في الاختبار في ضوء طبيعة المهمة أو الواجب المطلوب القيام بها. إلا أن هذا لا يعني التقليل من أهمية أن تتصف المفردات في هذا النوع من الاختبارات بدرجة معقولة من الصعوبة، ذلك أنه كلما زادت درجة الصعوبة في الاختبار المحكي المرجع، كان ذلك مؤشراً على أن التدريس فعال. والمفردات الهامة ليست دائماً متوسطة الصعوبة، قد تكون سهلة أو صعبة، ويمكن أن تكون اختباراً نقياً في مجال خاص، أي يجب أن يكون الاختبار متسقاً داخلياً.

اتفقت هذه النتائـــج جزئياً مع نتائج دراسة (العطاس، 2014) والتي استخدم فيها اختبار مستويات التفكير الهندسي لفان هيل، والتي أشارت إلى أن الاختبار يتمتع بمدى واسع لمؤشرات السهولة والصعوبة تراوحت ما بين 0.12 و0.88، وبمتوسط حسابي بلغ 0.65. ومع دراسة (لينا فؤاد، 2011) والتي تراوحت فيها مؤشرات صعوبات مفردات اختبار مستويات التفكير الهندســـي، الذي تم بناءه في ضوء نموذج (فان هيل)، ما بين 0.31 و 0.79.

وأشارت نتائج التحليل المبينة في الجدول رقم (23)، إلى أن مؤشرات قيم تمييز مفردات اختبار مستويات التفكير الهندسي باستخدام معادلة برينان (1972)، تراوحت مابين القيمة 0.21 و0.61، بمتوسط حسابي قدره 0.41، أي أن جميع المؤشرات ذات قيم موجبة. وهو ما تتطلبه مؤشرات التمييز لمفردات الاختبارات المحكيـــة، وعلى الرغم من أن بعض هذه المفردات ربما تحذف في الاختبارات معياريـــة المرجع بسبب عدم قدرتها على التمييز، إلا أنه في اختبارات الإتقان أو المحكيـــة المرجع تمد المعلم بمعلومات مفيدة عن مدى إتقان تلامذته للمهام التعليميـــة. وبالتالي فمعامل تتميز مفردات الاختبارات مرجعية المحـــك تتميز بإنخفاض معامل تمييزها، ذلك لأن الهدف من هذه الاختبارات ليس قياس الفروق الفردية بين الأفراد، بل قياس وتحديد ما إذا كان المتعلم قد وصل إلى مستوى الإتقـــان في المعرفة أو المهارة. ففي الطرق التقليدية تكون المفـــردات غير المميزة هي: السهلة جداً أو الصعبة جداً، أما بالنسبة للاختبارات مرجعية المحك، فإنه يجب أن تعدل أسس التمييز فلا تستبعد المفردة غير المميزة، إذا كانت تعكس صفة هامة للمحـــك. في حين أن المفردة موجبة التمييز مرغوب فيها في الاختبارات مرجعية المحك، والمفردة سالبة التمييز، يجب أن يتشكك فيها وقد يكتشف بعد فحص وتحليل دقيق عيب في هذه المفردة، و يرى كلٍ من هامبلتون ونوفك (1963) أن الاختبارات مرجعية المحـــك، عكس الاختبارات مرجعية المعيار، لا تبنى بصراحة لترفع تباين درجات الاختبار إلى الحد الأقصى، وإنما تميل الدرجات إلى أن تكون أكثر تجانســـاً، حيث يرى كذلك كلٍ من هامبلتون وسوامنيثان (Hambleton & Swaminathan, 1989)، بأنه إذا كانت العينة متجانسة نسبياً، فإن قيم مؤشرات التمييز تكون أقل من القيم التي نحصل عليها من عينة غير متجانسة.

اتفقت هذه النتائج جزئياً مع نتائج دراسة (لينا فؤاد، 2011)، والتي سجلت فيها قيم لمؤشرات تمميز مفردات اختبار مستويات التفكير الهندسي الذي صممته، ما بين 0.33 و 0.69. ودراسة (كريسليب وتشين، 2001)، والتي تراوحت فيها قيم مؤشرات تمييز مفردات اختبارات ولاية هاوي للقـــدرات الأساسيـــة وباستخدام معامل الارتباط بيسريال النقطي، واعتماد القيمة 0.40 كمحك، ما بين 0.31 و 0.50 وعلى هذا الأساس تم حذف 4 مفردات من الاختبار. ودراسة (العطاس، 2014) التي توصل فيها إلى قيم مؤشرات تمييز مفردات اختبار مستويات التفكير الهندسي لفان هيل، تراوحت ما بين 0.45 و 1 بمتوسط حسابي 0.83.

وتبين من خلال النتائج الموضحة في الجدول رقم (24)، لتقويم ثبـــات اختبار مستويات التفكير الهندســـي، باستخدام أسلوب مؤشر الاتساق الداخلي لمفردات الاختبار، بأن معاملات الارتباط بيرسون لحساب العلاقة بين الدرجة الكلية للاختبار ودرجة كل مفردة من مفردات الاختبار ككل، تراوحت ما بين 0.17 و0.51، و جميعها قيم ذات دلالة إحصائية عند مستوى 0.01، والتي تعد دليلاً على مدى فاعلية مفردات اختبار مستويات التفكير الهندســـي، وبأنها تعمل في إتجاه واحد، في تحديد التكوين الفرضي المقاس الذي ننوي قياسه، أي أنها تقيس سمة كامنة أحادية البعد، والتي يعتبرها البعض مؤشر للثبات.

في حين أشارت النتائج الموضحة في الجدول رقم (25)، إلى أن قيم معامل كودر- ريتشاردسون كانت قيمها منخفضة نوعاً ما في المستويات الأربعة المكونة للاختبار، إلا أن القيمة لهذا المعامل بلغت 0.71 بالنسبة للاختبار ككل، وهي قيمة مقبولـــة لمعامل الثبات. ويفسر الباحث ويرجع إنخفاض معامل الثبات باستخدام كودر- ريتشاردسون في المستويات الأربعة المكونة للاختبار، إلى أن هذا المعامل يعتمد على معامل الارتباط، حيث تزداد قيمه بزيادة تباين الدرجات، في حين أن الاختبارات المحكية يتم تفسير الدرجة فيها بالرجوع إلى مستوى أداء محدد، فيتوقع أن تكون مجموعة التلامذة متجانسة، كما يمكن أن يرد إلى عدد المفردات القليلة المكونة لكل مستوى، فكما نعلم أن زيادة عدد المفردات في الاختبـــار يعني أن الاختبار يقيس السمة المراد قياســـها قطاعاً عريضاً وأكثر اتساعاً، مما يقلل قدر الإمكان من تدخل عوامل المصادفة، وغيرها من العوامل المؤقتة التي تؤثر على الثبات، ويمكن تفسيره كذلك إلى تشتت درجات الاختبار حيث تؤدي زيادة المفردات شديدة الصعوبـــة والمفردات شديدة السهولة إلى إنخفاض في معامل الثبات، وهو ما ينطبق على مفردات اختبار مستويات التفكير الهندسي، ففكرة مستويات (فان هيل) مبنية على أساس أن تكون المفردات السهلة جداً في المستوى الأول (المستوى البصري أو الإدراكي) والذي فيه يتعرف التلميذ على الشكل الهندســـي ككل، والمفردات الصعبة جداً في المستوى الرابع (الاستدلالي) والذي يستخدم فيه التلميذ الاستنتاج والبرهان، وهذه النتيجة تأكيد آخر على البنية الهرمية لنموذج (فان هيل) لمستويات التفكير الهندسي، كما يمكن أن يرجع كذلك هذا الإنخفاض إلى التجانس بين عينة التلامذة المبحوثين، في المستوى التعليمي، و مستوى القدرة، و كذلك في أعمارهم الزمنية، فكما نعلم أن التجانس الشديد لعينة الأفراد التي يحسب بها الثبات من خلال أدائهما يظهر بانخفاض ملموس في

معامل ثبات الاختبار، ذلك أن التباين داخل هذه العينة المتجانسة يكون منخفضاً بقدر لا يسمح بتقدير التباين الحقيقي للاختبار أي ثباتـــه.

و يظهر من خلال النتائج الموضحة في الجدول رقم (25)، إلى أن قيم مؤشر الثبـــات ليفنجستون (1972)، جميعها مؤشرات مقبولة لمعامل ثبات المستويات الأربعة لاختبار مستويات التفكير الهندسي، إلا أن الباحث لاحظ اختلاف هذه القيم باختلاف المحك المستخدم، حيث ارتفعت قيم هذا المؤشر عند استخدام المحك 7/6 أي ما نسبته 86%، مقارنة باستخدام المحك 7/5 أي ما نسبته 71%، وهذا ما أكدت عليه عديد البحوث والدراسات التي حاولت التحقق من تأثير درجة المحك على قيمة مؤشر ليفنجستون في ثبات الاختبار المحكـــي.

اتفقت هذه النتائج مع توصلت إليه دراسة يوسيسكين (1982)، حيث أن من أحد أوجه الإنتقادات التي وجهت إلى اختبار (فان هيل) لمستويات التفكير الهندســـي المطور بمشروع جامعة شيكاغو، هو إنخفاض قيم معامل الثبات في مستوياته الخمسة المكونة له (Usiskin, 1982; Grawley, 1990 ; Wilson, 1990; Teppo, 1991)، وبنفس الأسلوب الذي استخدمه يوسيسكين استخدمنا في بحثنا الحالي معادلـــة سبيرمان و براون لتصحيح الطول، ووفقا لهذه المعادلة تراوحت قيم معاملات الثبات للمستويات الأربعة ما بين 0.60 و0.77، كما يوضحه الجدول رقم (25)، وهي قيم مقبولة للثبات.

اتفقت هذه النتائج جزئياً مع ما توصلت إليه دراسة الشويخ (2005) والتي استخدم فيها أسلوب معامل الثبات ألفا كرونباخ، والذي بلغت قيمه بحسب المستويات الأربعة الأولـــى لاختبار مستويات التفكير الهندســـي لفان هيل، كالتالي: 0.40؛ 0.09؛ 0.31؛ 0.23، و باستخدم أسلوب يوسيسكين لو كان لكل مستوى 25 مفردة، من خلال معادلة سبيرمان- وبراون والذي بلغت قيمه، كالتالي: 0.77؛ 0.33؛ 0.69؛ 0.60 على التوالي. ومع دراسة (مخلوف، 1994) الذي استخدم فيها اختبار لمستويات التفكير الهندسي وفق نموذج (فان هيل) من تصميمه، و اعتماد معادلة كودر- ريتشارسون21، حيث بلغت قيم الثبات في المستويات الخمسة: 0.11؛ 0.26؛ 0.12؛ 0.66؛ 0.76 على التوالي و 0.62 للاختبار ككل، كما استخدم مؤشر ليفنجستون والذي بلغ في المستويات الخمســـة القيم: 0.97؛ 0.95؛ 0.89؛ 0.91؛ 0.85 على التوالي والقيمة 0.97 للاختبار ككل. في حين توصلت دراسة (ماكبريد، 1995) وباستخدام اختبار (فان هيل) لمستويات التفكير الهندسي المعدل، إلى قيم الثبات بأسلوب معادلة كودر- ريتشاردسون20: 0.26؛

0.55؛ 0.47 للمستويات الثلاثة المكونة للاختبار على التوالي، والقيمة 0.68 للاختبار ككل، كما تراوحت القيم بأسلوب معادلة جتمان ما بين 0.80 و0.97، والقيمة 0.83 بثبات الاعادة بمعامل الارتباط بيرسون. وتوصلت دراسة (عفانة، 2002) والتي استخدم فيها اختبار (فان هيل) لمستويات التفكير الهندسي ومعادلة كودر- ريتشاردسون21 والذي بلغت قيمه في المستويات الأربعة الأولى ما يلي: 0.76؛ 0.86؛ 0.64؛ 0.51 على التوالي والقيمة 0.83 على الاختبار ككل. أما دراسة (رباب، 2008) فقد توصلت وباستخدام اختبار لمستويات التفكير الهندسي وفق نموذج (فان هيل) من تصميمها، إلى قيم معامل الثبات بأسلوب معادلة كودر- ريتشارسون21 للاختبار ككل القيمة 0.89، والقيمة 0.88 بأسلوب التجزئة النصفية. كما توصلت دراسة (لينا فؤاد، 2011) وباستخدام اختبار لمستويات التفكير الهندسي المبني وفق نموذج (فان هيل) من تصميمها، إلى قيم ثبات بمعادلة ألفا كرونباخ على المستويات الأربعة الأولى، كالتالي: 0.90؛ 0.87؛ 0.84؛ 0.89 على التوالي، والقيمة 0.91 للاختبار ككل. والتي لم تختلف كثيراً على ما توصلت إليه دراسة (مدحت نوري، 2015) من نتائج، والذي استخدم فيها اختبار لمستويات التفكير الهندسي من تصميمه، حيث بلغت قيم معامل الثبات بالتجزئة النصفية للمستويات الأربعة، التالي: 0.90؛ 0.86؛ 0.88؛ 0.82 على التوالي، و القيمة 0.87 للاختبار ككل.

وبالرغم من اعتبار الثبات شرط ضروري عند تصميم الاختبار، إلا أنه يبقى غير كافي، لتحقيق الصدق في قياساتنا، بالرغم من أنه يجب أن يتحلى بدرجة عالية من الثبات. لهذا، يجب أن نهتم ليس فقط بالاتساق في قياستنا، ولكن أيضاً بصدقها فالحصول على بينة عن الصدق هو جزء من عملية القياس.

بينت النتائج الموضحة في الجدول رقم (13)، والمتعلقة بالتحقق من الصدق الوصفي (المحتوى) لاختبار مستويات التفكير الهندسي بعد اتمام إجراءات بناءه وفقاً لنظرية (فان هيل) للتفكير الهندسي، و عرضه على مجموعة المحكمين خبراء المادة (الرياضيات)، إلى درجة مقبولة من الصدق الوصفي، باستخدام معادلة لاوش (1975) والتي بلغت قيمتها 0.74 وهي قيمة جيدة ومقبولة ودالة إحصائياً عند مستوى 0.05 حسب (لاوش)، حيث يمثل صدق المحتوى أهمية خاصة في التحقق من الاختبارات التكوينية التي تهتم بالإتقان في واحدة أو أكثر من الأهداف التعليمية النوعية أو الخاصة، وكذلك الاختبارات التحصيلية محكية المرجع.

في حين كشفت نتائج استخدام المعدل السنوي للتحصيل الدراسي في مادة الرياضيات كمحك للتحقق من الصدق التنبؤى لاختبار مستويات التفكير الهندســـي، والذي بلغ قيمته 0.26 وهي قيمة منخفضة نوعاً ما لمعامل الارتباط بيرسون ولكنها قيمة موجبة ودالـــة إحصائياً عند مستوى دلالة إحصائية 0.01، وهي قيمة بالإتجاه المتوقع منها منطقياً بحكم أن اختبار مستويات التفكير الهندســـي المبني، لا يقيس سواء جانباً من جوانب مادة الرياضيات والتي تحتوي اضافة للهندسة على الجبر والحساب والقياس والإحصاء.

وفي هذا السياق يرى علام (2000، ص.197)، أنه لا ينبغي توقع الحصول على قيم مرتفعة لمعامل الارتباط بين درجات الاختبارين التنبؤى والمحك (معدل مادة الرياضيات)، وذلك لتعقد وتشابك السلوك الإنساني.

كما كشفت النتائج الموضحة في الجدولين رقم (26 و27)، للتحقق من انتقاء النطاق السلوكـــي أو صدق التكوين الفرضي من خلال حساب معامل الارتباط بين درجة كل مفردة والدرجة الكلية لكل مستوى من مستويات (فان هيل) للتفكير الهندســـي بعد حذف درجة المفردة من البعد، إلى درجة مقبولة من التجانس أو الاتساق الداخلـــي للمفردات المكونة للاختبار. ويعد هذا دليلاً آخر على مدى فاعلية مفردات اختبار مستويات التفكير الهندســـي، في قياسها ما يقيسه البعد أو المستوى الذي تقع فيه، أي أن هناك إنسجاماً كبيراً بين ما تقيسه المفردة وما يقيسه البعد أو المستوى. كما اشارت نتائج قيم معامل الارتباط بيرسون بين الدرجة الكلية لكل مستوى من مستويات التفكير الهندسي الأربعة التي يقيسها الاختبار ودرجة الاختبار ككل، والتي تراوحت ما بين 0.64 و0.74، إلى درجة مقبولة من التجانس أو الاتساق الداخلي للمستويات الأربعة المكونة للاختبار. في حين أن قيم معاملات الارتباط بين المستويات الأربعة نفسها كانت ضعيفة بالرغم من دلالتها الإحصائية، وهذا دليل على أن الاختبار يقيس فعلاً أربعة مستويـــات متمايزة حسب عملية بنائـــه وفقاً لنظرية مستويات التفكير الهندســـي لفان هيل.

اتفقت هذه النتائج جزئياً مع ما توصلت إليه دراسة (مخلوف، 1994)، بعد تطبيق اختبار مستويات التفكير الهندســـي المصمم وفق نموذج (فان هيل)، من خلال قيم معامـــلات ارتباط بيرسون بين الدرجة الكلية لكل مستوى والدرجة الكلية للاختبار، والتي تراوحت ما بين 0.32 و 0.96. ومع دراسة (عفانة، 2002) التي توصلت باستخدام اختبار مستويات التفكير الهندسي لفان هيل لمعاملات ارتباط بيرسون بين درجات كل مستوى والدرجة الكلية للاختبار، ما بين 0.36 و 0.72 . ومع دراسة (لينا فؤاد، 2011) التي

توصلت من خلال تطبيق اختبار مستويات التفكير الهندسي مصمم وفق نموذج (فان هيل) لمعاملات الارتباط الثنائي النقطي ما بين 0.32 و 0.79 بالنسبة للدرجة الكلية للاختبار ودرجة كل مفردة، و ما بين 0.75 و0.90، بمتوسط حسابي بلغ 0.83 بالنسبة للدرجة الكلية للاختبار والدرجة الكلية لكل مستوى أو مجال. ومع دراسة (العطاس، 2014) التي توصلت بعد تطبيق اختبار مستويات التفكير الهندسي لمعاملات ارتباط بيرسون ما بين 0.44 و0.67 بالنسبة لمفردات كل مستوى والدرجة الكلية للمستوى الذي تنتمي إليه المفردة. ومع دراسة (مدحت، 2015) والتي توصل فيها باستخدام اختبار لمستويات التفكير الهندسي من تصميمه في ضوء نموذج (فان هيل) إلى قيم معاملات ارتباط بيرسون بين درجة كل مفردة في كل مجال والدرجة الكلية للمجال التي تنتمي إليه، ما بين 0.30-0.42.

وعليه وبناءً على ماسبق، يمكننا القول أن هذه الفرضية محققة، والتي تشير إلى أن اختبار مستويات التفكير الهندسي يتوفر على خصائص سيكومترية مقبولة باستخدام النظرية الكلاسيكية في القياس. بالرغم من اعتمادها في الحكم على هذه الخصائص على الدرجة الكلية للاختبار وليس على كل مفردة على حدة، لهذا لا يمكننا إضافة أو حذف أو تعديل مفردات إلى الاختبار مستقبلاً دون أن تتأثر الخصائص السيكومترية للاختبار ككل. وهذا الإشكال تجيب عليه نظرية الاستجابة للمفردة ونموذج (راش).

1-2- تفسير ومناقشة نتائج الفرضية الثانية:

والتي تنص على أن:" تتوفر لبيانات اختبار مستويات التفكير الهندسي المستمدة من استجابات العينة على ملاءمة مقبولة لنموذج راش".

أشارت نتائج التحليل والتي يمكن تلخيصها في الشكل رقم (24)، والمتعلقة بنتائج التحقق من افتراض أحادية البعد لاختبار مستويات التفكير الهندسي، وباستخدام التحليل العاملي الاستكشافي، من خلال إعتماد أسلوب التحليل الموازي، وتقنية العامل المتسارع (Acceleratiion Factor) لكلٍ من رايش وروبيل وبليس (Raiche, Roipel,and Blais, 2006)، إلى أن الاختبار يظهر عامل واحد متسارع والذي يدل على أحادية البعد للاختبار. بالرغم من أن نتائج عديد الأبحاث والدراسات أشارت إلى أن نماذج الاستجابة للمفردة متينة نسبياً لانتهاك فرضية أحادية البعد (Lau, 1996 ; Spray, Addel-fattah, Huamg and Lau, 1997)، وأن فرضية أحادية البعد تتخفف عادة في الاختبارات محكية المرجع. ويتبين كذلك من خلال نتائج الجدول رقم (28) والشكل رقم (25)، إلى أن المؤشرات الإحصائية المستخرجة باستخدام

نموذج راش مختلفة عن الإحصاءات المستخرجة بالطرق الكلاسيكية، لأن البعد في الإحصاء الكلاسيكي يعمل على الدرجات الخام، بينما تعمل نظرية الاستجابة للمفردة باستخدام أسلوب نموذج (راش) للمكونات الأساسية المعتمدة على تباين البواقي (PCAR) لقدرات الأفراد. وبعد تفحص الباحث للمحكات المعتمدة في التحليل تبين أن المحكات تحققت كلياً، وهو ما يؤكد على أن اختبار مستويات التفكير الهندسي له تكوين فرضي ذو فضاء كامن أحادي البعد.

وتعني أحادية البعد للاختبار، بأن العلامات تتنبأ بالأداء على المفردات، ومن ثم يكون من المنطقي تلخيص الأداء على الاختبار بعلامة منفردة (Kame, 2008). أما إذا لم يكن الاختبار أحادي البعد، فيعني أن هناك معلومات إضافية تتعلق بهذا الاختبار الفرعي لم يتم عكسها في العلامة الكلية، ومن ثم فعلامته لا تتنبأ بالأداء على المفردات أو مجموعات فرعية منها (Schilling, 2007).

ويرى وليد(2014) في هذا السياق نقلاً عن أمينة كاظم (1994، ص.119)، بأن أحادية القياس لا تعني بساطة ما تقيسه المفردة، وإنما تعني أن مفردات المقياس تكون متجانسة فيما بينها وتقيس في أساسها نفس الصفة. ومعنى هذا أن أي مفردة من هذه المفردات المتدرجة الصعوبة تتطلب في حلها نفس النوع من الإجراءات والعمليات السلوكية، ولكنها تختلف فيما بينها من حيث تدرج صعوبتها فقط. في حين يرى لورد (Lord)، أن افتراض أحادية البعد في الاختبارات العقلية والمقاييس الوجدانية لا يعد افتراضاً صحيحاً كلية، وإنما يمكن أن يؤدي إلى نتائج تقريبية؛ فالسؤال الذي يتم طرحه هو هل أن هناك أبعاد أخرى في البيانات تستوجب شطر مفردات الاختبار إلى بعدين مستقلين أو أكثر، ومن ثم إجراء تحليل مستقل لكل من تلك المجموعات أم لا؟

دعمت نتائج التحقق من افتراض تساوي التمييز لمفردات الاختبار، و التي يوضحها الجدول رقم (29)، ما يفترضه نموذج (راش) بأن جميع المفردات تتساوى في قدرتها على التمييز. حيث يرى علام (2000، ص.696)، أن مفردات الاختبار لا يجب أن تميز بين مستويات القدرة المقاسة تمييزاً دالاً فحسب، بل أن جميعها يجب أن تكون قدرتها على التمييز متساوية تقريباً. فقيم مؤشرات التمييز تشير إلى تحقق افتراض دالة استجابة المفردة، ويقصد بذلك تساوي الميل النسبي للمنحنيات المميزة للمفردات، وبالتالي تساوي قوة التمييز، وأفضل المفردات هي التي تكون زاوية الميل النسبي للمنحنى المميز للمفردة فيها يساوي $\theta = 45$، أي أن معامل تمييزها يساوي الواحد أو يحوم حول هذه القيمة (كاظم، 1988).

فالمنحنيات المميزة لخصائص المفردة تعتبر من الافتراضات الأساسية التي تقوم عليها نظرية الاستجابة للمفردة التي تتناول سمة أحادية البعد، وهي دالـــة رياضية تربط بين احتمال نجاح الفرد في إجابة المفردة وبين السمة أو القدرة التي تقيسها مجموعة المفردات، أو يقيسها اختبار يحوي هذه المفردة، أي هو دالـــة غير خطية لانحدار درجة المفردة على السمة الكامنة، أو القدرة التي يقيسها الاختبار، والاختلاف الرئيس بين نماذج الاستجابة للمفردة يعتمد على الصيغة الرياضية لمنحنى خاصية المفردة (Hambleton, 1989, p.151)، وهو ما يوضحه الشكل رقم (26)، بأن الدرجة على الاختبـــار تتزايد بزيادة السمـــة الكامنة، وبالتالي تزداد احتمالية التوصل إلى الاستجابة الصحيحة، وتكمن أهمية منحنيات خاصية المفردة في اظهار التباين في معامل صعوبـــة وتمييز المفردة، والتي تسمح لنا أن نرى كيف أن احتماليـــة الإجابة الصحيحة تعتمد على السمة الكامنـــة (Crocker and Algina, 1986, p.340). و يرى لينكر أن متوسط قيم تمييز المفردات والتي تتمثل في الميل (Slope) في نموذج (راش) يجب أن تكون قريبة من الواحد تقريباً (Linacre, 2012, p.597). وهو ما تشير إليه نتائج الجدول رقم (29) لمتوسط قيم مؤشرات التمييـــز الذي بلغ 0.98 وهي قيمة قريبة جداً من الواحد. والتي توضحه الرسوم البيانيـــة في الشكل رقم (26) لمنحنى احتمال التوصل إلى الاجابة الصحيحة على المفردة كدالـــة للسمة الكامنة في ضوء الأداء على مفردات الاختبار، وفي تطبيقات نظرية الاستجابـــة للمفردة يفترض أن منحنيات خصائص المفردة لها شكل حرف (S)، كما أن احتمال الاجابة الصحيحة يفترض في نماذج النظرية الكلاسيكية للقياس، بأن يكون على شكل المنحنى الاعتدالـــي في الاختبارات جماعية المرجع، وشكلاً ملتوياً التواءً سالباً في الاختبارات محكية المرجع، أما في نماذج نظرية الاستجابـــة للمفردة فيفترض أنها تتخذ شكل المنحنى اللوغاريتمـــي.

ومن المؤشرات الأخرى التي تم مراعاتها و التي تدل على تحقق تكافؤ مؤشرات التمييز لمفردات الاختبار ومطابقتها للنموذج، ما أوصى به هامبلتون وسوامسيشان (Hambelton and Swuminathan, 1985) وهو أن تكون قيم هذه المؤشرات واقعة ضمن حدود مدى متوسط مؤشرات التمييز يساوي 0.15، وبما أننا وجدنا قيمة متوسط معاملات التمييز 0.32 فإن حدود المدى تتراوح ما بين 0.18 و0.50. كما تدل عليه نتائج جدول رقم (34)، والتي توضح قيم معاملات الارتباط الثنائي المتسلسل أو الأصيل (ρ_{pbis})، والتي كانت جميعها ضمن المدى المقبول، وأن توزيعها كان إلى حد ما متجانساً وبالتالي تساوي مؤشرات

التمييز، مع تسجيل قيم صغيرة للانحراف المعياري لهذه المؤشرات، وهو مؤشر آخر على تحقق هذا الافتراض.

و أشارت نتائج التحليل الموضحة في الجدول رقم(30)، إلى أن قيم معاملات ارتباط البواقي المعيارية باستخدام المؤشر الإحصائي Q3، للتحقق من افتراض الاستقلال المحلـــي، و على الرغم من أن الاستقلال الموضعي يعد افتراضاً ربما يصعب تحققه في كثير من الاختبـــارات، وبخاصة الاختبارات العقلية نظراً لتداخل هذا الافتراض مع أحادية البعد. فالمقصود باستقلال موضع المفردة، هو أن تكون استجابات الفرد للمفردات في الاختبـــار مستقلة استقلالاً إحصائياً. حيث تراوحت قيمه ما بين −0.11 و−0.17 بمتوسط حسابي بلغ −0.12، وعموما فإن هذه القيم لم تتجاوز مستوى الإنتهاك 0.30 حسب زينسكي وهامبلتون وسيريسي (Zenisky and Hammbleton and Sireci, 2002)، وكذلك المجال الذي اقترحه (Chen and Thissen, 1997) وهو أن تكون القيم ما بين 0.20 و−0.20. وفي ضوء هذه النتائج، يمكننا الحكم على أن مفردات اختبار مستويات التفكير الهندســـي لا توجد فيها مفردات تتأثر الإجابة عنها بمفردات أخرى من مفردات نفس الاختبار، مما يعني عدم تداخل منحنيات الخصائص المميزة للمفردات، وهو ما يشار إليه بالاستقلال الموضعـــي، وبالتالي فإن هذا الافتراض متحقق.

ويتضح للباحث من خلال النتائج الموضحة في الجدول رقم (31)، إلى تحقق افتراض تدنـــي عامل التخمين، حيث أشارت جميع قيم مؤشرات التخمين الأدنى لمفردات اختبار مستويات التفكير الهندسي بأنها أقل من القيمة 0.10، ما عدا المفردة 4 والتي بلغت القيمة 0.12، وهي على العموم قيم منخفضة وقريبة من الصفر وفي المجال المقبول لقيم التخميـــن الذي حدده كل من (Linacre, 2012, p.112) (Steven, 2003, pp.164-184) (وليد، 2014، ص.68) (Baker, 2001, p.28).و بالنظر إلى الشكل رقم (26) والذي يظهر المنحنيات المميزة لخصائص مفردات الاختبـــار (خطوط الإنحـــدار غير الخطية للدرجات على الاختبار)، إلى أن الخط التقاربي الأدنـــى اقترب من الصفر عند المفردات جميعها، وهذا يحدث عندما لا يلجأ الفرد إلى الإجابة عشوائياً على الاختبـــار، وبذلك يمكن استخلاص خلو الاختبـــار تقريباً من عنصر التخميـــن، مما يعني أن مفرداته غطت معظم موضوعات المنهج (الهندسة)، و تخدم كل من بذل جهداً في أي جزء من أجزائه، وبالتالي يمكن القول أن هذا الافتراض من افتراضات شروط تطبيق نموذج (راش) في التدريج قد تحقق.

تحقق الباحث كذلك من افتراض عامل السرعة، بعد أن تأكد عملياً بعد تطبيق الاختبار من أن فشل التلامذة في الوصول إلى الإجابة الصحيحة عن مفردات الاختبار، يعود فقط لإنخفاض قـــدراتهم وليس لعامل السرعة.

لم تسفر نتائج التحليل في هذه المرحلة من مراحل تدريج الاختبار عن استبعاد أي مفردة من مفردات الاختبار 28، وهي مرحلة التحقق من مدى توفر بيـــانات الاختبـــار لافتراضات نموذج (راش) أحـــادي البارامتر.

لم تختلف هذه النتائج إلى ما توصلت إليه دراسة كلٍ من (ويلسون، 1990؛ سيرو، 2007؛ مايكل كلرن، 2012؛ ستولس، 2015؛ شونغ ونور، 2015؛ سيف خليل، 2017)، والذين استخدموا اختبارات لمستويات التفكير الهندســـي لفان هيل في دراساتهم سواء الاختبار الأصلي أو اختبارات تم بنائها.

وعليه وبناء على ماسبق من نتائج، يمكننا القول أن الفرضية محققة، والتي تشير إلى أنه تتوفر لبيانات اختبار مستويات التفكير الهندســـي المستمدة من استجابات تلامذة المرحلة الثانويـــة على ملاءمة مقبولة لنموذج (راش). وبالتالـــي أمكن للباحث وبكل اطمئنان أن يقوم بعملية تدريج الاختبـــار على أفراد عينة البحث وفقاً لنموذج (راش) أحادي البارامتر.

1-3- تفسير ومناقشـــة نتائج الفرضية الثالثـــة:

والتي تنص علـــى أن:" تتباين تقديرات صعوبة مفردات اختبار مستويات التفكير الهندسي وتدرجها باستخدام نموذج راش".

تبين من خلال تحليل نتائج عملية تدريج أو معايرة مفردات الاختبـــار والموضحة في الجدول رقم(32) والذي يمثل ملخص نتائج التحليل لقدرات الأفراد، والجدول رقم(33) الذي يوضح ملخص لنتائج التحليل لمعالم الصعوبة لمفردات اختبار مستويات التفكير الهندســـي قبل التدريج، والجدول رقم(34) الذي يوضح المفردات غير المطابقة، وقيم إحصائيات المطابقـــة الداخلية والخارجية، ومتوسطات المربعات الموزونة لمفردات اختبار مستويات التفكير الهندســـي في التدريج الأولى. والجدول رقم(35) الذي يلخص نتائج التحليل للقيم المتحررة من قدرات الأفراد بعد استبعـــاد الأفراد والمفردات غير المطابقة. والجدول رقم(36) الذي يوضح ملخص لنتائج التحليل المتحرر من صعوبـــة المفردات بعد التدريج. والجدول رقم(37) الذي يوضح صعوبـــة المفردات، وأخطاؤها المعيارية المرتبة تنازلياً وفق صعوبة المفردات.

والشكل رقم(27) الذي يوضح مؤشرات المطابقة للأفراد وحدود المطابقة ومواقع الأفراد من هذه الحدود وفقاً لإحصائي متوسط المربعات التباعدي والتقاربـــي (MNSQ)، والذي تراوحت قيمه ما بين 0.84 و 1.13، بمتوسط حسابـــي 1، وانحراف معيـــاري 0.04 بالنسبة لمؤشرات الملاءمة الإحصائية لمتوسط المربعات التقاربية (الداخلية)، وبمتوسط حسابـــي 0.99، وانحراف معيـــاري 0.08 بالنسبة لمؤشرات الملاءمة الإحصائية لمتوسط المربعات التباعدية (الخارجية). وهي قيم في مجال القياس المنتج الذي حدده كلٍ من (Wright and Linacre, 1994) بالقيم ما بين 0.70 و1.30 أو ما بين 0.50 و1.5 عموماً والقريبة من الواحد.

ويرى رايت وآخرون (Wright, Mead and Bell, 1980)، كما ورد عن كاظم (1988أ) أنه يجب أن تتوزع قيم (ت) أو الإحصائي (ZSTD)، للمفردات الملائمة إعتداليـــاً، بمتوسط قدره (صفر)، وانحراف معياري قدره (واحد). ومن الملاحظ أنه قد تنخفض قيمة المتوسط الحسابي إلى 0.5، وانحراف معياري قدره 0.6. وهو ما تحقق تقريباً في بحثنا، حيث بينت قيم إحصاءات المطابقة التقاربية (الداخلية) (Infit-ZSTD) للمتوسط الحسابـــي 0.1، وانحراف معياري 1.1، كما بينت قيم إحصاءات المطابقة التباعدية (الخارجية) (Outfit- ZSTD) للمتوسط الحسابـــي 0.1، وانحراف معياري 1.2.

وتدل هذه المؤشرات على اقتراب المنحنى الملاحظ من المنحنى المتوقع (أفضل منحنى له يطابق النموذج)، وبذلك يتوافر لهذا المنحنـــى ما يتوافر للمتوقع من النموذج من استقلال لمعلمة الصعوبـــة للمفردات عن خصائص توزيع القـــدرة لعينة التلامذة. كما تدل هذه المؤشرات فيما إذا كانت منحنيات الخصائص للمفردات المطابقة للنموذج ذات ميل أو انحناء متشابه، فعندما تكون معالم المفردات مستقلة عن العينة، فإن قدرة المفردات على التمييز تكون متساوية نسبياً وبذلك تكون مفردات ذات إنحناء متماثل وهذا ما يتضح من خلال الشكل رقم(28) الذي يظهر منحنيات خصائص مفردات اختبار مستويات التفكير الهندســـي بعد التدريج باستخدام نموذج راش. وبأن المنحنيات المميزة لمفردات الاختبـــار بعد تدريجه باستخدام نموذج (راش) أحادي البارامتر لا تتقاطع ولكن تختلف عن بعضها في الصعوبة فقط وتتلاقى فقط عند المحور الأفقي الذي يمثل متصل القدرة أو السمـــة. حيث ترى كاظم (1988أ) أن تساوي قوة التمييز لمفردات اختبار ما يرتبط بتحقق شرط استقلالية القياس، حيث تتساوى الصعوبة النسبية للمفردات عبر مستويات القدرة، و تعرف قوة تمييز المفردة بأنها معـــدل التغيير الذي يعرف بأنه الميـــل النسبـــي

للمنحنى المميز للمفردة على محور القدرة، وتقل قــوة المفردة على التمييز بين الأفراد على مستويــات القدرة المختلفة أو تزيد تبعاً لانحناء المنحنى المميز للمفردة.

وتتفق كل من النظرية الكلاسيكية في القيــاس ونظرية الاستجابة للمفردة ومنها نموذج راش على فرضية مؤداها وجود متصل للقدرة وإذا علم موقع الفرد على هذا المتصل، فإنه يمكن تقدير احتمال إجابة فرد ما إجابة صحيحة على مفردة اختبارية تقيس هذه القدرة، إلا أن كلا منهما يختلفان في الدالة التي يتم بواسطتها تحديد موقع الفرد وارتباطه باحتمال الإجابــة الصحيحة، حيث أن نماذج النظرية الكلاسيكيــة تفترض أن شكل الدالــة يتخذ شكل منحنى التوزيع الاعتدالي في الاختبارات جماعية المرجع، وشكلاً ملتوياً إلتواء سالباً في الاختبارات محكية المرجع. في حين أن نماذج نظرية الاستجابة للمفردة تفترض أنها تتخذ شكل المنحنى اللوغاريتمي، وبالتالي وجود دالة مميزة خاصة بكل مفردة من مفردات الاختبــار تعتمد في أساسها على موقع الفرد والمفردة معاً على متصل القدرة وتتخذ شكل المنحنى اللوغاريتمي التراكمي.

وأشارت نتائج دراسة الصدق البنائــي لاختبار مستويات التفكير الهندسي من خلال ما يظهره الشكل رقم(29)، والذي يوضح خريطة رايت لتدريج الأفراد ومفردات اختبار مستويات التفكير الهندسي. والتي تظهر بأن هناك سمة مهمة في الاختبار، وهي القدرة على التفكير الهندســي، وبأن الاختبار يقيس ما يراد قياســه وفق نظرية (فان هيل) في تطور مستويات التفكير الهندســي. ومن العناصر المهمة في الصدق البنائــي هو عنصر التمثيل، ونعني بالتمثيل هو مدى حساسية الاختبار للاختلافات في التكوين الفرضــي الذي يقيسه، حيث أن هناك ثلاث قضايا مهمة في عنصر التمثيل، فالقضية الأولى، تتعلق بما إذا كان عدد مفردات الاختبار كاف لقياس جميع جوانب التفكير الهندســي في ضوء نظرية (فان هيل) والتي يراد قياسها. والقضية الثانية، هي ما إذا كان التسلسل الهرمــي التجريبي للمفردات يظهر انتشاراً كافياً في المستويات الأربعة لمستويات التفكير الهندسي وفق نظرية (فان هيل). أما القضية الثالثة، وهي إمكانية وجود مسافات أو فراغات في التسلسل الهرمي للمفردات التجريبية.

فالناتج النهائي لعملية معايرة الاختبار يمثل تعريف تدريج مقياس القدرة، وفي نموذج راش نجد لهذا المقياس وحدة قياس 1 ونقطة المنتصف هي صفر. و تعتبر إمكانية وضع المفردات والمفحوصين في مقياس عام سمة قويــة لنظرية الاستجابــة للمفردة، بحيث تسمح هذه السمة للفرد بتفسير نتائج معايرة

الاختبار وفقاً لمرجعية واحدة وتعطي معنى لقيم تقديرات المعالم. وهذا ما أشارت إليه خريطة رايت في الشكل رقم (29) وبصورة جلية، إلى وجود تناسب تام بين توزيع تقدير مفردات اختبار مستويات التفكير الهندسي، وتدريج المفردات، مما يشير إلى مصداقية مفردات اختبار مستويات التفكير الهندسي وثباتها، ومناسبتها واستيعابها، وصلاحيتها للقياس والتقويـــم التربوي. وتمت دراسة هذه القضايا من خلال خريطة (رايت) للمفردات والأشخاص، وقد أظهرت هذه الخريطة العلاقات الخطية بين تعيير (راش) لـــ 566 من المختبرين، و 22 مفردة من مفردات اختبار مستويات التفكير الهندسي، والتي أشارت إلى وجود تداخل مناسب بين مستويات قـــدرة الأفراد وصعوبـــة المفردات، حيث توزعت صعوبة المفردات، وقدرات الممتحنين على امتداد لوغاريتميين من −0.2 إلى +0.2 تقريباً، كما أشارت نتائج التدريج إلى وقوع المفردة 27 في أعلى الخريطة وهي أصعب المفردات بـــ1.46 لوجيت، وهي تخص مفردة البرهنة في المستوى الرابع وهو المستوى الاستدلالي، والذي يعتبر أعلى مستويات التفكير الهندسي وفق نظرية (فان هيل)، والذي هو عبارة عن جملتين منطقيتين حول " كيفية إثبات أن الشكل مستطيـــل من خلال أقطاره "، كما أظهرت بعض المفردات أنها خارج نطاق قدرات المختبرين من حيث السهولة في الجانب السفلي من الشكل الهرمي، وهي المفردة 3 والتي تخص " التعرف على المثلث"، والمفردة 7 والتي تخص" التعرف على المثلث المتساوي الأضلاع "، والمفردة 12 والتي تخص " خصائص متوازي الأضلاع "، ويعني هذا أن جميع المختبرين أجابوا إجابـــة صحيحة على هذه المفـــردات في تلك المنطقة السفلى من الاختبار لسهولتها، وبالتالي تعتبر أقل استهدافا لكفاية التفكير الهندســـي التي يراد قياسها. والتي تظهر بأن نمط استجابات التلامذة لم يكن متسق مع توقعاتنا وافتراضات النظرية، حسب ترتيب صعوبة المفردات أو (المستويات) وهذا يعتمد كذلك على صدق استجابات الأفراد. وبالرغم من انتظام تدريج صعوبة المفردات بصورة واضحة من المدى −1.61 إلى 1.46 لوجيت، إلا أنه بدت بعض الفراغات الضيقة عند بعض المستويات الأخرى من المتغير والتي تظهرها خريطة (رايت)، وباستخدام الخطأ المعياري، والفرق بين صعوبـــة كل مفردتين مكونتين للفراغ، أظهرت النتائج أن هناك ثلاثة مسافـــات فارغة لم يتم تغطيتها بمجموع الخطأ المعياري للقياس لصعوبة المفردتين و لكن بفرق ضئيل جداً، وعلى الرغم من ذلك يمكن اعتبار انتظام تدرج المفردات على مدى ميزان القياس، بما يعني تعريف مفردات الاختبار للمتغير الذي يمثل بمدى هذا المتصل وهو القدرة على التفكير الهندسي. مما يدل على جودة تدريج مفردات الاختبار،

وهو ما يمكننا من إضافـــة مفردات أخرى للاختبار مستقبلاً لتغطية هذه الفراغات، ويكون ذلك من خلال إعداد بنك أسئلة لاختبار مستويات التفكير الهندســـي وفقاً لنظرية (فان هيل)، بما يساعد في توفير صور متكافئة من الاختبار عند الاستخدام.

وتضيف كاظم (1988أ) في هذا السياق أن معرفة العلاقة بين صعوبات المفردات وقدرات الأفراد، من الممكن أن تفيد في انتقاء مجموعات من المفردات تكون اختبارات مختلفة الصعوبة، تناسب المستويات العليا أو الوسطى أو الدنيا من القدرة للأفراد.

وللتحقق من القضية الثانية من قضايا الصدق البنائي من خلال ما تظهره مفردات اختبار مستويات التفكير الهندســـي من ترتيب هرمي متسلسل أو أنها منفصلة، وبالرجوع إلى خريطة (رايت) بعد التدريج النهائي لاختبار مستويات التفكير الهندســـي في الشكل رقم(29)، تظهر بأن المفردات المصممة لاختبار كل مستوى من مستويات التفكير الهندسي الأربعة لـــ(فان هيل) بأنها تميل إلى أن تكون متسلسلة، فالزيادة في متوسط عتبة صعوبة المفردات لكل مستوى تدعم الأدلة البحثية الأخرى أن المستويات تشكل التسلسل الهرمي. إلا أن هناك بعض التداخل في عتبات الصعوبة بالنسبة لبعض المفردات والتي تشير إلى أنها ليست منفصلة، وقد يعزى جزئياً لبعض المفردات والتي لم تكن دقيقة في اختبار التفكير الهندســـي المبني وفق نظرية (فان هيل)، بالرغم من أن المستويات الأربعة أظهرت ميلاً بأنها تكون بنية هرمية طبيعياً.

وتشير هذه النتائج إلى عكس ما افترضه (فان هيل)، والذي يرى بأن النمو من خلال المستويات ليس متداخلاً وإنما فجائياً إلا أن بعض الباحثين عارضه في ذلك وأيده آخرون. فقد وجد بيج ودافري (Pegg and Davery, 1998) تبريراً للفصل بين المستويات واعتبراه منطقياً إذا أخذ بالنظرة العريضة والمثالية للنظرية، أما برغر و شاونيسي (Burger and Shaughnessy, 1986) و فيوز (Fuys et al,1988) فوجد أن التلامذة يمكن أن يتحركوا داخل المستويات جيئة وذهاباً. وهو ما تدعمه النتائج الحالية لهذه الفرضية، وبالتالي يعتبر الباحث أن مستويات التفكير الهندســـي بدلا من كونها ذات طبيعة ثابتة ومنفصلة، هي في الواقع أكثر ديناميكية وليست منفصلة، وأن التلامذة تتزايد قدراتهم في عدة مستويات في وقت واحد. إذ أن الطبيعة المعقدة لعمليات التفكير البشري، لاتجعلهم يفكرون بطريقة بسيطة وخطية. وهو ما تدعمه خريطة (رايت) في الشكل رقم (29) في عدم انتظام توزيع المفردات عبر المستويات، وقد يكون هذا راجعاً إلى أن المفردات المصممة غير كافية.

وبشكل عام يمكننا القول أن مفردات اختبار مستويات التفكير الهندســـي التي تم تصميمها تقيس مستويات (فان هيل) للتفكير الهندســـي، إلا أن هناك بعض الشذوذ يستحق مزيداً من التحقق في دراسات وبحوث لاحقة.

ويتضح من خلال النتائج السابقة، أنه أمكن و باستخدام نموذج راش اللوغاريتمي أحادي البارامتر من تدريج مفردات اختبار مستويات التفكير الهندسي و تعريفها لمتغير التفكير الهندسي، وبأن المفردات تتدرج من حيث صعوبتها بحيث تعرف متغيراً واحداً فيما بينها، وهو ما يدل على صدق تدرج قدرات الأفراد على متصل هذا المتغير، مما يعني أن تدرج قدرات الأفراد على المتغير محددة لتقديرات أدائهم على هذا الاختبار.

اتفقت هذه النتائج مع نتائج دراسة كلٍ من ستولس (STOLS, 2015)، و مايبري (Mayberry, 1981, p.135)، و (الشويخ، 2005) والتي بينت أن التلامذة يتأرجحون في تفكيرهم الهندســـي بين مستويين متتاليين أو أكثر، وهذا يعني أن مستويات (فان هيل) ليست منفصلة. و تبدو أن المستويات ديناميكية أكثر منها ثابتة، ومتصلة أكثر منها منفصلة، حيث يتذبذب تقدم التلامذة أو تراجعهم بين مستويين أثناء انتقالهم من المستوى الأدنى للأعلى. ويؤكد كل من برغر وشونيسي (Burger & Shaugnessy, 1986) على أن هذا التذبذب يتضح أكثر بين المستويين الثاني والثالث، ولكن في بحثنا الحالي أتضح أن التذبذب حدث بين المستويين الأول والثاني، كما أنه يحدث بين ثلاثة مستويات متجاورة. في حين أظهرت نتائج دراسة ويلسن (Wilson, 1990) عند تدريج اختبار مستويات التفكير الهندسي لفان هيل، إلى انحصار قيم حدود الملائمة المعيارية ما بين −2 و2 لوغاريتم، والتي تتناسب مع مستوى الدلالة الإحصائية 0.05، والتي أكدت على أن مستويـــات (فان هيل) متسلسلة بحسب توقعات النظرية ما عدا المستوى الخامس، وهذا ما اتفق مع ملاحظات مصممي الاختبار حول قدرة الاختبار على رصد المستوى الخامس (التجريد الصارم)، وبينت النتائج كذلك أن هناك نوعاً من التداخل بين المستويات. وتوصلت دراسة مايكل (Michiel, 2012) إلى أن جميع مفردات الاختبار مطابقة لنموذج راش، و قيم صعوبتها مناسبة، حيث انحصرت قيم حدود الملائمة المعيارية لها ما بين −2 و2 لوغاريتم، مع حذف المفردة 3 والتي أجاب عنها جميع التلامذة، كما أكدت على التسلسل الهرمي للمفردات وفقاً للمستويات الأربعة، مع وجود تداخل بين قيم صعوبة المفردات بين المستويات. في حيت توصلت دراسة (STOLS, 2015) والذي استخدم فيها اختبار (فان هيل)

لمستويات التفكير الهندســـي، و على الرغم مما أظهرته خريطة رايت فرد- مفردة من تكافؤ في توزيع العتبات بينهما، إلا أن نتائج المعايرة أشارت إلى أن بعض المفردات كانت خارج مكانها المفترض حسب نظرية (فان هيل) الهرميـــة، حتى بعد تصحيحها من أثر التخمين بقيت الإشكالية قائمة. وعلى الرغم من وجود نمط تنازلي واضح في الإجابات الصحيحة ضمن المستويات كما تنبأت به نظرية (فان هيل). إلا أنها أكدت إلى حد ما قوة نظرية (فان هيل) في تنمية التفكير الهندســـي.

أما دراسة سيرو (Serow, 2007) والتي أظهرت نتائجها بعد تدريج مفردات الاختبار، بأنها وقعت ضمن حدود الملائمة المقبولـــة، وبأنها تتدرج تدرجاً هرميـــاً واضحاً ومتسلســـلاً وفقاً لمستويات نظرية مخرجات التعلم الملاحظ (SOLO)، حيث سمحت هذه النتائج بالمقارنـــة بين صعوبة المفردات من خلال فئة الاستجابات بشأن العلاقات بين الأشكال وبين خصائصها ضمن سياق المثلثات والأشكال الرباعية، والاستجابات حول المفردات المتعلقة بالعلاقات بين خصائص الأشكال.

كما تشابهت نتائج بحثنا الحالي مع نتائج دراسة (القهوجي، 2015) والتي توصل فيها إلى أن تدريج اختبار مهارات التفكير العلمي باستخدام نموذج راش إلى حذف 16 مفردة غير مطابقة للنموذج. و دراسة (عوض الله، 2000) والتي استخدم فيها نموذج راش في تدريج اختبار للذكاء، وحذف 76 مفردة كانت غير مطابقة لنموذج راش. في حين توصلت دراسة كرسليب وشين (Crislip and Chin, 2001) بعد تدريج اختبارات ولاية هاواي للقـــدرات الأساسية، وباستخدام مؤشر إحصائي المطابقة كاي مربع (كا^2) أن ما نسبته 37% من مفردات الاختبار غير ملائمة لنموذج راش.

إلا أنها اختلفت مع دراسة (سيف، 2017) والتي أظهرت نتائجها أن تدريج صعوبة مفردات اختبار مستويات التفكير الهندسي كان مطابقاً لنموذج راش استناداً لقيم مربع كاي لحسن المطابقة عند مستوى دلالة إحصائية 0.05 والتي لم تستبعد فيه أي مفردة في هذا المرحلة. وكذلك نتائج دراسة (بوسالم، 2008) عند تدريج مفردات اختبار القياس النفسي والتي تبين أنه يقيس متغير أحادي البعد، وأن تدريج المفردات لم يتغير برغم تغير مستويات قدرة الطلبة المستخدمين في التدريج. وأن استخدام نموذج راش اللوغاريتمي في بناء اختبار موضوعي في وحدة القيـــاس النفسي تتدرج مفرداته بوحدة قياس مطلقة قدم تفسيراً موضوعياً لإداء الطلبة على الاختبار.

وعليه وبناء على كل ما سبق من نتائج، يمكننا القول أن هذه الفرضية محققة، والتي تشير إلى أنه تتباين تقديرات صعوبة مفردات اختبار مستويات التفكير الهندسي وتدرجها باستخدام نموذج راش. والتي تسمح بتمثيل قدرات الأفراد ومواقع المفردات على تدريج خطي متصل فتري بحسب صعوبتها، من خلال تحديد التناظر بين البيانات المشاهدة ومواقع الأشخاص على المتغير الكامن. فإذا ما حددنا مواقع الأفراد على المتغير الكامن، نستطيع عندئذ مقارنتهم ببعضهم بعضاً. هذا التمثيل الذي يشير إلى أحد خصائص القياس الموضوعي التي جاءت به نظرية الاستجابة للمفردة ونماذجها المتعددة ومنها نموذج راش أحادي البارامتر، حيث يمكن تحديد موقع كل فرد على متصل القدرة مما يسمح بإجراء عملية تشخيص لجوانب القوة والضعف في تعلم التلميذ، ومعرفة الأهداف التي حققها والتي لم يحققها التلميذ لوحده بغض النظر عن مستوى الجماعة التي ينتمي إليها، وهو الهدف الأساسي للتربية الحديثة المبنية على تحديد المهارات والكفايات المطلوب إتقانها.

1-4- تفسير ومناقشـــة نتائج الفرضية الرابعـــة:

والتي تنص علـــى أن:" تتباين تقديرات قدرات الأفراد ودالـــة المعلومات المقابلة لكل درجة كلية خام محتملة على مفردات اختبار مستويات التفكير الهندسي باستخدام نموذج راش".

تبين من خلال تحليل النتائج الموضحة في الجدول رقم (38)، إلى أن قدرات التلامذة تراوحت ما بين -3.44 و3.38 لوجيت، حيث أن التلميذ الذي قدرته -3.44 لوجيت يحصل على الدرجة الكلية الخام 1 في الاختبار، أي أن هذا الاختبار يعتبر صعباً بالنسبة لقدرة هذا التلميذ فهو لم ينجح في الإجابة إلا على مفردة واحدة من مفردات الاختبار. والتلميذ الذي يكون في متوسط القدرة 0.02 لوجيت وهي تقريباً صفراً لوجيت يحصل على الدرجة الكلية الخام 11 في الاختبار، أي أنها عند أقصى دالة للمعلومات في الاختبار أي أن صعوبة هذا الاختبار قريبة من مستوى قدرة التلميذ، حيث استطاع هذا التلميذ أن يجيب على نصف مفردات الاختبار. والتلميذ الذي قدرته 3.38 لوجيت يحصل على الدرجة الكلية الخام 21 في الاختبار، أي أن هذا الاختبار يعتبر سهلاً بالنسبة لقدرة هذا التلميذ فهو نجح في الإجابة على مفردات الاختبار 21. وهو ما يظهره الشكل رقم (30) بأكثر وضوح للعلاقة بين الدرجة الكلية الخـــام والمحتملة على الاختبار والقدرة المقابلة لها، بأنها علاقة طرديـــة موجبة والتي كانت في شكل دالـــة منتظمة متزايدة، حيث تزيد القدرة كلما ازدادت الدرجة الكلية الخـــام المقابلة لها. فالدور الأساسي للمنحنى المميز للاختبار في نظرية

الاستجابة للمفردة كما هو معلوم، هو امدادنا بوسائل لتحويل درجات القدرة إلى درجات حقيقيـــة. فالدرجة الحقيقيـــة في النظرية الكلاسيكية حسب بيكر(2010)، هي الدرجات المتوقعة التي سوف تتجمع حول المتوسط الحسابي بعد أخذ الاختبار عدة مرات. في حين أن الدرجة الحقيقيـــة في ظل نظرية الاستجابة للمفردة تعتمد على احتمال الاستجابـــة الصحيحة على كل مفردة من مفردات الاختبار. فالمنحنى المميز للاختبار هو العلاقة الوظيفية (التقييسيـــة) بين الدرجة الحقيقية ومقياس القدرة. لهذا يمكننا إيجاد الدرجة الحقيقية المقابلة لأي مستوى محدد من القدرة من خلال المنحنى المميز للاختبار، والذي يعتمد شكله على عدد من العوامل منها عدد مفردات الاختبـــار ونموذج المنحنى المميز للمفردات المستخدمة، وقيم معالم المفردات. فمنتصف الدرجة الحقيقية يحدد صعوبة الاختبار في صيغة رقمية، لكن من الأفضل أن يعرف إنحدار المنحنى المميز للاختبار لفظياً. بالنسبة لمعظم أغراض التفسير، فإن هذين الوصفين كافيان لمناقشة المنحنى المميز للاختبار الذي تم رسمه، وأمكن التحقق منه بصريـــاً من خلال الشكل رقم (30) الذي يوضح العلاقة المتجمعة بين الدرجة الكلية الخـــام المحتملة على الاختبار والقدرة المقابلة لها باللوجيت. حيث أن منحنى معلومات الاختبار جاء متماثلاً بطريقة مقبولة وذا ارتفاع واضح جداً في الوسط، وبالتالي فهو يأخذ الشكل الاعتدالـــي المتوقع لمثل هذه المنحنيات، ويشير شكل المنحنى إلى أنه يتم تقدير القدرة بأعظم دقة بمصاحبة متوسط مقياس القدرة. وبما أن القدرة متغير متصل تكون المعلومات متغيراً مستمراً أيضاً. و يتضح من خلال الشكل رقم (31)، أن ذروة دالـــة معلومات الاختبار حدثت في نقطة أقل بدرجة طفيفة من النقطة المتوسطة لمقيـــاس القدرة. فمتوسط القدرة يساوي تقريباً صفراً، أي أنها عند أقصى دالـــة للمعلومات، فالاختبار يعطي أقصى دالـــة للمعلومات عند مستوى القـــدرة المتوسطة، حيث بلغت قيمتها 4.58، وخطأ معياري للقياس 0.47، والمقابلة لمنتصف الدرجة الحقيقية 11، والتي تقابل صفر التدريج.

والقيمة المتوسطة لمعالم صعوبـــة المفردات هي مؤشر على المكان الذي سيقوم الاختبار فيه بوظيفته على مقياس القدرة، كما أشارت نتائج الأخطاء المعياريـــة للقياس، من خلال ما يوضحه الشكل رقم (32) لخريطة المسار(The Pathway Map) أوالمخطط الفقاعي، للأخطاء المعياريـــة للقياس، لكل من صعوبة مفردات اختبار مستويات التفكير الهندســـي، وقدرات الأفراد. أن بعض الفقعات كان حجمها كبير وهي عبارة عن الدرجات المتطرفة والتي تكون أخطاؤها المعيارية أكبر، فبدل تقديم خطأ معياري واحد للقياس

ينطبق على جميع المختبرين بغض النظر عن درجاتهم في الاختبـــار بحسب النظرية الكلاسيكية للقياس، تجعل نظرية الاستجابة للمفردة الاختبارية ومنها نموذج راش أنه من الممكن تقديم تقديرات منفصلة للخطأ المعياري لكل فرد، ولكل مستوى قدرة معين، لذلك فإن الخطأ المعيـــاري للقيـــاس في هذه الحالة يتميز بامكانية التعميـــم على مجتمعات من الأفراد مختلفى القدرة، ويمكن إيجاد متوسط الأخطاء المعيارية الذي يعد مؤشراً لهذه الأخطاء في مجتمع هؤلاء الأفراد. فقيم الخطأ المعياري للقيـــاس عموماً كانت منخفضة نوعاً ما، فهذه القيم لها علاقة بعدد المفردات فكلما كان عدد مفردات الاختبار كبير في نموذج راش، فإن الخطأ المعياري للقيـــاس للقدرة يكون أصغر $SE(\theta)$، مما يؤدي إلى تقديرات للقدرة تكون أكثر دقة. باعتبار أن دالـــة معلومات الاختبار، هي مجموع دوال معلومات مفردات الاختبار، مما يجعلها تساهم في دالـــة معلومات الاختبار ككل، كما أن الخطأ المعياري للقياس يقل كلما قلت قيمة التخمين.

حيث يرى هامبيلتون (Hambleton, 1991) في هذا السياق أن كمية المعلومـــات المقدمة بواسطة الاختبار عند مستوى القدرة (θ) تكون مرتبطة عكسياً مع الدقة عند أي نقطة على متصل القدرة (θ)، وهي الخطأ المعيـــاري لتقدير القدرة ($SE(\theta)$) والذي يحصل عليه عندما يحصل على تقديرات الاحتمال الأقصى، وبصفة عامة فإن حجم الخطأ المعيـــاري يعتمد على مجموع مفردات الاختبار، أي أن أقل الأخطاء المعيارية ترتبط بالاختبارات الطويلة؛ ومدى المطابقة بين صعوبة المفردة وقدرة الممتحنين، كما أن أقل الأخطاء المعيارية ترتبط بالاختبارات المكونة من مفردات مع معالم الصعوبة مساويةً تقريباً لمعلم القدرة وهذا بخلاف الاختبارات السهلة أو الصعبة نسبياً؛ كذلك فإن كفاءة مفردات الاختبار تعني أن أقل الأخطاء المعياريـــة ترتبط بالمفردات مرتفعة التمييز أي أن الاجابات الصحيحة لا يمكن الوصول إليها بالتخمين. وكلما كانت الأخطاء المعيارية صغيرة كانت المعلومات التي يشتمل عليها الاختبار دقيقة، وهو ما تحقق من نتائج التحليل حيث تراوحت قيم الأخطاء المعيارية للقياس لمفردات اختبار مستويات التفكير الهندسي، كما يظهرها الجدول رقم (37) بعد تدريجها ما بين 0.09 و 0.12، وبمتوسط حسابي قدره 0.10 وانحراف معياري قدره 0.01 وهي قيم صغيرة جداً وقريبة من الصفر. وبالتالي فإن انخفاض قيم تقدير الخطأ المعياري للقياس لتدريجات صعوبة مفردات اختبار مستويات التفكير الهندســـي بوجه عام و تقدير قدرة كل فرد والتي تراوحت قيمه ما بين 0.92 و0.93 لوجيت. يدل على إرتفاع قيم دالـــة المعلومات للمفردات ومن ثم الاختبار ككل، فالأخطاء المعيارية للقياس لتدريجات المفردات هي دلالات

لثبات المفردات، ومؤشراً لدقة القياس وكلما كانت الأخطاء المعيارية صغيرة كانت المعلومات التي يشتمل عليها الاختبار دقيقة.

ويرى بيكر (2010) في هذا السياق أن دالـــة المعلومات نمطية إلى حد ما والتي من خلالها يتم تقدير مستويات القدرة المختلفة بدرجات مختلفة من الدقـــة، مما يعطيها أهمية عظمى لكل من واضع الاختبار ومستخدم الاختبار، لأن الدقة التي يتم من خلالها تقدير قدرة المفحوص تعتمد على أين تم وضع قدرة المفحوص في مقياس القدرة.(ص99)

لم تختلف هذه النتيجة مع نتائج دراسة كل من (علي زكري، 2009) والذي توصل فيها إلى وجود علاقة طردية موجبة بين الدرجة الكلية الحقيقيـــة على اختبار القدرات العقليـــة (أوتيس- لينون) والقدرة المقابلة لها، وكذلك مع دراسة (أيمن، 2014) والذي توصل فيها إلى أن أفضل تقديرات لقدرات الأفراد المقابلة لكل درجة محتملـــة على اختبار مهارات البحث العلمي، وإلى وجود علاقة طرديـــة موجبة بين الدرجة الكلية الحقيقية والقدرة المقابلة لها.

تشابهت كذلك مع نتائج دراسة (الخياط، 2012)، والتي تبين فيها أن تكرارات الأفراد حسب الدرجة الكلية على الاختبـــار، وتقدير مستويات القـــدرة، والخطأ المعياري المرتبط بتقدير القدرة الحقيقية كان مناسباً، في حين أن مستويات القـــدرة تراوحت ما بين -3.33 لوجيت للدرجة الكلية 2 و3.83 لوجيت للدرجة الكلية 39 على اختبار المهارات الرياضية (TIMMS)، والتي دلت على توزيع أفراد عينة الدراسة على سلم مستويات القدرة ولجميع الأفراد.

وعليه وبناء على ما سبق من نتائج، يمكن القول أن الفرضيـــة محققة، والتي تشير إلى أنه تتباين تقديرات قدرات الأفراد ودالـــة المعلومات المقابلة لكل درجة كليـــة خام على مفردات اختبار مستويات التفكير الهندســـي. وبالتالي تم تقدير قدرات الأفراد على الاختبار، مما يحقق موضوعية تقدير القدرة لهذا الاختبار الذي تم تدريجه باستخدام نموذج راش.

1-5- تفسير ومناقشـــة نتائج الفرضية الخامســـة:

والتي تنص علـــى أن:" يحقق تدريج اختبـــار مستويات التفكيـــر الهندســـي باستخدام نموذج راش استقلالية القياس بإختلاف عينة التدريج و الاختبار المستخدم".

أشارت نتائج التحليل المبينة في الجدول رقم (39) والشكل رقم (33)، والمتعلقة بمدى تحقق استقلالية القيـــاس عن قدرة العينة التي تؤدي الاختبـــار، بعد ما تم تقسيم عينة التدريج إلى عينتين إحداهما "مرتفعة المستوى"، والأخرى "منخفضة المستوى" بعد ترتيب أفراد عينة التحليل الكلية البالغة 566 تلميذاً وتلميذة، وعمل المقارنات بين التقديـــرات الناتجة عن العينتين و العلاقات التقييسية بينها ومدى تكافئهما إحصائياً، مع الأخذ بالاعتبار الخطأ المعيـــاري للقياس لهذه التقديرات، والتحقق من مدى تحرر قدرة الفرد المقدرة بهذا الاختبـــار عن قدرة باقي الأفراد الذين يجيبون عليه، وإلى تحرر صعوبة المفردة عن قدرة الأفراد الذين يجيبون على الاختبار.

كذلك نتائج التحليل الموضحة في الجدول رقم (40) والشكل رقم (34)، المتعلق باستقلالية القيـــاس عن مجموعة المفردات التي أجاب عليها الأفراد أن تقديرات صعوبة المفردات المشتقة من تحليل أداء الأفراد على الاختبار الكلي، بأنها تتكافأ إحصائياً مع التقديرات المتناظرة لصعوبة المفردات المشتقة من تحليل أداء نفس الأفراد على الاختبار السهل و الاختبار الصعب، مما يعني لنا عدم تأثر تقدير صعوبة المفردات باختلاف مجموعة المفردات المستخدمة في التحليل. وبالتالي تحرر تقدير صعوبة المفردة من صعوبة باقي المفردات المستخدمة في الاختبار.

وهو ما أشارت إليه كذلك نتائج التحليل الموضحة في الجدول رقم (42) للتقديرات المتناظرة لقدرة الفرد المشتقة من كل من الاختبار السهل والاختبار الصعب (بعد تعديل التدريج) وكذلك تلك المشتقة من الاختبار الكلـــي، بعد إجراء عملية الموازنة بين صعوبة مفردات كل من الاختبارين الصعب والسهل والتدريج المرجعي بعد إجراء التعـــادل الرأســـي. بأنها متكافئة والتي تدل على عدم تأثر تقدير قدرة الفرد بمجموعة المفردات المستخدمة طالما أنها ملائمة للنموذج ومناسبة لمستوى قدرة الأفراد. إضافة إلى ذلك تبين أن متوسط الصعوبة للمجموعة المشتركة بلغ 0.0005 بالنسبة للاختبار المشترك، و0.54 بالنسبة للاختبار السهل، و-0.43 بالنسبة للاختبار الصعب. مما يدل على أن متوسط القدرة للاختبار المشترك يشير إلى ما ينبغي فعله، ويشير متوسط الاختبار السهل إلى أن متوسط القدرة يفوق متوسط صعوبة مفردات الاختبار، وهذا ما ينبغي فعله تماماً وفي النهاية كان متوسط القدرة للاختبار الصعب أقل من متوسط صعوبة المفردات ومرة أخرى يعتبر هذا ما يتوقعه الفرد. وبما أنه تم قياس صعوبة المفردة والقدرة بنفس القياس لذلك يمكننا استخدام متوسط القدرات لوضع الاختبارات في المقياس العام. ويعتبر ما

تم انجازه فيما سبق كمعادلـــة للاختبار، أي وضع الاختبارات الثلاثة في مقياس عام. والتي تعتبر من إحدى المزايا الرئيسة لنظرية الاستجابة للمفردة والتي تتميز بها عن نظرية القياس الكلاسيكية. وبناء على هذه النتائج يمكن القول أن الاختبار قد حقق فرضيات نموذج راش، وهي: أحادية البعد، واستقلالية القياس من خلال تحرر تقديرات كل من صعوبة المفردة وقدرة الفرد، ومن قدرة الأفراد الذين يؤدون الاختبار، أي تحرر القيـــاس من توزيع أداء العينة، وتحرر تقديرات كل من صعوبـــة المفردة وقدرة الفرد، من مجموعة المفردات المستخدمة، أي تحرر القياس من الاختبار المستخدم. ويتضمن هذا ثبـــات التقدير لكل من قدرة الفرد و صعوبة المفردة، بالرغم من اختلاف عينات الأفراد أو الاختبارات المستخدمة. وتعد قاعدة الثبات إحدى أساسيات تكافؤ الاختبارات في إطار نظرية الاستجابة للمفردة.

حيث يرى بيكر (2010) في هذا السياق، أن ثبات معالم المفردة (Invariance of Item paramter) للمجموعة هو أحد ملامح القوة لنظرية الاستجابـــة للمفردة، والتي تفترض أن قيم معالم المفردة تخص المفردة ولا تخص المجموعة التي تجيب على المفردة. أما في ظل النظرية الكلاسيكيـــة فالعكس هو الصحيح تماماً.

وتضيف كاظم (1988أ) إنه بتحقيق هذا الفرض، أي استقلال تقديرات المفردات عن عينة الأفراد، يعني أنه يمكن التغلب على المشكلات المتصلة بعينة التقنيـــن، فليس من الضروري أن تكون عينة التقنين المستخدمة في تدريج مفردات الاختبار ممثلة للمجتمع. كما ليس هناك ضرورة أن يكون توزيع الأداء لعينة التقنيـــن يخضع لشكل معين من التوزيعات، وذلك طالما أن تدرج المفردات لا يتأثر بتوزيع الأداء لعينة التقنيـــن.

لم تختلف هذه النتائج مع ما توصلت إليه دراسة (كاظم، 1988ب)، والتي وجدت تقارب في مستوى الاختبـــار الذي صممته مع مستوى عينة التحليل وصولاً إلى دقة التقديرات المؤدية إلى الموضوعية في عملية التعـــادل الرأســـي في تدريج مفردات الاختبار، حيث اصبحت خصائصه مشتركة وصالحة لجميع خصائص الأفراد، وبالتالي تحرر القياس من صعوبة المفردات وقدرات الأفراد. كما اتفقت هذه النتائج مع نتائج دراسة (بوسالم، 2008)، الذي توصل فيها إلى أن اختبار القياس النفســـي حقق استقلالية عملية القياس من ناحيتين استقلالية أو تحرر عملية القياس في حد ذاتها من قدرات الأفراد، و تدريج المفردات لم يتغير برغم تغير مستويات قدرة الطلبة المستخدمين في عملية التدريـــج.

إلا أنها اختلفت مع نتائج ما توصلت إليه دراسة (سيف، 2017) التي تبين فيها أن هناك مفردتين لم تكن متكافئتين للتقديـــرات الإحصائية المناظرة لها، وعليه استبعدت هاتين المفردتين من اختبار مستويات التفكير الهندســـي لفان هيل لتحقيق استقلالية القياس بما يحقق موضوعية القياس، كما تمثل بنموذج راش.

وعليه وبناء على ما سبق من النتائج، يمكننا القول بأن هذه الفرضية محققة، والتي تشير إلى أن تدريج اختبار مستويات التفكير الهندســـي باستخدام نموذج راش يحقق استقلالية القياس باختلاف عينة التدريج.

1-6- تفسير ومناقشة نتائج الفرضية السادســـة:

والتي تنص علـــى أن:" يمكن اشتقاق معايير كمية تفسر تقديـــرات الأفراد على اختبار مستويات التفكير الهندســـي في صورته النهائية بعد التدريج على تلامذة المرحلة الثانوية".

أشارت نتائج التحليل في الجدول رقم (44)، الذي يوضح تقدير صعوبة المفردات المقدرة باللوجيت والواط مرتبة تصاعدياً تبعاً للصعوبة بعد ما تم تدريج اختبار مستويات التفكير الهندســـي باستخدام نموذج (راش)؛ والجدول رقم (45) الذي يوضح معايير القـــدرة لعينة التقنـــين 566 تلميذاً وتلميذة، باستخدام نموذج راش بوحدة اللوجيت وتحويلها إلى وحدة الواط، و النظرية الكلاسيكية في القياس (المعايير التائية – الرتب المئينية)؛ والجدول رقم (46) الذي يمثل توزيع متغير القدرة باللوجيت على مجموعة المعايرة، وتوزيع متغير درجة صعوبة مفردات اختبار مستويات التفكير الهندسي باللوجيت بعد التدريج. وذلك بعد تحقق محكات الدقة، وبعد تقدير قدرات التلامذة على هذا الاختبار بوحدة اللوجيت، وقيامنا بتحويل تدريج وحدات اللوجيت لصعوبة مفردات الاختبار في صورته النهائية، إلى تدرج وحدات مئوية بوحدة الواط، إضافة إلى ذلك استخراج الباحث لمعايير الدرجات التائية، وكذلك معايير الرتب المئينية المقابلة لكل درجة من الدرجات الخام التى حصل عليها الأفراد في عينة التقنيـــن. وهي معايير الطريقة الكلاسيكية المألوفة قدمها الباحث لإمكانية إجراء مقارنـــة بين كل من المعايير المستخرجة وفق النظريتين الكلاسيكية في القياس، ونظرية الاستجابة للمفردة (نموذج راش).

وفي ضوء جداول المعاييـــر السابقة الذكر، يمكن الحكم على مستوى قدرة التلميذ وتفسير مستوى قدرته بالنسبة لإقرانه، وتحديد ترتيبه بين أقرانه، ودرجته التائيـــة، ورتبته المئينية. وبالعودة إلى الجدول رقم (45) يمكن الحكم على مستوى قدرة التلميذ وتفسير مستوى قدرته بالنسبة لأقرانه، فمثلاً: نجد تقدير

قدرة التلميذ على الاختبار الكلي عند درجة خام 11، و0.02 لوجيت، و50 واط، تقابلها في جدول المعايير الدرجة التائية 47، الرتبة الميئنية 40. كما يتضح للباحث أن مدى القدرة التي يقيسها الاختبار كان واسعاً بعد التدريج، كما أن الدرجات الخام أشارت إلى وجود مدى واسع من درجات المفحوصين و هذا دليل على أن إجابات أفراد عينة البحث كانت متباينة ومتنوعة، مما يدل على أن الاختبار حساس لرصد التباين في القدرة على التفكير الهندســـي وفقاً لمستويات (فان هيل) التي يقيسها، وبأنه اختبار جيد ومتوافق مع الأفراد الذين تمتد مستوياتهم إلى مدى واسع.

وبعد إجراء الباحث لمقارنـــات بين قدرات الأفراد ومفردات الاختبار ومجموعات المعايرة، لتسهيل عملية تفسير قيم صعوبة المفردات التي تم الحصول عليها وقدرة المفحوص في هذا القياس. ظهر أن الأفراد الذين يندرجون ضمن مستوى قدرة معين يحصلون على درجة حقيقية على الاختبار، والتي تعد بمثابة المتوسط النظري (المفترض نظرياً) لكل الدرجات الخام للاختبار، والذين سيحصلون عليها المفحوصون في هذا المستوى من القدرة. ويتم تفسير قيم صعوبة المفردات التي تم الحصول عليها وقدرة المفحوص في هذا القيـــاس، وبهذا تتم معايرة الاختبار في صورته النهائية. والناتج النهائي لعملية معايرة الاختبار يمثل تعريف تدريج مقياس القدرة وفي نموذج راش نجد لهذا المقياس وحدة قياس 1 ونقطة للمنتصف هي صفر.

وباستقراء النتائج الموضحة في الجدول رقم (46)، والشكل رقم (35) يظهر أن 522 تلميذاً وتلميذة من أصل 566 من العينة الكلية تمكنوا من الإجابة بسهولة نوعاً ما على مفردات الاختبار، حيث أن صعوبة المفردات غطت في معظمها مجال القدرة مابين –1.44 إلى 1.44 لوجيت، في حين وقع تلميذين فقط في المجال المنخفض للقدرة ما بين –1.44 إلى–2.15 لوجيت، و42 تلميذاً وتلميذة وقعا في المجال أو النطاق الأعلى للقدرة مابين 1.44 إلى4.63 لوجيت. ويتضح من هذه النتائج إلى أن الاختبار في حاجة لبعض المفردات أكثر صعوبة لتقدير قدرات التلامذة العليا وهم (التلامذة الذين حصلوا على درجات مابين 18 إلى 22 على اختبار مستويات التفكير الهندسي المبني وفقاً لنموذج (فان هيل)، في أي مراجعة لاحقة للاختبار لتغطية قدرات التلامذة.

اتفقت هذه النتائج مع ما توصلت إليه دراسة (سيف خليل، 2017) في استخدم نموذج (راش) في تدريج اختبار مستويات التفكير الهندســـي لفان هيل، وايجاد معايير الأداء لتقديرات صعوبة المفردات

وقدرات الأفراد وتحويلها من وحدة اللوجيت لوحدة الواط. ودراسة (العطاس، 2014) والتي تم فيها توفير معايير الأداء على اختبار مستويات التفكير الهندســـي لفان هيل وفقاً للنظرية الكلاسيكية، تمثلت في الدرجات المعيارية والدرجات التائية والرتب المئينية. ودراسة (علي زكري، 2009) في استخرج معايير اختبار (أوتيس- لينون) للقدرة العقلية، لقدرات الطلبة بوحدة اللوجيت ووحدة المنف والدرجات التائية والرتب المئينية.

وعليه وبناء على ما سبق من نتائــج، يمكن القول بأن الفرضية محققة، والتي تشير إلى أنه يمكن اشتقاق معاييــر كمية تفسر تقديرات الأفراد على اختبار مستويات التفكير الهندســـي في صورته النهائية بعد التدريج على تلامذة المرحلة الثانوية.

1-7- تفسير ومناقشـــة نتائج الفرضية السابعـــة:

والتي تنص علـــى أن:" يتحقق صدق وثبـــات اختبار مستويات التفكير الهندســـي في صورته النهائية بعد التدريج باستخدام نموذج راش".

كشفت نتائج التحليل في شقه المتعلق بمدى صدق الاختبـــار بعد تدريجه، أن صعوبات المفردات وقدرات الأفراد تتدرج على متصل واحد يمثل متغيراً واحداً وهو التفكير الهندســـي. مما يعني أن تدرج قدرات الأفراد على المتغير محددة لتقديرات أدائهم على هذا الاختبار، فصدق القياس يعني تحقق أول مطالب الموضوعية في أداة القياس التي أنشئت باستخدام نموذج (راش)، وهو أن تعرف مفردات الاختبار فيما بينها متغيراً واحداً. وبالعودة إلى نتائج الجدول رقم (37)، والشكل رقم (27)، اللذان يشيران إلى أن قيم مؤشرات الملاءمة الإحصائية لمتوسط المربعات التباعدية والتقاربية لمفردات اختبار مستويات التفكير الهندســـي كانت قريبة من الوضع المثالي الذي يفترضه النموذج وهو القيمة 1، والتي تراوحت ما بين 0.84 و 1.13 بمتوسط حسابي بلغ 1، وإنحراف معياري 0.04 بالنسبة لمؤشرات الملاءمة الإحصائية لمتوسط المربعات التقاربية (الداخلية)، وبمتوسط حسابي بلغ 0.99، وإنحراف معياري 0.08 بالنسبة لمؤشرات الملاءمة الإحصائية لمتوسط المربعات التباعدية (الخارجية). و بالتالي فإن هذه القيم لم تخرج عن مجال القيـــاس المنتج الذي حدده كلٍ من رايت ولينكر (Wright and Linacre, 1994) بالمجال مابين 0.70 و1.30 أو المجال ما بين 0.50 و1.5 عموماً والقيم القريبة من الواحد تكون أفضل. وتتعلق هذه المؤشرات بصدق تدرج مفردات الاختبار في تعريفها للمتغير موضوع القياس. كما تتعلق أيضاً بصدق

تدرج قدرات الأفراد على متصل هذا المتغير، الذي يقوم على صدق استجابات الأفراد على الاختبار. ولإضفاء المزيد من المصداقيـــة على الاختبار الحالي، قام الباحث بالتحقق من قطبيـــة المفردات والتي تختبر مدى عمل المفردات في الإتجاه الواحد لتحديد التكوين الفرضـــي المقاس. أي أن مجال قيم معامل الارتباط المتسلسل (مرتبة من الأدنى إلى الأعلى)، كما تظهره نتائج التحليل الموضحة في الجدول رقم (37)، أن نقطة قياس معاملات الارتباط لكل مفردات اختبار مستويات التفكير الهندســـي كانت موجبة (point measure correlation PT-MEA CORR). والتي تراوحت قيمها ما بين 0.27 للمفردتين7 و8، والقيمة 0.47 للمفردة 15، مما يدل على أن هذه المفردات تقيس التكوين الفرضـــي الذي ينوي الباحث قياسه.

كما تبين من خلال نتائج التحليل الموضحة في الجدول رقم (47)، والشكل رقم (36)، إلى تحقق أحادية البعد باستخدام أسلوب المكونات الأساسية المعتمدة على البواقي (PCAR)، والذي يظهر الإختلافات بين الأبعاد، وإلى ارتفاع مؤشر القيمة الملاحظة للتباين الذي فسره العامل الرئيسي بـــ 23.5 % بعد التدريج و حذف ستة 6 مفردات من الاختبار، عن قيمته السابقة والتي كانت 21.8 % قبل التدريج، كذلك تحقق المحك الثاني وهو نسبة التبـــاين الذي يفسره العامل الثاني الذي كان في المجال المقبول، وبالتالي تحقق أحادية البعد لاختبار مستويات التفكير الهندســـي.

في حين كشفت نتائج تحليل هذه الفرضية في شقها الثاني المتعلق بالثبـــات، إلى تحقق استقلالية القياس للاختبار – في ما سبق من هذا البحث – من ناحيتين وهما: تحرر القياس من قدرة العينة التي تؤدي الاختبار؛ و تحرر القياس من المفردات التي يجيب عليها الأفراد.

وأشارت كذلك نتائج التحقق من اللاتغاير بعد التقسيم العشوائي لعينة التدريج والمقارنة بين تقديرات معالم المفرة لتحديد درجة الخطيـــة فيها. بأن تقديرات مواقع المفردات في العينة الفرعية الأولى ليست مساوية لنظيرتها في العينة الفرعية الثانية بالرغم من أنهما يظهران بالترتيب نفسه. وعند حساب معامل الارتباط لبيرسون بين هاتين المجموعتين من تقديـــرات معالم المفردات فقد بلغ 0.994، ولإستكمال الصورة فقد بلغ معامل ارتباط بيرسون بين تقديرات العينة الفرعية الأولى وتلك في العينة الكلية 566 تلميذاً وتلميذة، القيمة 0.998، وبين تقديرات المواقع في العينة الفرعية الثانية ونظيرتها في العينة الكلية أيضاً القيمة 0.998. و تدل القيم المرتفعة لحجم معامل الارتباط بيرسون بأنه يمكننا استخدام تحويل خطي

لتحويل التقديرات على المقياس إلى ما يناظرها على مقياس آخر دون أن نفقد أية معلومات تتعلق بمطابقة البيانات للنموذج، أو تقديرات مواقع الأفراد والمفردات، ويتضح لنا من خلال هذه النتائج أنه لدينا بينة أخرى عن اللاتغاير (بمعنى"الثبات") بالنسبة لاختبار مستويات التفكير الهندســـي في صورته النهائية (دي إيالا، 2017، ص.76).

ويتضح كذلك من خلال تحليل نتائج الجدول رقم (48) للتحقق من مؤشرات ثبات الاختبار بعد تدريجه، إلى أن معامل ثبات مفردات الاختبار بلغ 0.99، وهي قيمة مرتفعة تدل على كفاية عينة الأفراد في الفصل بين المفردات، وبالتالي تعريف متصل السمة الذي تقيسه هذه المفردات، وهو مستقل عن طول الاختبار. في حين بلغ معامل فصل المفردات (GI) 9.20 وهي قيمة أكبر من القيمة 2، وبالتالي التأكيد على التسلسل الهرمـــي لصعوبات مفردات اختبار مستويات التفكير الهندسي في التدريج النهائي. حيث يرى لينكر (Linacre, 2012, p.644) في هذا السياق أن أقل قيمة لفصل المفردات هي 3، وأقل قيمة لثبات المفردات هي 0.90، لكي تعد هذه المفردات كافية لتعريف متصل السمة التي تقيسها وهي التفكير الهندســـي، وكفاية عينة الأفراد لتأكيد هرمية صعوبة المفردات، وهي تكافئ الصدق البنائي للأداة. كما أن معامل الفصل يشير إلى عدد الطبقات التي تكون صعوبة مفردات الاختبار، ولهذا فإن صعوبات المفردات تنتشر بشكل جيد على سلم لوغاريتمـــي مع ثبات عالي. كما بلغت قيمة معامل الثبات للأفراد 0.65، والتي تدل على أن نطاق القدرة ضيق نوعاً ما، أي أن هناك ضعف في تباين قدرات الأفراد، و بالتالي فالاختبار في حاجة لمزيد من المفردات، وهو مستقل عن حجم العينة. وقد يكون السبب في ذلك إلى توزع أداء التلامذة على أربعة مستويات متباينة الصعوبة. و بلغت قيمة معامل الفصل للأفراد (GP) 1.32 وهي قيمة قريبة من القيمة 2 نوعاً ما، حيث أن أقل قيمة مفضلة لفصل الأفراد هي 2، وثبات الأفراد 0.80 لكي تكون عينة الأفراد حساسة للتمييز بين الأداء العالي والمنخفض، وفي بحثنا الحالي تعتبر قيم كلٍ من معامل الفصل والثبات قيم مقبولة نوعاً ما، إلا أنها تشير إلى حاجة اختبار مستويات التفكير الهندسي لمزيد من المفردات، واعتمادا على معامل فصل المفردات، ومعامل فصل الأفراد، تم حساب معامل الثبات لكل من المفردات والأفراد، وفق الصيغة الرياضية الآتية: ($R=G^2/1+G^2$) (Andrich, 1982, pp.95-104)، حيث يشير الرمز (G) إلى معامل الفصل، و (R) إلى معامل الثبات. وتتراوح نظرياً قيم مؤشرات الفصل من صفر إلى ما لانهاية والقيم الأعلى هي الأفضل.

وتجدر الإشارة إلى أن قيمة معامل ثبات الأفراد في نموذج (راش) يكافي ثبات الاختبار في النظرية الكلاسيكية ألفا كرونباخ (α) أو كودر–ريتشارسون (KR_{20}) تقريباً (Linacre, 2012, p.645)، والذي يمثل الحد الأدنى لمعامل الثبات (Julian, 1988). والمشار إليه في الجدول رقم (48) بالقيمة 0.64، بخطأ معياري للقياس بلغ 2.09 للاختبار ككل، وهي قيمة منخفضة وقريبة من الصفر مما يعني اقتراب الدرجة الفعلية على الاختبار من الدرجة الحقيقية وفقاً للتفسير المعياري.

إضافة إلى ذلك حدد الباحث عدد الطبقات الإحصائية المتمايزة لكل من المفردات والأفراد باستخدام الصيغة الرياضية التالية: H=(4*G+1) /3، حيث يرمز (H) إلى عدد الطبقات الإحصائية، و يرمز (G) لمعامل الفصل وبالتعويض في المعادلة السابقة بالنسبة للأفراد فإن H=(4*1.32+1) /3؛ وكذلك بالنسبة للمفردات H=(4*9.2+1) /3، وبالتالي بلغ عدد الطبقات الإحصائية للافراد القيمة 2، مما يدل على قدرة المفردات في إظهار الفروق الفردية بين الأفراد في درجة إمتلاكهم للسمة بشكل كبير إلى طبقتين لقدرات الأفراد (المرتفعة والمنخفضة) حسب الخطأ المعيـــاري للقياس، في حين بلغ عدد الطبقات الإحصائية للمفردات القيمة 13، والتي تدل على أن هناك 13 ثلاثة عشر مستوى لدى هؤلاء الأفراد. (Linacre, 2012, p.313)

ويتضح للباحث من خلال هذه النتائج، أن قيمة الثبات المتعلقة بالأفراد لم تكن كبيرة بدرجة كافية، إذ أنها لم تكن قادرة على إظهار سوى طبقتيـــن إحصائيتين متمايزتين للأفراد، أي مستويين للقدرة على التفكير الهندســـي لدى التلامذة. وقد يعزى ذلك إلى طول الاختبار بصورته النهائية 22 مفردة، والذي تراوح فيه مدى صعوبة المفردات ما بين −1.61 و1.46 لوجيت. و لزيادة معامل الثبات المتعلق بالأفراد يمكن إضافة مفردات سهلة للتلامذة الذين تقل قدراتهم عن القيمة −1.61 لوجيت، وأخرى صعبة للتلامذة الذين تزيد قدراتهم عن القيمة 1.46 لوجيت. في حين أن قيمة معامل الثبات المتعلق بالمفردات كان كبيراً جداً، فقد تبين أن المفردات قادرة على إظهار ثلاثة عشر طبقة إحصائية متمايزة تقريباً لصعوبة المفردات لدى أفراد عينة البحث؛ أي إظهار الفروق الفرديـــة بين الأفراد بشكل كبير. وقد يعزى ذلك إلى جودة المفردات المكونة للاختبار بصورته النهائية.

وبالعودة إلى نتائج الجدول رقم (48) الذي يوضح قيم جذر مربع متوسط الخطأ المعياري للثبات الحقيقي للأفراد والمفردات؛ ولمعامل الفصل للأفراد والمفردات (REAL RMSE)، وهو تباين الخطأ

الحقيقي الذي يمثل أسوأ حالة للثبات أو أدنى قيمة لثبات الاختبار، حيث سجلنا القيم التالية: 0.53 و0.52 على التوالي، و كذلك قيم جذر مربع متوسط الخطأ المعياري لثبات نموذج كل من الأفراد والمفردات وهو تباين خطأ النموذج (MODEL RMSE)، والذي يحتسب على أساس أن البيانات تناسب النموذج وأن جميع الأخطاء في البيانات هي مجرد إنعكاس للطبيعة العشوائية للنموذج، وهو أفضل حالة أو أعلى قيمة لثبات الاختبار، حيث سجلنا القيمة 0.10 لكليهما، فقيمة جذر مربع متوسط الخطأ المعيـــاري للنموذج بالنسبة للأفراد تكافئ قيمة الخطأ المعيـــاري للقياس في النظرية الكلاسيكية للقيـــاس، والذي بلغ القيمة 2.09. (Linacre, 2012, p.316)

اتفقت هذه النتائج مع ما توصلت إليه (كاظم، 1988ب) في دراستها عند بناء اختبار تحصيلي مدخل في علم النفس، وهو أن مفردات الاختبـــار عرفت متغيراً أحـــادي البعد، و حققت استقلاليـــة القياس من ناحيتين، تحرر القياس من قدرة العينة التي تؤدي الاختبار، وتحرر القياس من المفردات التي يجيب عليها الأفراد. كما بلغت قيمة ثبات قدرات الأفراد 0.90، وثبات صعوبة المفردات القيمة 0.82. وهو ما توصل إليه كذلك (بوسالم، 2008) في دراسته، إلى أن مفردات اختبار القياس النفسي تعرف متغير أحادي البعد، ويحقق استقلالية عملية القياس من ناحيتين، وهي تحرر عملية القيـــاس في حد ذاتها من قدرات الأفراد، وتحرر عملية القيـــاس من قدرات الطلبة من أثر صعوبة مفردات الاختبار.

في حين اختلفت مع نتائج دراسة (سيف خليل، 2017)، والذي توصل فيها إلى أن اختبار (فان هيل) لمستويات التفكير الهندســـي لم يحقق استقلالية القياس بما يحقق موضوعية القياس كما تمثل بنموذج راش، حيث تبين له أن هناك مفردتين لم تكن متكافئتين للتقديرات الإحصائية المناظرة لها، وعليه استبعدت هاتين المفردتين من الاختبـــار في التدريج النهائي في دراسته. ومع دراسة ستولز (STOLS, 2015) التي استخدم فيها اختبار فان هيل للتفكير الهندسي، حيث سجل انخفاضاً في مؤشر معامل الفصل لدى الافراد 0.49، بمتوسط حسابي لموقع الافراد −0.15 وهي قيمة قريبة من الصفر، أما الانحراف المعياري فكان 0.70 وهو أقل من المتوقع أي الواحد، على الرغم من أنهما ضمن الحدود المقبولة. ومع دراسة (أيمن، 2014) الذي توصل فيها إلى أن اختبار مهارات البحث العلمي يتمتع بثبات عالي، حيث بلغ معامل ثبات المفردات 0.92 ومعامل الفصل 4.52، وعدد طبقات قيم تقديرات الصعوبة 6.36، في حين بلغ معامل ثبات الأفراد 0.91، ومعامل الفصل 3.21، وعدد طبقات تقديرات القدرة 4.61. ومع دراسة

(علي زكري، 2009) التي توصل فيها إلى تحقق صـــدق اختبار (أوتيس- لينون) للقدرة العقلية من خلال ما يوفره نموذج (راش) من أحاديـــة البعد في القياس، ومعامل ثبات لقدرات الطلاب بلغ 0.97، ومعامل ثبات صعوبة المفردات بلغ 0.95.

وعليه ومن خلال عرض النتائج السابقة، فإنه يمكننا القول أن الفرضية محققة، والتي تشير إلى أنه يتحقق صدق وثبات اختبار مستويات التفكير الهندســـي في صورته النهائية باستخدام نموذج (راش)، وهو ما يعني توفر متطلبات الموضوعية في قياس متغير القدرة موضوع القيـــاس. ويعني كذلك هذا ضمنياً توفره لشرطي الصدق والثبات لتقديرات كل من صعوبـــات مفردات الاختبار، وقـــدرات الأفراد، أي أن الاختبار حقق شرطي الصدق وثبات القياس عند تدريجه في صورته النهائية.

1-8- تفسير ومناقشـــة نتائـــج الفرضية الثامنـــة:

والتي تنص علـــى أن:" لا يظهر تدريـــج صعوبـــة مفردات اختبـــار مستويات التفكير الهندســـي في صورته النهائية أداءً تفاضلياً تبعاً لمتغير الجنس".

تبين من خلال تحليل النتائج الموضحة في الجدول رقم (50)، والمتعلقة بالتحقق من الأداء التفاضلي لمفردات الاختبار والتي تعزى لمتغير الجنس (الذكور، والإناث)، والذي يعتبر كجزء من عملية التحقق من الصـــدق (صدق المفهوم)، بعد ضبط الفروق في مواقع الأشخاص للتحقق من التكافؤ بين الذكور والإناث في القدرة على التفكير الهندســـي، واختبار احتمالية إن كانت الاستجابة الصحيحة للمفردة واحدة بينهما. ووفقاً للمؤشرات الإحصائية لأداء المجموعة المرجعية والمجموعة المستهدفة على مفردات اختبار مستويات التفكير الهندســـي في صورته النهائية بعد التدريج، تبين أن مفردات الاختبار تعمل في الإتجاه المفترض منها، وبالتالي لم يستبعد الباحث أي مفردة من مفرداته في هذه المرحلة من مراحل تصميم الاختبار وفقاً لنموذج راش، أي أن مفردات اختبار مستويات التفكير الهندســـي لم تبدي أي شذوذاً غير محتمل باعتبار أنها تقيس التكوين الفرضـــي نفسه بين المجموعتين المرجعية والمستهدفة، هو ما يوضحه الشكل رقم (37) الذي يمثل منحنى بياني للأداء التفاضلي لاختبار مستويات التفكير الهندســـي في صورته النهائية بعد تدريجه، بالنسبة للاختبار ككل (DTF)، والشكل رقم (38) الذي يوضح منحنى بياني للأداء التفاضلـــي لمفردات اختبار مستويات التفكير الهندســـي، بحسب مفرداته (DIF) بين المجموعة المرجعية (الذكور)، والمجموعة المستهدفة (الإناث). والتي تعتبر مؤشـــرات تدل على تساوي القدرة فيما يتعلق

بالتكوين الفرضـــي الذي صمم الاختبـــار لتقييمه. وبناءً على هذه المؤشرات يمكننا الإحتفاظ بجميع مفردات الاختبار وبكل إطمئنان، لأن المفردات لم تبدي أداءً تفاضلياً يعزى لمتغير الجنس.

وباستخدام طريقة حساب معامل الارتباط بيرسون بين قيم بارامتر الصعوبة للمجموعتين المرجعية و المستهدفة والتي سجلت القيمة 0.98، وإلى معامل تحديد بلغ القيمة 0.96، وهو ما يشير إلى أنه معامل ارتباط قوي يترجم العلاقة التفسيرية القوية بين الجنسين في أدائهم على مفردات اختبار مستويات التفكير الهندسي. وهذا ما يعني ضمنياً لأمرين: الأول هو تأكيد افتراض أحاديـــة البعد للاختبار؛ والآخر تحقق خاصية اللاتغاير (الثبات)، و يشير كذلك إلى عدم وجود تغير في فترات الثقة حول معالم الاختبار، وفقاً لنموذج راش، وبالتالي فإن تحقق خاصية اللاتغاير تعني ضمنياً عدم وجود أداء تفاضلي لمفرات الاختبار. مما يعني أنه لن يكون هناك ظلم كبير بين المجموعتين في الأداء على الاختبار، وهي نتيجة مهمة في عملية تطوير الباحث لاختبار مستويات التفكير الهندســـي وفقاً لنموذج (فان هيل)، لهذا يمكننا الإطمئنان في استخدامه بدقة لتقييم مدى امتلاك تلامذة المرحلة الثانوية الشعب العلمية للقدرة على التفكير الهندســـي في ضوء نظرية (فان هيل)، وكذلك في إتخاذ قرارات مهمة عن واقع تدريس الرياضيـــات وبالأخص الجزء المتعلق بالهندســـة في المنهاج الجزائري، وعن مدى ملائمة المناهج الحالية لهذه المادة ولقـــدرات التلامذة، ومدى تدريسها وفقاً لنموذج (فان هيل) لمستويات التفكير الهندســـي.

اتفقت هذه النتائج جزئياً مع نتائج دراسة يوسيسكين بمشروع جامعة شيكاغو(1982)، والتي توصل فيها إلى أنه لا توجد فروق ذات دلالة إحصائية بين الجنسين في الأداء والقدرة في تعلم الهندســـة من الحقائق حتى البرهـــان. و مع دراسة كليمنتس و أخرون (Clements et al, 1997) التي توصل فيها عند دراسة تطور التفكير المكانـــي للتلامذة من خلال العمل على " مفهوم المساحة والحركة في الهندسة، باستخدام الكمبيوتر" أن كلا الجنسين يستفيدون من أنشطة تتطلب قدرة مكانية. و مع دراسة (سيف خليل، 2017) التي توصل فيها بعد تدريج اختبار مستويات التفكير الهندسي لفان هيل بأنه لا توجد فروق ذات دلالة إحصائية بين الذكور والإناث في مستوى التفكير الهندســـي.

في حين أن هذه النتائج اختلفت مع ما توصلت إليه دراسة كلٍ من غامر وإنغلهارد (1999) والتي هدفت للكشف عن الأداء التفاضلي للمفردة تبعا لمتغير الجنس، بعد تطبيق اختبار للقدرة الرياضية، والتي أظهرت تفوق الذكور على الإناث في موضوعي الهندســـة والقياس وتحليل البيانات، بينما تفوق الإناث

على الذكور في موضوع الجبر. و مع دراسة هاريس وكارلتون (1993)، بعد تطبيق اختبار الاستعداد المدرسي(SAT) والمكون من مفردات في الهندسة والحساب، والذي استخدم طريقة (مانتل–هانزل) للكشف عن الأداء التفاضلي للمفردة تبعاً لمتغير الجنس، والتي خلصت إلى أن متوسط علامات الذكور كان أعلى منه عن متوسط علامات الإناث على الاختبار وبدلالة إحصائية، كما كانت مفردات الاختبار ككل أسهل على الذكور منه على الإناث.

في حين اختلفت مع ما توصلت إليه دراسة كاثرين وشوان (2001)، والتي هدفت للكشف عن المفردات التي تظهر أداءً تفاضلياً بحسب متغير الجنس بعد تطبيق اختبار مستويات الرياضيات، والتي تبين فيها أن مفردات موضوع الجبر كانت أسهل على الجنسين، في حين أن في موضوعي الهندسة والمثلثات كانت أسهل لدى الذكور منه لدى الإناث. ونفس النتائج تقريباً توصلت إليها دراسة عنابي ودودين (2006) والتي هدفت لتقصي المفردات التي تظهر أداءً تفاضلياً بحسب متغير الجنس بعد تطبيق اختبار المهارات الرياضيات للدراسة الدولية (TIMMS, 1999) على عينة اردنية، والتي استخدم فيها طريقة (مانتل–هانزل) للكشف عن الأداء التفاضلي، تبين أن أغلب المفردات التي أظهرت أداءً تفاضلياً لصالح الإناث تنتمي لموضوع الجبر والإحصاء، أما المفردات التي كانت سلبية باتجاه الإناث فقد كانت موضوعاتها تتطلب المخاطرة في الحل مثل موضوعات التقدير، التوقع، التقريب، كما أن أغلب المفردات التي أظهرت أداءً تفاضلياً لصالح الذكور تنتمي لموضوع الهندسة والقياس.

لم تتفق هذه النتائج مع ما توصلت إليه دراسة عبد العزيز(2010)، والتي كان هدفها الكشف عن الأداء التفاضلي لمفردات اختبار في الرياضيات تبعاً لمتغير الجنس بماليزيا، والتي استخدم فيها ثلاث طرق للكشف، وهي: طريقة مانتل– هانزل؛ والصعوبة المحولة؛ وفرق معلم الصعوبة (منحنى خصائص المفردة)، والتي كان فيها التوافق الأكثر بطريقة مانتل–هانزل، بينما كان التوافق الكلي بين طريقتي الصعوبة المحولة وفرق معلم الصعوبة، وبشكل عام أظهرت النتائج أن مفردات الهندسة تبدي أداءً تفاضلياً لصالح الذكور، يضاف إلى ذلك بعض المفردات التي كانت تقيس المفاهيم المجردة الجبرية والتعريفات والحساب العددي، بينما أظهرت مفردات الجبر وبعض التطبيقات الحياتية أداءً تفاضلياً لصالح الإناث. ونفس النتيجة تقريباً توصلت إليها دراسة معتصم العكور وأخرون (2005) والتي هدفت للكشف عن الفروق بين أداء الذكور والإناث، لدى العينة الأردنية على أحد الاختبارات الدولية المتعلقة

بالرياضيات (PISA, 2012) والتي كانت نتائجها مغايرة تماماً لنتائج معظم الدول الأخرى المشاركة في هذا الاختبار، و هي تفوق الإناث على الذكور وبدلالة إحصائية في الأداء الكلي، بالإضافة إلى تفوقهن في معظم مهارات محتوى الرياضيات ومعظم مهارات العمليات الرياضية، أما في ما يخص أداء التلامذة الأردنيين على مستوى الكفاءات الست والتي يقيسها الاختبار فقد أشارت النتائج إلى عدم وجود فروق دالة إحصائياً بين الجنسين في المستويات العليا (من المستوى الثالث إلى السادس). وهو ما أشارت اليه كذلك دراسة عبد الله شطناوي وحسان غازي العمري (2014) عند تقصي الأداء التفاضلـــي لمفردات الاختبار الوطني الأردني لمادة الرياضيـــات للصف العاشر تبعا لمتغير الجنس، باستخدام (مؤشر الأداء التفاضلي اللاتعويضي NCDIF)، عن وجود أداء تفاضلي منتظم لـــ 9 مفردات من أصل 16 مفردة وهي عدد المفردات المشتركة بين الطلاب والطالبات، وعن وجود أداء تفاضلـــي غير منتظم لـــ 15 مفردة من أصل 26 مفردة وهي عدد المفردات المشتركة بين الطلاب والطالبات.

في حين تبين من خلال نتائج دراسة التحصيل الوطني في الرياضيـــات في الولايات المتحدة (Fennema and Carpenter, 1981) أن هناك ضعف عام في أوساط التلميذات مقارنة بالذكور من نفس العمر 17 سنة على جميع المستويات المعرفية في الهندسة بشكل خاص. وإحدى التفسيرات الممكنة والتي قدمت لهذا الضعف هو أن تمارين أو أسئلة الهندســـة ترتبط بمهارات التصور المكاني، حيث أن أداء الإناث في هذه المهارة أضعف منه من الأداء لدى الذكور.

وعليه وبناء على ما سبق من نتائج، يمكننا القول أن الفرضية محققة، والتي تشير إلى أن تدريج صعوبة مفردات اختبار مستويات التفكير الهندسي في صورته النهائية بعد التدريج، لا يظهر أداءً تفاضلياً تبعاً لمتغير الجنس (الذكور، والإناث). وبالتالي فإن عدم وجود أداء تفاضلي لمفرات الاختبار يعني ضمنياً على أن اختبار مستويات التفكير الهندســـي يحقق خاصية اللاتغاير (الثبات)، وإلى صدق المفهوم.

1-9- تفسير ومناقشـــة نتائـــج الفرضية التاسعـــة:

والتي تنص علـــى أن:" يبلغ تلامذة المرحلة الثانوية مستويات متقدمة في مستويات (فان هيل) للتفكير الهندســـي بعد تطبيق اختبار مستويات التفكير الهندســـي في صورته النهائية".

تبين من خلال عرض النتائج الموضحة في الجدول رقم (51)، إلى أن ما نسبته 78% من تلامذة السنة الأولى ثانوي؛ و ما نسبته 88.9 % من تلامذة السنة الثانية ثانوي؛ و ما نسبته 85.8 % من تلامذة

السنة الثالثة ثانوي، أي ما نسبته84.61% من العينة الكلية، أمكن تصنيفهم على مستويات (فان هيل)، بمعنى أن النظريـــة نجحت في تصنيفهم. في حين أن ما نسبته 22% من تلامذة السنة الأولى ثانوي؛ و11.10% من تلامذة السنة الثانيـــة ثانوي؛ و14.20% من تلامذة السنة الثالثـــة ثانوي، أي ما نسبته 15.37% من العينة الكلية لم يصنفوا حسب معايير النظرية والمحكات المستخدمة في بحثنا الحالي، وهذه نقطة قوية لصالح النظريـــة. أما النقطة الضعيفة في النظرية فهي أن تصنيف التلامذة إلى مستويات يعتمد على المعيار أو المحك المستخدم في الحكم على تمكن التلميذ من المستوى أم لا.

وأشارت نتائج التحليل كذلك، إلى أن توزيع التلامذة على المستويـــات الأربعة الأولى لـــ(فان هيل) كان بشكل هرمي نوعا ما، ويظهر هذا التوزيع بشكل واضح في المستوى الأول مع تداخل طفيف في ما بين المستويات الثلاثة الأخرى. وهو ما يتفق مع أحد الخصائص الأساسية لنظرية (فان هيل) لمستويات التفكير الهندسي. حيث يرى فان هيل (1986) في هذا السياق، أن النمو من خلال المستويات ليس متداخلاً وإنما فجائياً، في حين أن بعض الباحثين عارضه في ذلك وأيده آخرون. فقد وجد بيج وديفري (Pegg and Davery, 1998) تبريراً للفصل بين المستويـــات واعتبراه منطقياً إذا ما أخذ بالنظرة العريضة والمثالية للنظرية، أما بورغر وشوغنيسي (Burger and Shaughnessy, 1986)، وكذلك فيوز وآخرون (Fuys et al, 1988) فوجدوا أن التلامذة يمكن أن يتحركوا داخل المستويـــات جيئة وذهاباً، وهو ما يتفق وتدعمه نتائج بحثنا الحالي.

وبالعودة لنتائـــج الجدول رقم (51)، والشكل رقم (39)، واللذان يظهران أن مستويات التفكير الهندســـي في ضوء نظرية (فان هيل) لغالبية تلامذة المرحلة الثانوية كان متدنياً جداً، والتي توزع فيها التلامذة في المستويين الأول و الثاني من مستويات (فان هيل) بما نسبته 60.06%، حيث حقق 272 تلميذاً وتلميذة فقط من أصل 566 تلميذاً وتلميذة بما نسبته 47.34% المستوى الأول من مستويات (فان هيل) للتفكير الهندسي وهو المستوى البصري أو الإدراكي؛ في حين حقق 72 تلميذاً فقط من أصل 566 تلميذاً وتلميذة وما نسبته 12.72% المستوى الثاني وهو المستوى التحليلي أو الوصفي؛ وحقق 77 تلميذاً فقط من أصل 566 تلميذاً وتلميذة وما نسبته 13.6% المستوى الثالث وهو المستوى شبه الاستدلالي؛ كما حقق 62 تلميذاً فقط من أصل 566 تلميذاً وتلميذة وما نسبته 10.95% المستوى الرابع وهو المستوى الاستدلالي. في حين تبين أن 83 تلميذاً وتلميذة من أصل 566 تلميذاً وتلميذة، أي ما نسبته 14.66% من عينة البحث

لم يصنفوا بحسب نظرية (فان هيل) للتفكير الهندســـي. وأشارت العديد من البحوث والدراسات السابقة والتي استخدمت اختبار مستويات التفكير الهندسي لـــ (فان هيل) أن المستوى الرابع نادراً ما كان يظهر لدى التلامذة في المراحل التعليمية الثلاثة قبل الجامعية (الإبتدائية؛ المتوسطة؛ والثانوية)، في حين أن المستوى الخامس لا يظهر إطلاقاً في هذه المراحل.

و يفسر الباحث نتيجة أن معظم عينة البحث من تلامذة المرحلة الثانوية، قد حققوا المستوى الأول وهو بما يسمى المستوى البصري أو الإدراكي وبنسبة 47.34%، بأن التلامذة في هذا المستوى قادرين فقط على القيام بالعمليات المتعلقة بالمستوى البصري كملاحظة الصورة أو الشكل الهندســـي ككل دون إدراك لخواصه أو عناصره، فهم يتعرفون على شكل المستطيل لأنه يشبه الباب مثلاً، وهذا نفس الشكل لأنه يبدو متشابه، وهذا مربع وليس مستطيل لأن المستطيل أطول، ويكون توجه التلامذة بقوة في هذا المستوى نحو الشكل، كما يميلون إلى تحديد المربع فقط من جانبيه الأفقي والعمودي، كما يمكنهم أن يتعلموا المصطلحات بشرط أن تكون في لغة محسوسة، وبأسلوب ملموس من خلال التعرف على هيئة الشكل وهو في أوضاع مختلفة وتحديد بعض الأشكال الهندسية في صورتها الفيزيقية أو المرسومة، و رسم بعض الأشكال الهندسية البسيطة، كما يسمون أشكالاً بأسماء عامة، و يميزون بين الأشكال بحسب مظهرها ويصفوها بالكلام، و يتعرفون على أجزاء شكل و ينظرون لكل شكل أن أضلاعه مستقيمة وشكل محيطه على شكل منحنيات، ولكنهم لا يميزون بين الأشكال من نفس النوع.

أما بالنسبة للمستويات الثلاثة الأخرى والتي أطلقت عليها (دينا فان هيل) مسميات: شكل الهندسة، جوهر الهندسة، التبصر في نظرية الهندسة، التبصر العلمي في الهندسة. فقد أظهرت النتائج تدني مستوى التفكير الهندســـي لدى أفراد عينة البحث في هذه المستويات، حيث لم يتجاوز المستوى الثاني التحليلي أو الوصفي ما نسبته 12.72% من مجموع التلامذة، وما نسبته 13.6% كذلك بالنسبة للمستوى الثالث شبه الاستدلالي، في حين لم يحقق المستوى الرابع أي الاستدلالي سوى 10.95% فقط من عينة البحث. وهذا ما يعني أن معظم تلامذة المرحلة الثانوية الشعب العلمية في هذه المستويات، يعانون من صعوبة في توظيف العلاقـــات المتداخلة بين خصائص الأشكال الهندســـة لإعطاء براهين معينة للوصول إلى حل المسألة الهندســـة أو كتابة التعاريف الهندسية وبرهنة بعض المسائل الهندسية أو إكمال برهان هندســـي معين، أو أن يعي التلميذ المسلمات والقيام بالاستدلال الشكلـــي في فهم العلاقات المتداخلة بين النظريات،

أو أن يكتب التلميذ برهاناً قائماً على الرموز الهندسيـــة ويستبعد الشروط غير الضرورية أو الكافية في برهنة مسألة هندسيـــة أو القيام باستنتاج نظريات هندسية معتمدة على مسلمـــات سبق للتلميذ معرفتها، وإجراء عمليات مقارنة بين تلك المسلمـــات لإكتشاف مسلمات جديدة، وهو مؤشر على عدم توظيفها التوظيف المناسب ضمن برنامج أو منهاج الرياضيات الحالي الموجه لتلامذة المرحلة الثانوية وخاصة بشقه المتعلق بكيفية تدريس الهندســـة. مما يعني كذلك أن المفاهيم التي تعرض لها التلامذة عند دراستهم لمادة الرياضيات في الجزء المتعلق بالهندســـة، لم تنمي مستويات التفكير الهندسي لديهم في شكل تسلسل هرمـــي، مما يدل على أن التلميذ لا يستطيع التقدم من مستوى تفكير إلى مستوى أعلى منه وتقديم الأداء المطلوب منه، إلا إذا كانت لديه خبـــرات كافية في المستوى السابق له والتي تؤهله إلى المستوى الجديد، وقد يعود كذلك إلى عدم ملائمة موضوعات الهندســـة وأساليب تدريسها لقـــدرات التلامذة وميولهم واستعداداتهم، أو عدم وجود تفاعل بين المعلم والمتعلم، ووجود حاجز لغوي بينهما كما عبر عن ذلك (فان هيل)، وهذا يتفق مع ما توصلت إليه عديد الدراسات السابقة في هذا الموضوع، و أكدت عليه نظرية (فان هيل).

وبالعودة للدراسات السابقـــة تبين أن نتائجنا اتفقت جزئياً مع ما توصلت إليه دراسة يوسيسكين في مشروع جامعة شيكاغو (1982)، والتي كان معظم عينتها من تلامذة الصفين العاشر56% والحادي عشر26%، والتي من أبرز نتائجها أن ما نسبته 71% من التلامذة الذين تقدموا لاختبار (فان هيل)، أمكن تصنيفهم على مستويات (فان هيل)، حيث توزعوا على مستويات التفكير الهندسي الخمسة، كالتالي: ما نسبته 32% على المستوى الأول؛ وما نسبته21% في المستوى الثاني؛ وما نسبته 9% للمستوى الثالث؛ وتسجيل نسبة 2% على المستوى الرابع؛ و ما نسبته1% على المستوى الخامس. في حين أن ما نسبته 29% من التلامذة لم يصنفوا حسب معايير هذه النظرية. كما توصلت دراسة لـــ(Gutierrez and Saine, 1998) إلى أن ما نسبته 83% من تلامذة السنة الأولى ثانوي حققوا المستوى الأول؛ و ما نسبته 38% حققوا المستوى الثاني؛ و 16% حققوا المستوى الثالث؛ و 3% حققوا المستوى الرابع. وتوزع تلامذة السنة الثانية ثانوي على المستويات الأربعة، بما نسبته 89%؛ 43%؛ 14%؛ و3% على التوالي. كما توزع تلامذة السنة الثالثة ثانوي على هذه المستويات، بما نسبته 90%؛ 57%؛ 20%؛ و4% على التوالي. ومع دراسة عياصرة (2002) التي توزع فيها تلامذة المرحلة الأساسية العليا في الأردن على مستويات

(فان هيل) الأربعة الأولى بما نسبته 36.5% في المستوى الأول؛ و24.9% في المستوى الثاني؛ و12.9% في المستوى الثالث؛ و0.8% في المستوى الرابع؛ وما نسبته 24.9% من التلامذة لم يصنفوا وفق المعايير التي استخدمها في دراسته.

في حين لم تتفق نتائجنا مع نتائج دراسة الطيطي (2001) التي أجراها على عينة من تلامذة سنة الأولى ثانوي بفلسطين، والتي توزع فيها أداء التلامذة على مستويات (فان هيل) بما نسبته 14% في المستوى الأول؛ و46.2% في المستوى الثاني؛ و14.4% في المستوى الثالث؛ و15.5% في المستوى الرابع؛ و3.4% في المستوى الخامس؛ وما نسبته 6.4% من التلامذة لم يصنفوا وفق المعايير المستخدمة.

وبشكل عام، لقد أجمعت غالبية الدراسات والبحوث السابقة في هذا المجال، بأن موضوع الهندسة يعتبر من أكثر فروع الرياضيات التي يواجه تعلمها صعوبات بالنسبة للتلامذة، وأن أغلبها أرجع هذه الصعوبات إلى الطرق والأساليب التدريسية التقليدية التي يتم من خلالها تعليم موضوعات الهندسة، والتي لم تعتمد نظريات أثبتت نجاحها على هذا الصعيد على غرار نظرية (فان هيل) لمستويات التفكير الهندسي، والذين أكدوا على ضرورة استخدام استراتجيات تدريسية مناسبة للخبرات وجعلها ممتعة في تدريسها ودراستها نظرياً وعملياً بما يتماشى مع طبيعة التلامذة وقدراتهم. وهو ما أكد عليه كذلك يوسيسكين (1987)، بأن خبرات معلمي الرياضيات مازالت توحي بوجود صعوبات لدى تلامذتهم في تعلم الهندسة وخاصة عندما يتطلب الأمر القيام بالإثباتات والبراهين الهندسية لإثبات الحقائق. فمعظم الكتابات في الهندسة في مراحل التعليم العام تتبع من جهتين، أولاً ضعف أداء التلامذة بشكل عام في الهندسة، وثانياً في المناهج المتبعة و المنتهية الصلاحية والتاريخ.

وعليه وبناء على ما سبق من نتائج، يمكننا القول بأن الفرضية محققة جزئياً، والتي تشير إلى أنه يبلغ تلامذة المرحلة الثانوية مستويات متقدمة من مستويات (فان هيل) للتفكير الهندسي من خلال تطبيق اختبار مستويات التفكير الهندسي في صورته النهائية.

خلاصة نتائج البحث:

نظراً لأهمية الاختبارات بشكل عام، واختبارات التحصيل واختبارات مستويات التفكير الهندسي المصممة وفق نموذج (فان هيل) بشكل خاص، اهتم الباحثون في مجال القياس والتقويم النفسي والتربوي، باجراء العديد من الدراسات والبحوث المتعلقة ببناء مفردات هذه الاختبارات وانتقائها بالاعتماد على نظرية القياس الكلاسيكية، ونتيجة قيام الكثير من العلماء مثل ثيرستون (Thurstone, 1995)، ولورد (Lord, 1953)، وبيرنبوم (Birnbaum, 1958) بتوجيه مجموعة من الانتقادات لنظرية القياس الكلاسيكية، بسبب قصورها في تحقيق الموضوعية في القياس، والمتمثلة في اعتماد القياس على خصائص الاختبار المستخدم (الصدق، والثبات، والموضوعية)، أي بمعنى آخر اختلاف نتيجة القياس باختلاف الاختبار المستخدم، وكذلك اختلاف نتيجة القياس باختلاف مستوى الأفراد الذين يطبق عليهم الاختبار، بالإضافة إلى عدم خطية القياس؛ أي عدم وجود وحدات متساوية على مدى متصل القدرة التي يتدرج عليها المتغير موضوع القياس. ظهرت عدة نماذج لنظرية الاستجابة للمفردة تسمح بالقياس الموضوعي ومنها نموذج (راش) أحادي البارامتر، والتي يتم فيها التركيز على المفردات الاختبارية بدل الاختبار ككل، مما يسمح باضافة مفردات، أو حذفها، أو تعديلها، دون أن يتأثر الاختبار ككل، مما يساعد في التغلب على مشكلات القياس الكلاسيكي نظرياً وتطبيقياً .

وانطلاقاً من هدف بحثنا الحالي، وهو مقارنة فعالية استخدام نموذج (راش) أحادي البارامتر ونظرية القياس الكلاسيكي في بناء اختبار موضوعي تشخيصي محكي المرجع في ضوء نموذج (فان هيل) لمستويات التفكير الهندسي، لتفسير أداء تلامذة المرحلة الثانوية (الشعب العلمية) على هذا الاختبار، ثم التحقق من موضوعية هذا التفسير كما يتمثل في فرضيات نموذج (راش)، وتقييم مدى صدق النموذج في تحقيقه لموضوعية القياس.

خلص الاختبار في صورته النهائية بعد ما تمت إجراءات بنائه وتحليل وتدريج مفرداته باستخدام المقاربة الكلاسيكية في القياس إلى 28 مفردة، وإلى 22 مفردة باستخدام نموذج راش، أي حذف 6 مفردات و11 فرداً، كانت مؤشرات المطابقة لهما خارج حدود المطابقة الداخلية والخارجية المطلوبة. كما اقتربت قيمة إحصائي المطابقة الخارجية للمفردات بعد هذا الحذف من الصفر، والانحراف المعياري لهما من الواحد، مما دل على أن هناك اتساقاً بين الاستجابات الملاحظة للافراد عن المفردات ودرجاتهم

الكلية على الاختبار، وهذا يشير إلى أن هناك اتفاقاً بين مستويات التفكير الهندســـي التي تعبر عنها هذه المفردات والقدرة التي تعبر عنها باقي المفردات، وذلك ضمن العينة كلها، ومعنى هذا مطابقة المفردات وتمتع الاختبار بدرجة عالية من الثبــات، كما تمتع الاختبار بالصــدق في قياس أربعة مستويات متمايزة نوعاً ما. كما أشارت النتائج كذلك إلى تحرر القيــاس من خصائص توزيع القدرة لأفراد العينة، أي عدم اختلاف قيم الصعوبة النسبية للمفــردات عند معظم الأفراد، وعبر المستويات المختلفة للقدرة أو السمة.

وعلى أساس ما تم التطرق له في الفصول النظريــة، جاءت نتائج الفصول الميدانيــة للبحث موافقة وداعمة له تماماً، وعليه أمكن للباحث وباستخدام التحليل الإحصائــي وفقاً للمقاربتين الكلاسيكية ونموذج (راش) في القياس، أن يتحقق إجرائياً من الفرضيات التي انطلق منها، والتي أسفرت نتائجها، على التالي:

1. توفر اختبار مستويات التفكير الهندسي على خصائص سيكومترية مقبولة باستخدام النظرية الكلاسيكية في القيــاس. بالرغم من اعتمادها في الحكم على هذه الخصائص من خلال الدرجة الكلية للاختبار وليس على درجة كل مفردة على حدة.

2. توفر بيانات اختبار مستويات التفكير الهندسي المستمدة من استجابات تلامذة المرحلة الثانوية (الشعب العلمية) على ملاءمة مقبولة لشروط تطبيق نموذج (راش). والتي سمحت للباحث بعملية تدريج الاختبار لاحقاً على أفراد عينة البحث وفقاً لنموذج (راش) أحادي البارامتر.

3. تباينت تقديرات صعوبة مفردات اختبار مستويات التفكير الهندســي وتدرجها باستخدام نموذج راش. والتي سمحت بتمثيل قدرات الأفراد ومواقع المفردات على تدريج خطــي متصل فتري بحسب صعوبتها، من خلال تحديد التناظر بين البيانات المشاهدة ومواقع الأشخاص على المتغير الكامن.

4. تباينت تقديرات قدرات الأفراد ودالــة المعلومــات المقابلة لكل درجة كلية خام على مفردات اختبار مستويات التفكير الهندســي باستخدام نموذج راش. وبالتالي تحققت موضوعية تقدير القدرة لهذا الاختبار الذي تم تدريجه باستخدام نموذج راش.

5. حقق تدريج اختبار مستويات التفكير الهندســي باستخدام نموذج راش استقلالية القياس، باختلاف عينة التدريج والاختبار المستخدم.

6. أمكن اشتقاق معاييــر كمية تفسر تقديرات الأفراد على اختبار مستويات التفكير الهندسي في صورته النهائية بعد تدريجه على تلامذة المرحلة الثانوية.

7. تحقق الصدق والثبات لاختبـــار مستويات التفكير الهندســـي في صورته النهائية بعد تدريجه باستخدام نموذج (راش)، وهو ما يعني أنه يوفر متطلبات الموضوعية في قيـــاس متغير القدرة موضوع القيـــاس.

8. أشارت نتائج تدريج صعوبة مفردات اختبار مستويات التفكير الهندســـي في صورته النهائية، بأنه لا يظهر أداءً تفاضلياً تبعاً لمتغير الجنس (الذكور، والإناث). وبالتالي عدم وجود أداء تفاضلـــي لمفرات الاختبار يعني ضمنياً على أن اختبار مستويات التفكير الهندســـي يحقق خاصية اللاتغاير أي (الثبات)، وكذلك إلى صدق المفهوم.

9. أشارت عملية تحليل استجابات تلامذة المرحلة الثانوية (الشعب العلمية) بعد تطبيق اختبار مستويات التفكير الهندســـي عليهم في صورته النهائية إلى مستويات متدنية وفقاً لنموذج (فان هيل) لمستويات التفكير الهندســـي.

أثبتت نتائج البحث من خلال استخدام نموذج (راش) ونظرية القياس الكلاسيكـــي، في عملية بناء اختبار موضوعي تشخيصي محكي المرجع لمستويات التفكير الهندســـي في ضوء نموذج (فان هيل)، أفضلية نموذج (راش) في اختبار ثبات الاختبار وصدقه، من خلال دراسة مساهمة كل مفردة أو فرد على حدة في جودته، حيث تدرجت مفرداته بوحدة قيـــاس مطلقة، تصل بنا إلى التفسير الموضوعـــي لدرجة التلميذ على الاختبار، أو أي اختبار فرعي آخر مأخوذ من هذا الاختبار. والذي يصبح بمثابة بنـــك لأسئلة اختبارات مستويات التفكير الهندســـي المصممة وفقاً لنموذج (فان هيل) مستقبلاً، بحيث نستطيع مستقبلاً أن نسحب منه أي عدد من المفردات التي تحقق أهداف القيـــاس، كما يمكننا أيضاً أن نضيف إليه مفردات جديدة تشترك معه في تدريج واحد مشترك وصفر واحد مشترك. وبهذا يمكن التغلب على مشكلة من أهم مشكلات تقويم التحصيل الدراســـي، وهي موضوعية التقدير بالرغم من اختلاف الاختبار المستخدم في القياس، كما يصل بنا أيضاً للتفسير الموضوعي لأداء التلميذ بالرغم من اختلاف مستوى المجموعات أو الشعب التي ينتمي إليها، وبذلك يمكن أن نقترب باحتمال كبير إلى العدالـــة والموضوعيـــة في تقديرنا للتحصيل الدراسي في مستويات التفكير الهندســـي للتلميذ بل والأداء السلوكي للتلامذة بوجه عام.

توصيات البحث:

في ضوء ما توصل إليه بحثنا الحالـــي، ومن أجل تجسيد هذه النتائج ميدانياً، نتوجه بالتوصيات التالية:

1. استغلال اختبار مستويات التفكير الهندســـي الذي تم بناءه على ضوء نموذج (فان هيل)، والمطور باستخدام نموذج (راش) أحادي البارامتر، لما يتمتع به هذا الأخير من موضوعية في القياس بعد تدريجه على تلامذة المرحلة الثانوية (الشعب العلمية)، في تشخيصهم لصعوبات تعلم الهندســـة، وتصنيفهم وفقاً لمستويات (فان هيل) للتفكير الهندسي، نظراً لأهمية هذه المرحلة التعليمية التتويجية قبل التعليم الجامعي.

2. إعادة تصميم وبناء مناهج الهندسة في مراحل التعليم العام، وفق نموذج (فان هيل) في تدريس الهندسة.

3. استخدام نماذج نظرية الاستجابـــة للمفردة وخاصة نموذج (راش) نظراً لبساطة افتراضاتـــه وسهولة التعامل معه، في تطوير وبناء اختبارات للقدرات العقلية والتحصيلية في ضوء نماذج متطورة في القياس التربوي، للتغلب على أوجه النقد الموجه لهذا النوع من الاختبارات التي تم تطويرها وبنائها وفق نظرية القياس الكلاسيكـــي.

4. تنظيم دورات تدريبية على البرامج الإحصائية المستخدمة في تحليل بيانات نماذج نظرية الاستجابة للمفردة، يشرف عليها خبراء القيـــاس والتقويـــم والإحصاء موجهة لكافة الأساتذة وطلبة الدراسات العليا.

5. التفكير في إنشاء مركزاً للقيـــاس والتقويـــم بوزارة التعليم العالي والبحث العلمي بالجزائر على غرار مركز المملكة العربية السعودية ومركز جمهورية مصر العربية، مهمته إعداد وتطوير الاختبارات العقلية والتحصيلية، وبنوك للأسئلـــة، باستخدام الأساليب العلمية المتطورة في هذا المجال.

6. تكوين الأساتذة والمعلمين خاصة خريجـــي المدارس العليا للأساتذة، في مجال بنـــاء الاختبارات التحصيلية المدرسية، وإطلاعهم على التطورات الحاصلة في هذا المجال على المستوى العالمي.

مقترحـــــات البحث:

انطلاقاً من النتائج المشجعة المتوصل إليها في بحثنا الحالـــي، ومن أجل جمع المزيد من المعلومات حول خصائص اختبار مستويات التفكير الهندسي الذي تم بناءه على ضوء نموذج (فان هيل)، نتوجه بالاقتراحات التالية:

1. إجراء بحوث ودراسات سيكومترية مقارنة لاختبار مستويات التفكير الهندسي المبني على ضوء نموذج (فان هيل) باستخدام النظرية الكلاسيكيـــة في القيـــاس ونموذج راش، لدى تلامذة المرحلتين التعليميتين (الإبتدائية، والمتوسطة).

2. إجراء بحوث ودراسات للخصائص السيكومترية لاختبار مستويات التفكير الهندسي المبني على ضوء نموذج (فان هيل) باستخدام نماذج نظرية الاستجابة للمفردة (النموذج ثنائي البارامتر لبيرنبوم؛ و الثلاثي البارامتر للورد).

3. إجراء بحوث ودراسات سيكومترية مقارنة لاختبار مستويات التفكير الهندسي المبني على ضوء نموذج (فان هيل)، باستخدام نموذج راش أحادي البارامتر، ونموذج بيرنبوم ثنائي البارامتر، ونموذج لورد ثلاثي البارامتر، من حيث فعالية ودقة قياس وموضوعية كل نموذج.

4. إجراء بحوث ودراسات لتقصي الأداء التفاضلي لمفردات اختبار مستويات التفكير الهندسي المبني على ضوء نموذج (فان هيل)، تبعاً لمتغير (العمر، والمستويات والمراحل الدراسية).

خاتمـــة البحث:

رغم أصالة نتائج البحث الحالي في البيئة العربية والجزائرية وربما على المستوى العالمي، ونجاحنا في تصميم اختبار موضوعي محكي وتشخيصي يقيس مستويات التفكير الهندسي وفق نموذج (فان هيل)، وتحقيقه لاستقلالية تقديرات مفرداته عن عينة الأفراد، وبالتالي التغلب على المشكلات المتعلقة بالتقنيـــن، الأمر الذي يضفي على نتائجه نوعاً من المصداقيـــة وإمكانية تعميـــم نتائجه على أفراد متشابهين في الخصائص، يبقى هذا البحث يحتاج إلى مزيد من التمحيص والتدقيق والتأكيد. إذ ليس من السهل التحكم في مفهوم مجرد مثل متغير التفكير الهندسي، وكذلك متغيرات البحث الكامنة المستقلة في نموذج (راش) المتمثلة في قدرة الفرد، وصعوبة المفردة، والمتغير التابع متمثلاً في احتمال إجابة الفرد إجابة صحيحة على المفردة، من خلال تطبيق الاختبار مرة واحدة وما يتبعه من عدم الجديـــة في الاستجابـــة لبعض أفراد العينة، وكذلك الظروف النفسية والفيزيقية والمدرسية المؤثرة في أداء التلامذة على هذا الاختبار.

المـــــراجع

قائمـــة المراجع

المراجـــع:

1- أناستازي وسوسن، أرينا (2015). القيـــاس النفســـي. (صلاح الدين محمود علام، مترجم).عمان: دار الفكر.

2- أمين على محمد سليمان، مراجعة (رجاء محمود أبو علام) (2009). القيـــاس والتقويـــم في العلوم الانسانية: أسسه وأدواته وتطبيقاته. القاهرة: دار الكتاب الحديث.

3- أبو ناهية، صلاح الدين (1994). القيـــاس التربـــوي. مصر: مكتبة الآنجلو مصرية.

4- أحمد، محمد عبد السلام (1960). القيـــاس النفســـي والتربوي. القاهرة: مكتبة النهضة المصرية.

5- أبوعلام، رجاء محمود (2006). مناهج البحث في العلوم النفسيـــة والتربويـــة.القاهرة: دار النشر للجامعات.

6- أبو ملوح، محمد (2002). تنمية التفكير في الهندســـة واختزال القلق نحوها لدى طلبة الصف الثامن الأساسي بمحافظة غزة في ضوء مدخل فـــان هايل ومخططات المفاهيـــم. رسالة دكتوراه. كلية التربية جامعة عين شمس.

7- أبو لوم، خالد (2005). الهندســـة وأساليب تدريسها. عمان: دار المسيرة للنشر والتوزيع.

8- أيمن سليمان، سلمان القهوجي (2014). تطوير اختبـــار في مهارات البحث العلمي لطلبة الجامعات والتحقق من خصائصه السيكومترية وفق النظرية الكلاسيكية في القياس ونموذج راش. رسالة ماجستير غير منشورة، الجامعة الأردنية.

9- بشير، معمرية (2007). القياس النفسي وتصميم أدواته للطلاب والباحثين.ط2.الجزائر:منشورات الحبر.

10- البنا، مكة (1994). برنامج مقترح لتنمية التفكير في الهندســـة في ضوء نموذج فان هايل. رسالة ماجستير غير منشورة. كلية التربية. جامعة عين شمس.

11- بوسالم، عبد العزيز (2008). توظيف نموذج راش أحـــادي البارامتر في بناء اختبار تحصيلي في القيـــاس النفسي وتحقيق التفسير الموضوعي لنتائجه بالمقارنة مع النظرية الكلاسيكية: دراسة سيكومترية

مقارنة بين نظرية السمات الكامنة والنظرية الكلاسيكية في بنـــاء الاختبـــارات والتحقق من موضوعيتها. أطروحة دكتوراه غير منشورة. كلية العلوم الانسانية والاجتماعية. جامعة الجزائر.

12- بوسالم، عبد العزيز (2014). القيـــاس في علم النفس و التربيـــة الأسس النظرية والمبادئ التطبيقية. منشورات مخبر القياس والدراسات النفسية. الجزائر: قرطبة للنشر والتوزيع.

13- بوسالم، عبد العزيز و كريش، أحمد (2015). الأداء التفاضلي للبند في الاختبارات النفسية مصادره وتفسيره . مجلة دراسات نفسية وتربوية . العدد (12)، ص61-78.

14- بوسنة، محمود (2007). علـــم النـــفس القياســـي المبادئ الاساسية. الجزائر: ديوان المطبوعات الجامعية.

15- بيكر، فرانـــك (2010). أسس نظريـــة الاستجابـــة للمفردة (الطريري، أبوهاشم، سوسن، مترجم). السعودية: النشر العلمي والمطابع جامعة الملك سعود .

16- التقي، أحمد محمد (2013). النظرية الحديثة في القياس. ط2. عمان: دار المسيرة للنشر والتوزيع.

17- تيغزة، أمحمد (2009). البنية المنطقية لمعامل ألفـــا كرونباخ ومدى دقته في تقدير الثبـــات في ضوء افتراضات نماذج القيـــاس. مجلة العلوم التربويـــة والدراسات الأسلامية. جامعة الملك سعود. 21(3). ص637-688.

18- تيغزة، أمحمد (2011). اختبار صحة البنية العامليـــة للمتغيرات الكامنـــة في البحوث: منحى التحليل والتحقق. بحث علمي محكم. قسم علم النفس. كلية التربية. جامعة الملك سعود. الرياض.

19- الشويخ، جهاد (2005). أنماط التفكير الهندسي لدى الطلبة الفلسطنيين. رسالة ماجستير غير منشورة. كلية الدراسات العليا. جامعة بيرزيت. فلسطين.

20- حسن، شوكارة و أحمد، أفينشل و أحمد، الزيتي و حميد، البجطة (2007). ملف تربـــوي تكوينـــي حول مجزوءة: الانشاءات الهندسيـــة في الاعدادي. وزارة التربية الوطنية والتعليم بالمملكة المغربية. ص01-39.

21- حمادنة، اياد محمد ذياب (2007). الأداء التفاضلي لفقرات اختبارات تحديد الكفاءة اللغوية في اللغة الانجليزية في الجامعات الأردنية: دراسة مقارنة. رسالة دكتوراه. كلية التربية. جامعة اليرموك. الاردن.

22- الحمري، هند (2011). دراسة استكشافية لملاءمة نماذج نظرية الاستجابة للمفردة. مجلة العلوم التربوية والنفسية. البحرين. 12 (2)،61-62.

23- حوالة ، مصطفي (2005). إعداد المعلم تنميته وتدريبه. عمان: دار الفكر.

24- دي إيالا، ر.ج (2017). النظرية والتطبيق في نظرية الاستجابة للفقرة.(عبد الله زيد الكيلاني وإسماعيل سلامة البرصان، مترجم). الرياض: دار جامعة الملك سعود للنشر.

25- ديانا فهمي، علي حماد (2010). علاقة الحكمة الاختبارية بالأداء الناتج من اختبار تحصيلي ذي اختيار من متعدد مبني وفق نموذج راش لدى طالبات كلية التربية للبنات الأقسام الأدبية بجامعة أم القرى. الرسالة دكتوراه غير منشورة. كلية التربية. قسم علم النفس. جامعة أم القرى.

26- دياب، سهيل (2000). تعلم مهارات التفكير وتعلمها في الرياضيات لطلبة المرحلة الابتدائية العليا. غزة : دار المنارة.

27- راضي، الوقفي (2003). مقدمة في علم النفس. عمان: دار الشروق للنشر والتوزيع.

28- رباب، الطنة (2008). تحليل محتوى منهاج الرياضيات للصف الثامن الأساسي في ضوء مستويات التفكير الهندسي لفان هيل. رسالة ماجستير غير منشورة. كلية التربية بالجامعة الاسلامية. غزة.

29- رضوان، محمد نصر الدين (2011). المدخل الى القياس في التربية البدنية والرياضة. ط2. القاهرة: مركز الكتاب للنشر والتوزيع.

30- رلى يوسف، فهد (2001). صعوبات تعلم الهندسة لدى طلبة الصف الثالث الاعدادى فى البحرين و تفسيرها فى ضوء مستويات التفكير الهندسي. مجلة العلوم التربوية و النفسية. المجلد(2). العدد(2). ص176-178.

31- سعد الله، أبوبكر خالد (2009). سند تكويني موجه لمفتشي التعليم الابتدائي لمادة الحساب والهندسة. المعهد الوطني لتكوين مستخدمي التربية وتحسين مستواهم بالحراش. الجزائر.

32- سعيد، ردمان محمد (2007). مدى اتساق محتوى الهندسة في كتب الرياضيات للصفوف من 7-9 في الجمهورية اليمنية مع الأسس التعليمية لنظرية فان هيل للتفكير الهندسي. مجلة العلوم التربوية والنفسية، البحرين. المجلد08. العدد(03). ص165-185.

33- سلامة، حسن (1995). طرق تدريس الرياضيـــات بين النظرية والتطبيق. القاهرة : دار الفجر.

34- سوسن، شاكر مجيد (2005). أساسيات بناء الاختبارات والمقاييس النفسية والتربوية. عمان: دار رسلان للطباعة والنشر والتوزيع.

35- سوسن، شاكر مجيد (2014). أساسيات بناء الاختبارات والمقاييس النفسية والتربوية. عمان: مركز ديبونو لتعليم التفكير.

36- السيد، فؤاد .(1994). الذكـــاء. ط5. القاهرة: دار الفكر العربي.

37- السيد، محمد على (2011). موسوعة المصطلحات التربويـــة. عمان: دار المسيرة للنشر والتوزيع.

38- سيف خليل، إسماعيل (2017). إعداد اختبار فان هيل للتفكير الهندســـي لطلبة المرحلة المتوسطة على وفق نظرية الاستجابة للفقرة. رسالة ماجستير غير منشورة، قسم العلوم التربوية والنفسية. كلية التربية ابن رشد للعلوم الإنسانية، بغداد، العراق.

39- شحاتة، ساميه سمير (2012). دروس في القيـــاس النفسي والتربوي. القاهرة: مكتبة ايتراك للنشر والتوزيع .

40- الصادق، اسماعيل (2001). طرق تدريس الرياضيات. القاهرة: دار الفكر العربي.

41- الطريري، عبد الرحمان بن سليمان بن سعود (1997). القياس النفسي والتربوي: نظريته و أسسه و تطبيقاته. الرياض: مكتبة الرشد للنشر والتوزيع.

42- طلال، سعد الحربي (2003). اتجاهات وأساليب معلمي رياضيات المرحلة المتوسطـــة في تدريس الهندســـة وارتباطها بمستويـــات فـــان هيل. مجلة مركز البحوث التربويـــة. جامعة قطر. السنة (12) العدد(24)، ص29-59.

43- العطاس، أحمد (2014). دلالات الصدق والثبات لاختبار مستويـــات التفكير الهندســـي في ضوء نموذج فان هيل لطلاب الصف الثاني ثانوي في مدينة مكة المكرمة. رسالة ماجستير غير منشورة. كلية التربية. قسم علم النفس جامعة أم القرى.

44- عباس، حسن (2007). أثر ثلاثة استراتجيات في طرح الأسئلة على التفكير في الهندســـة واختزال القلـــق نحوها لدى طلاب الصف التاسع أساسي بغزة. رسالة ماجستير غير منشورة. الجامعة الاسلامية بغزة.

45-عبد الحفيظ، مقدم (2011). الإحصـــاء والقيـــاس النفســـي والتربوي. الجزائر: ديوان المطبوعات الجامعية.

46-عبد القادر، أيمن مصطفى (1997). فهم الأشكال الهندسية وخواصها لدى التلاميذ معلمي الرياضيات وعلاقته بمستويـــات فـــان هيل للتفكير الهندســـي. رسالة ماجستير غير منشورة. كلية التربية. جامعة الاسكندرية.

47- عبد القوي، مصطفى (2007). فاعلية استراتجية التدريس بحل المشكلة في تنمية التفكير الهندســـي والتحصيل لدى تلاميذ الصف الأول الثـــانوي. مجلة دراســـات في المناهج وطرق التدريس. مصر. المجلد(125)، ص162-202.

48- عبد المسيح، عماد (1991). استخدام نموذج راش اللوغاريتمي أحـــادي البارمتر في تحليل مفردات الاختبارات المعرفية مرجعية المعيار ثنائية القطب (دراسة تجريبية). مجلة البحث في التربية وعلم النفس. العدد (04) أفريل. ص443-475.

49- عبيد، وليم (1993). تقرير عن مؤتمر الكونجرس العالمي لتعليم الرياضيات المنعقد بكندا في الفترة من 17-23 أغسطس1992. المجلة التربويـــة. جامعة الكويت. 8(27). ص193-204.

50- عبيد، وليم (2004). تعليم الرياضيات لجميع الأطفال في ضوء متطلبات المعايير وثقافة التفكير. عمان: دار المسيرة.

51- عبيد، وليم و عفانة، عزو (2003). التفكيـــر والمناهج المدرسي. الكويت: مكتبة الفلاح.

52- عدنان، يوسف العتوم و عبد الناصر، ذياب الجراح و موفق، بشارة (2009). تنمية مهارات التفكير نماذج نظرية وتطبيقات عملية. ط2. عمان: دار المسيرة للنشر والتوزيع.

53- عفانة، عزو إسماعيل (2002). تقويـــم مقرر الرياضيـــات المطور للصف السادس الأساسي في فلسطين في ضوء مستويات التفكير الهندســـي لفان هيل.الجمعية المصرية لتربويات الرياضيات. المؤتمر العلمي السنوي الثاني. البحث في تربويات الرياضيات. جامعة عين شمس. القاهرة.

54- علام، صلاح الدين محمود (1995). الاختبـــارات التشخيصية: مرجعية المحـــك في المجالات التربويـــة والنفسيـــة والتدريبية. القاهرة : دار الفكر العربي.

55- علام، صلاح الدين محمود (2000). القيـــاس والتقويـــم التربـــوي والنفســـي أساسياته وتطبيقاته وتوجهاته المعاصرة. القاهرة: دار الفكر العربي.

56- علام، صلاح الدين محمود (2001). الاختبارات التشخيصية: مرجعية المحك في المجالات التربوية والنفسية والتدريبية. ط2. القاهرة: دار الفكر العربي.

57- علام، صلاح الدين محمود (2013). نماذج الاستجابة للمفردة الاختباريـــة أحاديـــة البعد ومتعددة الأبعاد وتطبيقاتها في القياس النفسي والتربوي. القاهرة: دار الفكر العربي.

58- علي بن محمد، عبد الله زكري (2009). الخصائص السيكومترية لاختبار (أوتيس-لينون) للقدرة العقلية مقدرة وفق القياس الكلاسيكي ونموذج راش لدى طلبة المرحلة المتوسطة بمحافظة صبيا التعليمية. أطروحة دكتوراه غير منشورة. قسم علم النفس كلية التربية جامعة أم القرى. السعودية.

59- عوض لله، محمد عبد الرحيم (2000) مقارنة بين أسلوبي أنموذج راش والطريقة التقليدية في بناء اختبارات الذكاء باستخدام محك التنبؤ بالتحصيل الدراسي. أطروحة دكتوراه غير منشورة. جامعة بغداد.

60- فارس، راتب الأشقر (2011). فلسفة التفكير ونظريات في التعلم والتعليم. عمان: دار زهران للنشر والتوزيع.

61- فاضل سلامة، شطناوي (2008). أسس الرياضيـــات والمفاهيم الهندسية الأساسية. عمان. الاردن: دار المسيرة للنشر والتوزيع.

62- فتحي، جروان (2007). تعليم التفكير مفاهيم وتطبيقات. ط3، عمان: دار الفكر.

63- فؤاد، أبو حطب و سيد، عثمان و آمال، صادق (2008). التقويـــم النفســـي. ط**4**، القاهرة: مكتبة الأنجلو مصرية.

64- كاظم ، معصومة وآخرون (1970). أساسيات تدريس الرياضيات الحديثة. ط2. القاهرة: دار المعارف المصرية.

65- كاظم، أمينة محمد (1988أ). دراسة نظرية نقدية حول القيـــاس الموضوعي للسلوك نموذج راش. مؤسسة الكويت للتقدم العلمي.

66- كاظم، أمينة محمد (1988ب). استخدام نموذج راش في بنـــاء اختبـــار تحصيلي في علم النفس وتحقيق التفسير الموضوعي للنتائج. مطبوعات جامعة الكويت.

67- كاظم، أمينة محمد (1994). تدريج ومعايرة المقاييس. في محمد عماد الدين اسماعيل وأمينة كاظم وآخرون. معايير نمو طفل ما قبل المدرسة. القاهرة: المجلس القومي للأمومة والطفولة.

68- الكحلوت، أحمد اسماعيل (1991). الخصائص السيكومترية لمقياس (مايرز - بريجز) لأنماط الشخصية وقدرة المقياس على التنبؤ بالتحصيل الدراسي. أطروحة دكتوراه. كلية الدراسات العليا. الجامعة الاردنية.

69- لويس، ر أيكين (2007). الاختبارات والامتحانات قياس القدرات والأداء. الرياض: مكتبة العبيكان.

70- لينا، فؤاد جواد (2011). مستويات التفكير الهندسي لدى طلبة قسم الرياضيات في كلية التربية الأساسية بالجامعة المستنصرية بغداد. مجلة البحوث التربوية والنفسية. العدد (31). ص429-466.

71- ماجد، محمد الخياط (2011). درجة مطابقة اختبار تحصيلي وفق نموذج راش أحادي المعلمة في الكشف عن مستوى المعرفة العلمية في المهارات الرياضية لدى طلبة الصف الثامن الأساسي. مجلة جامعة الأقصى (سلسلة العلوم الإنسانية). المجلد (10). العدد(1). ص 87 - 111.

72- مجدي عزيز، ابراهيم (2005). التفكير من منظور تربوي: تعريفه- طبيعته- مهارته- تنميته- أنماطه. القاهرة: دارعالم الكتاب للطباعة والنشر والتوزيع.

73- محاسنة، ابراهيم محمد (2013). القياس النفسي في ظل النظرية التقليدية والنظرية الحديثة. عمان. الاردن: دار جرير للنشر والتوزيع.

74- مخلوف، لطفي عمارة (1994). مستويات التفكير الهندسي لدى الطلاب المعلمين طبقا لنموذج "فان هيل". كلية التربية جامعة المنصورة . دراسة تحليلية. مجلة كلية التربية. العدد (26). ص451-480.

75- مدحت نوري، جليل (2015). مستويات التفكير الهندسي لدى طلبة المعهد التقني بعقوبة. مجلة الفتح. العدد (64). ص428- 457.

76- مراد، صلاح أحمد و سليمان، أمين على (2013). الاختبارات والمقاييس في العلوم النفسية والتربوية خطوات إعدادها وخصائصها. ط**2**، الجزائر: دار الكتاب الحديث.

77- مطر، نعيم (2004). أثر استخدام مخططات المفاهيم في تنمية التفكير الرياضي لدى طلاب الصف الثامن الأساسي بغزة. رسالة ماجستير غير منشورة. الجامعة الاسلامية بغزة.

78- ميمي السيد، أحمد اسماعيل (2007). الخصائص السيكومترية لاختبار القدرة العقلية باستخدام نموذج راش لدى طلبة المرحلة الثانوية العامة. رسالة ماجستير غير منشورة. كلية التربية. جامعة الزقازيق.

79- ميمي السيد، أحمد اسماعيل (2014). التوجهات الحديثة في القياس النفسي والتربوي. القاهرة: دار الكتاب الحديث.

80- نظلة، حسن و أحمد، خضر (1985). أصول تدريس الرياضيات. ط3. مصر: مكتبة الآنجلو المصرية.

81- النبهان، موسى (2013). أساسيات القياس في العلوم السلوكية. ط**2**. عمان. الناشر: دار الشروق للنشر والتوزيع.

82- النعيمي، عز الدين (2011). أثر الزيادة في عدد المفردات المرتبطة على الخصائص السيكومترية للمفردة والاختبار. مجلة اتحاد الجامعات العربية للتربية وعلم النفس. 9 (3). 158- 178.

83-الهاشمي، انعام هاشم سلطان (2005). مقارنة بين أسلوبي أنموذج راش ونظرية القياس التقليدي لبناء اختبارالذكاء اللغوي لتلامذة المرحلة الابتائية في قدرته على التنبؤ بتحصيلهم في مادة اللغة العربية. أطروحة دكتوراه غير منشورة. كلية التربية. ابن رشد. جامعة بغداد.

84- وليد، مسعود (2014). محاضرات في القياس نظرية الاستجابة للمفردة مع تطبيقات عملية. كلية التربية قسم علم النفس. جامعة أم القرى.

85- وائل محمد، أحمد مبارك (2010). الأداء التفاضلي لاختبار الرياضيات في الدراسة الدولية بيزا 2006. أطروحة دكتوراه غير منشورة. جامعة اليرموك. إربد. الاردن.

86- Ayala, R,j .(2009). The Theory and Practice of item Response Theory. A division of Guilford Publications , Ine 72 Spring, Street ,New York, Ny 10012. www.guilfford.com

87- Anne, Teppo. (1991). Van Hiele levels of geometric thought Revisited . Natienal Council of Teachers of Mathimatics .http : www.jstor.org/stable/27967094.

88- Abdalaziz, N.(2010). A Gender Related Differential Item Fonctioning of Mathematics Test Items. The international journal of Educational and Psychological Assessment ,Vol-5,pp101-116.

89- Andrich, David .(1982). An Index of Person Separation in Latent Trait Theory, the Traditional KR-20 Index, and the Guttman Scale Response Pattern. Education Research and Perspectives, 9:1, 95-104.

90- Baker, F.B.(2001).The Basics of Item Response Theory. ERIC clearinghouse on Assessment and Evaluation, University of Maryland,College Park,MD.

91- Brennan, R. L. (1972). A generalized upper-lower item discrimination index. Educational and Psychological Measurement, 32(2), 289-303.

92- Brown, J.D. & Hudson, T.(2002). Criterion-Referenced Language Testing. Cambridge University Press. https://www.amazon.com/Criterion-Referenced-Language-Testing-Cambridge - Linguistics/dp/0521000831

93- BURTON ,R et Detheux-jehin, M.(1999). Les éléves Du Premier Degré secondaire .Sont-Ils .Préts à. Démentrer en Géométrrie ? S-ynthése de la recherche en pédagogie n° 02/97. Service de Pédagogie expérimentale. Université de liége Sart- Tilman-Bàt.B32.4000 Liége. pp02-20.

94- C H Lawshe .(1975). A quantitive approach to content validity. Personnel Psychology ,28 (4) :563-575.

95- Cauffman, E., & MacIntosh, R. (2006). A Rasch Differential Item Functioning Analysis of the Massachusetts Youth Screening Instrument: Identifying Race and Gender Differential Item Functioning among Juvenile Offenders. Educational and Psychological Measurement, 66,502-521. http://dx.doi.org/10.1177/0013164405282460

96- Callingham, R., & Bond, T. (2006). Research in Mathematics Education and Rasch Measurement. Mathematics Education Research Journal, 18, 1-10. http://dx.doi.org/10.1007/BF03217432

97- Chappell, Michaele F. (2001).Spot Light in the Standards Geometry in the Middle Grades : From its Past to the Present.Mathematics Theaching in the Middle School ,no,6,pp.516-519.

98- Charles, S ; Brian, D.(2012). Handbook on measurement, assessment, and evaluation in higher education. First published by Routledge 711 Third Avenue, New York , NY 10017.

99- Chong chin and Nor Mohd,T.(2015).Assessing First year Pre- service Teach ers' Geometric Reasoning Ability on Two-Dimensional Shapes. Educatum . journal of science, Mathematics and Technology.vol.2.no.1.p1 -12.

100- Chrisitine, D.(2010). Item Response Theory Understanding Statistics Measurement .Oxford University Press, Inc.

101- Crislip, M.A. & Chin-Chance, S. (2001). Using Traditional Psychometric Methodologies and the Rasch Model in Designing a Test. Paper presented at the Annual Meeting of the American Educational Research Association, Seattle, WA. (ERIC Document Reproduction Service No. ED (453-231) .https://files.eric.ed.gov/fulltext/ED453231.pdf

102- Crocker ,L and Algina.(1986). introdication To Classical and Modern Test theory. Harcount Brace jovanovich College Pubishers,U.S.A.

103- Cronbach , L .J . (1970). Essentials of Psychological testing. New york : Harper and Row publishers .

104- Dimitrov, Dimiter M. (2012). Statistical methods for validation of assessment scale data in counseling and related felds. American Counseling Association .5999 Stevenson Avenue Alexandria, VA 22304.www.counseling.org

105- EL – Korashy , AF . (1995) : Applying the Rasch model to the selection of items for a mental ability test. Educational and psychological Measurement ,Vol .55 , No . 5 , pp 753 -763 .
106- Embretson, S.E. and Reise, S.P. (2000) .Item Response Theory for Psychologists. Lawrence Erlbaum Associates, Inc., Mahwah.
107- Francis ,Tuerlinckx ; Paul, DeBoeck.(2001). Non – modeled item interactions lead to distrorted discrimination parameters : A case Study . Methods of Psychological Research Online 2001,Vol 6,No 2 ,p159-174. http://www.mpr-online.de
108- Frank. K , Lester, jr.(2007). Second Handbook of Research on Mathematics Teaching and learning .National council of Teachers of mathematics (NCTM). information Age Publis hing inc.
109- Fuys, D ; Geddes, D ;Tischler, R.(1988). The van hiele Model of Thinking in Geometry Among Adolescent.The National Council of Teachers of Mathematics, INC.journal for Research in Mathematics Education, Monograph Numer 03 .pp.01-195.
110- Gamer, M. and Englhard, G.(1999). Gender differences in performance on multiple choice and Constructed response mathematics items . Applied Measurement in Education, 12(1), 29-43.
111- Geisinger, K. F. (1991). Using standard-setting data to establish cutoff scores. Educational Measurement: Issues and Practice, 10(2), 17-22.
112- Gronlund N.E. (1981). Measurement and evaluation in teaching. New York: Macmillan Publishing Co. Inc:
113- Hambleton ,R.K .(1989). Principles and applications of item response Theory in R.L.Linn (Ed). Educational measurement (3rd edition ;pp147-200).New york.
114- Hambleton, R. K. (1980). Test score validity and standard-setting methods. Criterion-referenced measurement: The state of the art, 80, 123.
115- Hambleton, R. K. (1978). On the use of cut-off scores with criterion-referenced tests in instructional settings. Journal of Educational Measurement, 15(4), 277-290.
116- Harris, A. and Carlton, S.(1993). Pattens of gender differences on mathematics itmes on the scholastic aptitude . Applied Measurement in Education, 6(2), 137-150 .
117- Hatfield, M ; Edwards, N ; Bitter, G ; Mornow, j. (2001). Mathematics Methods for Elementary and Middle School Teachers. (4) Edition New yord : johen and Sons, inc. .
118- Innabi, H and Dodeen, H .(2006). Content Analysis of Gender Related Differential Item Fonctioning (TIMSS) Items in Mathematics in jorden School Science and Mathematics,106 ,8 ; pp328-337.
119- Juhaida, A et al .(2016).Differential Item functioning in Online Pearning inst- rument (EPFun). Creative Education, 7, pp180-188.http://dx.doi.org/10.4236/ce.2016.71018
120- Linacre, J. M .(2012). A User's Guide to Winsteps Ministep Rasch- Model Computer Programs. Winsteps.com
121- Linacre, J. M .(2008a). Winsteps (version 3.92.1) . [computer software] Retrieved from http://www.Winsteps.com
122- Linacre, J. M .(2006b). Auser's guide to WINSTPES MINISTEP Rasch model computer Programme manual 3.92.1 .Retrieved from http://www.Winsteps.com.

123- Katherine, R and Shuwan, C.(2001). An Examination of Item Context effects, DIF, and gender DIF ,Applied Measurement in Education, 14(1),pp73-90.

124- Kelley, T. L. (1939). The selection of upper and lower groups for the validation of test items. Journal of Educational Psychology, 30, 17-24.

125- Kline, Rex B .(2011). Principles and practice of structural equation modeling. 3rd ed. A Division of Guilford Publications, Inc. 72 Spring Street, New York, NY 10012.

126- Louis L. Thurstone. (1925). A Method of Scaling Psychological and Educational Tests. Journal of Educational Psychology. 16: 433- 451.

127- Masters G.N. (1988). Item discrimination: when more is worse. Journal of Educational Measurement 25:1, 15-29.

128- Matthew, G ; Ray, C.(2013). Determining the Number of Factors to Retain in EFA : Using the SPSS R-Menu v2.0 to Make More judicious Estimations. Practical Assessment Research and Evaluation,Vol 18,No 8,p1-14

129- Mayberry, j.w .(1981). An investigation of the Van Hiele levels of geometric thought in undergraduate presevice teachers. Doctoral thesis,University of Georgia.

130- McBride, B ; Carifio, J.(1995). Empricial Results of Using an Analytic versus Holistic Scoring Method to Scare Geometric Proof : Linking and Assessing Greeno,Bloom,and van hiele views of Student Abilities to do Proof .paper presented at the Annual Meeting of the American Educational Research Association (San Fracisco),p1-57

131- Michel, klaren .(2012). Zwakke centrale coherentie en de van hiele hivea us, Master of Science and Education, Eindhoven University.
http://alexandria.tue.nl/extra2/afstversl/esoe/748067.pdf

132- Mislevy, R. T. (1990). Modeling Item Response When Different Subjects Employ Different Solations Strategies, Netherlands: Psychometrica, Vol. 55, No., PP. 195-215.

133- Pardikaris, S.(1994). " Markov chains and van hiele levels :Amethod of distinguishing different types of student geometric reas processes " ,inc .j.Math Educ.Sci.Technol,vol,25,no.4.

134- Pegg, j ; Davey, G.(1998).interpreting student understanding in geometry : A synthesis of two models.in D .Chazan (Ed), Designing environments for developing understanding of geometry and space (pp.109-135), Mahwah, NJ : Lawrence Erlbaum Associates.
http://www.cimm.ucr.ac.cr/ciaem/articulos/universitario/materiales/The%20Van%20Hiele%20Theory.

135- Prieto, L et al .(2003). Classical test theory versus Rasch analysis for quality of life question nnaire reduction. Health and Quality of life Outcomes ,1-27,p1-13.
http://www.hqlo.com/content/1/1/27

136- Randall, s. (1998). Company measurement theories, paper present at the annual meeting of the American educational research association san Diego, CA, April, 13-17.

137- Raykov, T. (1997). Estimation of composite reliability for congeneric measures. Applied Psychological Measurement, 21(2), 173-184.

138- Satterthwaite, F.E. (1946). An Approximate Distribution of Estimates of Variance Components. Biometrics Bulletin, 2: 110-114

139- Serow, Penelope.(2007). Utilising the Rasch Model to Goin insight into Students Unerstandings of class inclusion Concepts in Geometry . Mathematics Essential Research Essential Practice. Volum. 2. 651-660.

140- Sheldem, Zebeck .(2014). APA Dictionary of Statistics and Research Methods. American Psychological Association Washington, DG.

141- Steven P. Reise, Niels G. Walle .(2003). How Many IRT Parameters Does It Take to Model Psychopathology Items?, Psychological Methods, 2003, 8, 2, 164-184.

142- Stols ,G ; Caroline ,L ; Tim ,D.(2015). An application of the Rasch measurement thery to anassessment of geometry thinking levels. African journal of Research in Mathemathics Science and Technology Education, V ,19.issue1.p69-81.

143- Usiskin, Z. (1982). Van Hiele levels and achievement in secondary school geometry (Final report of the Cognitive Development and Achievement in Secondary School Geometry Project). Chicago: University of Chicago, Department of Education. (ERIC Document Reproduction Service No. ED 220 288) .ucsmp.uchicago.edu/resources/van_hiele_levels.pdf

144- Van de Walle, J.A. (2015). Elementary and Middle school mathematics: Teaching developmentally. Ninth edition. New York: Longman.

145- Van Hiele, Pierre M.(1999). Developing Geometric Thinking Through Activities That Begin with Play. Theaching Children Mathematics .pp 16-310. Natienal Council of Teachers of Mathimatics,inc.www.nctm.org.

146- Van hiele, Pierr M.(1986). Structure and Insight a theory of Mathematics Education, New York,Academic Press.

147- Welch, B. L. (1947). "The generalization of "Student's" problem when several different population variances are involved.", Biometrika, 34: 28-35
http://www.winsteps.com/winman/t-statistics.htm

148- William j ; Boone et al .(2014). Rasch Analysis in the Human Sciences. Spinger Dordrecht Heidelberg New York London . http:// extras.springer.com

149- Wilson, Mark .(2005). Constructing measures: an item response modeling approach. Lawrence Erlbaum Associates, Inc., Publishers 10 Industrial Avenue Mahwah, NewJersey, 07430, www.erlbaum.com

150- Wilson, M.(1990). Meassuring a van hiele Geometry sequence : A Reanalysis. journal for Research in Mathemathics Eduction, 21(3),230-237.

151- Wright, B. & Stone, M. (1979). Best test design. MESA Press: Chicago, IL

152- Wu, M. & Adams, R. (2007). Applying the Rasch model to psycho-social measurement: A Practiceal approach. Educational Measurement Solutions, Melbourne. pp01-87.

153- Yen, W. M. (1984). Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. Applied Psychological Measurement, 8, 125-145.

154- Yoke Mooi one.(2010). Understanding differential functioning by Gender in Mathematics Assessment ,thesis Phd .The University of Manchester School of Education.

155- Zwick, R, Thayer,D,T, Lewis, C.(1999). An Empirical Bayes Approach to Mantel-Haenszel DIF Analysis .journal of Educational Measurement,36,1,1-28.

الملاحق

الملحق رقم (01)

استمارة تحكيم اختبار مستويات التفكير الهندسي لفان هيل – في صورته الأولية–

معلومات شخصية عن الأستاذ المحكم

الإسم واللقب : ..... **الوظيفة :** ..... **الدرجة العلمية :** .....

أخي المحكم/ أختي المحكمة:

السلام عليكم ورحمة الله وبركاته

في اطار إعداد الباحث لأطروحة دكتوراه علوم في علم النفس تخصص القيـــاس والتقويم بجامعة وهران2، التي موضوعها (استخدام نموذج راش اللوغاريتمي أحادي البارامتر ونظرية القياس الكلاسيكية في تحليل وتدريج بنود اختبار مستويات التفكير الهندسي مبني وفق نموذج فان هيل)، وللتأكد من صلاحية المفردات المقترحة لبناء اختبار لمستويات التفكير الهندســـي وفق نموذج (فان هيل)، والموجه إلى تلامذة المرحلة الثانوية الشعب العلمية (تقني رياضي، ورياضيات، وعلوم تجريبية) ، أرجوا من حضرتكم التكرم بإبداء الرأي حولها وذلك وفق الخطوات التالية:

أولاً: وضع إشارة (X) أمام كل مفردة: (1) تقبل المفردة، أو (2) تعدل المفردة، أو (3) تحذف المفردة تماماً.

ثانياً: اقتراح تعديلات حول المفردات أو إضافة مفردة تراها ضرورية، على ظهر الورقة مباشرة.

ثالثاً: التأكد من مفتاح الإجابات الصحيحة المرفق مع الإستمارة.

التعريف بنظرية مستويات التفكير الهندســـي لفان هيل:

قدم "بيير فان هيل" وزوجته "دينا فان هيل" سنة 1957 ما يسمى بنظرية (فان هيل) والتي استندا فيها إلى دراستين لهما عن الصعوبات التي يواجهها التلامذة في دراسة الهندسة (بهولندا)، حيث اشاروا إلى أن التفكير الهندســـي وتعلم الهندسة يسير في مستويات متتابعة تتضمن نمواً في طرق ونوعية التفكير وأن لكل مستوى لغته ومصطلحاته التي يمكن استخدامها، وأن تعلم مستوى معين يتطلب تعلما للمستوى السابق له وأن الإنتقال من مستوى لآخر يتطلب وقتا لنضوجه قبل الإنتقال إلى المستوى التالي، وحذرت

النظرية من أنه إذا كان المتعلم مازال في مستوى معين وكان التدريس في مستوى أعلى، فإنه لا يحدث تعلم ولكن من الممكن فقط أن المتعلم يستظهر المعلومات وقد يتذكرها ولكن دون فهم، ذلك أنه من وجهة نظر فان هيل: " تنمو مستويات التفكير من خلال التعليم والتعلم وليس من خلال العمليات البيولوجية".

وقد صنف في البداية (فان هيل) مستويات التفكير الهندسي إلى خمسة مستويات كالآتي:

❖ **مستوى(1) التعرف البصري الكلي (Visual –Holistic Reasoning Level):**

في هذا المستوى يتعلم الطفل الأسماء ويميز بين الأشكال ككيان متكامل دون إدراك لخواص الشكل فهو يدركها بصريا ولكن لا يدرك خواصها يمكن للطفل في هذا المستوى أن يقوم بالآتي:

- يتعرف على هيئة الشكل وهو في أوضاع مختلفة.
- يرسم بعض الاشكال الهندسية البسيطة.
- يسمي أشكالا بأسماء عامة (مثل المستطيل على شكل الباب).
- يميز بين الأشكال بحسب مظهرها ويصفها بالكلام.
- يتعرف على أجزاء الشكل.
- ينظر لكل شكل على حدة بدون تعميم.
- يميز بين شكل أضلاعه مستقيمة (مربع مثلا) وشكل محيطه على شكل منحنيات ولا يميز بين الأشكال من نفس النوع.

❖ **مستوى (2) التفكير التحليلي (Analytic - componential reasoning):**

في هذا المستوى يميز الطفل خواص الأشكال ولكن دون إدراك علاقات بين هذه الخواص وهو أيضا لا يمكنه فهم أو استعياب التعاريف التي تعطى للأشكال، الشكل هنا بالنسبة له مجموعة من الخواص وليس مجرد هيئة أو صورة، ويمكن للطفل في هذا المستوى أن يقوم بالآتي:

- يميز بين الأشكال بحسب خواصها ومكوناتها.
- يستخدم ألفاظا لفظية وكلامية.
- يتعرف على الشكل من خواصه ويختبرها بالقياس.
- يستخدم الخواص في رسم الشكل.
- يعمم خواصاً على مجموعة من الأشكال (المربعات لها أربع أضلاع لها أربع زوايا قائمة).

- يحل بعض التمارين على خواص مثل مجموع قياسات زوايا المثلث.
- لا يرى حاجة لإثبات صحة الخواص التي يدركها فيكفي القياس مثلا.

❖ **مستوى (3) الإستدلال بطرق غير شكلية (formal deduction Proof):**

في هذا المستوى يصنف الأشكال عن طريق خصائصها، ويدرك تعاريف مجردة ويستخدم ألفاظا لها طابع منطقي مثل "بعض" "كل" ويمكنه أن يستدل على خاصية ما بدون حاجة لبرهان منطقي (مجموع الزوايا للشكل الرباعي 360 درجة – يكفي الإستدلال على ذلك أنه مكون من مثلثين وكل مثلث مجموع زواياه 180 درجة). ويمكن للطفل في هذا المستوى أن يقوم بالآتي:

- يرتب أشكالا هندسية في ضوء خواصها ولكن دون الإستناد إلى برهان منطقي.
- يدرك الخصائص التي تكفي لتمييز شكل عن آخر.
- يستنتج بعض خواص العلاقات مثل إذا كان أ=ب فإن ب=أ وأنه إذا كان أ=ب، ب=ج فإن أ=ج.
- يصل إلى نتائج من معطيات ويدلل على صحتها بطرق غير شكلية.
- يتابع برهانا منطقيا ولكنه لا يقيمه بنفسه.
- يدرك الفرق بين نظرية هندسية ومعكوسها ويشرحها بطرق غير شكلية.
- لا يدرك استقلالية مجموعة من المسلمات أو النظريات.

❖ **مستوى (4) الإستنباط الشكلي(formal deduction Proof):**

ويطلق عليه أيضاً الاستدلال المجرد، في هذا المستوى يستطيع المتعلم أن يفكر نظرياً ويقيم براهين منطقية، ويدرك العلاقات بين الخواص كما يدرك أهمية الإستنتاج ذهنياً واستخلاص نتائج من خواص ومعطيات معطاة ويمكن للمتعلم في هذا المستوى أن يقوم بالأتي:

- يميز بين المصطلحات المعرفة والغير المعرفة؛ وبين العبارة التي تُقبل كمسلمة وتلك اللازم برهنتها (النظرية).
- يُنتج تتابعا من العبارة التي يستنتج فيها كل عبارة من السابقة لها وحتى يصل إلى نتيجة مطلوبة أو تساعد في الوصول إلى المطلوب إثباته بالبرهان

- يدرك معنى الشرط اللازم والشرط الكافي (مثلا :إذا كان "أب ج د" متوازي أضلاع فإن أب//ج د. كون "أ ب ج د" متوازي أضلاع فهذا الشرط كاف لأن يكون أ ب//ج د. ولكن إذا كان أ ب//ج د. فهذا شرط لازم ولكنه ليس كاف لأن يكون "أ ب ج د" متوازي أضلاع).
- يقيم برهانا يستند إلى مجموعة من المسلمات أو النظريات.
- يدرك خواص عامة تجمع بين مجموعة من الأشكال أو مجموعة من النظريات.

❖ **مستوى (5) الدقة البالغة (Rigor Dedction):**

ويسمى أيضاً الاستدلال المجرد الكامل، في هذا المستوى يمكن للمتعلم المقارنة بين أنظمة هندسية مختلفة (هندسة اقليدية، هندسة غير اقليدية، هندسة محايدة لا تعتمد على مسلمة التوازي الإقليدية ولا على مسلمات التوازي اللااقليدية) ويكون المتعلم على وعي وفهم لدور المنطق والطرق المختلفة للبرهان وأسانيده في المنطق الشكلي مثل البرهان المباشر والغير مباشر وذلك الذي يعتمد على رفض التعارض .ويمكن للمتعلم في هذه المستوى أن يقوم بالآتي:

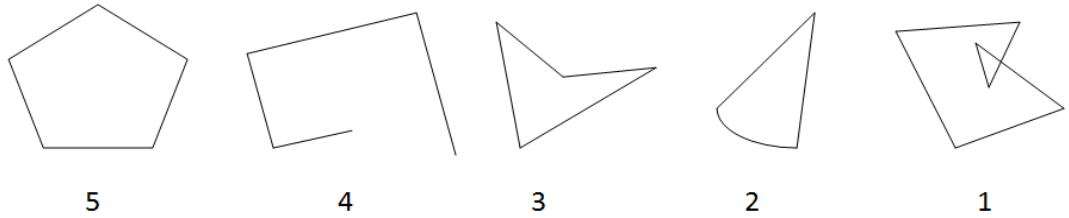
- استنتاج واثبات بعض النظريات في مختلف أنظمة المسلمات الهندسية (الأقليدية، واللاإقليدية).
- يبرهن على بعض النظريات بعدة طرق (بحسب طبيعة النظرية) مثل البرهان المباشر من المعطيات إلى المطلوب (بما أن...إذن).
- البرهان بإستنفاذ جميع الحالات (في المواقف محدودة الإمكانات).
- رفض النقيض (إذا لم يكن المطلوب صحيحا فإنه يؤدي إلى تناقض أو بالخلف).
- يدرك أهمية استقلال المسلمات التي يبنى عليها نظام هندسي معين.
- يمكن إدراك أي تناقض أو عدم إتساق بين مجموعة من العبارات أو الخصائص.
- يمكنه التعامل مع أشكال ثلاثية الأبعاد بمعالجات نظرية.

ملاحظة: اقتصر الباحث في البحث الحالي، في عملية تصميم و بناء اختبار مستويات التفكير الهندســـي، على المستويات الأربعة الأولى فقط لنموذج (فان هيل).

وأخيراً نشكركم على حسن تعاونكم معنا، وأن تتقبلوا منا فائق التقدير والإحترام.

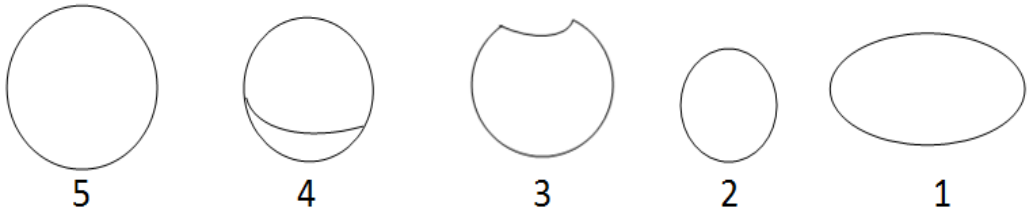
المفردات المقترحة للمستوى الأول: المستوى البصري أو التعرف (الادراكي)

المفردة 01: أي شكل من الأشكال يمثل مضلع؟



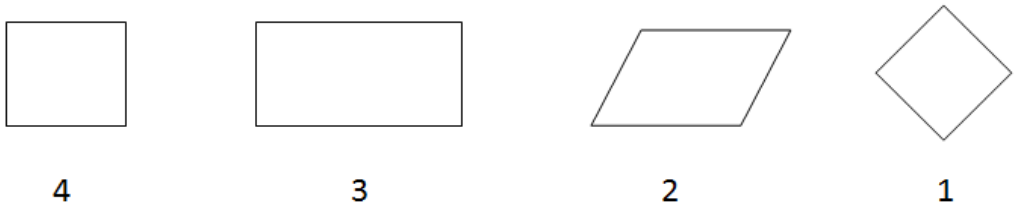
A). الشكل 1 و3 و5 فقط B).الشكل 3 و5 فقط C). الشكل 2 و4 فقط D). كل ما سبق مضلعات

المفردة 02: أي شكل من الأشكال التالية يمثل دائرة؟



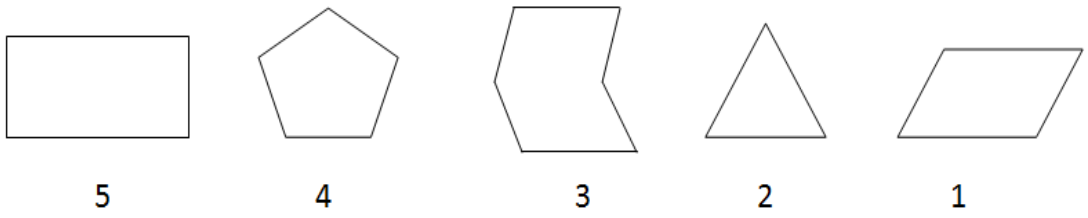
A). الشكل 2 و4 و5 فقط B). الشكل 1 و2 و5 فقط C). الشكل 2 و5 فقط D). كل ماسبق

المفردة 03: أي شكل من الأشكال التالية يمثل المربع؟



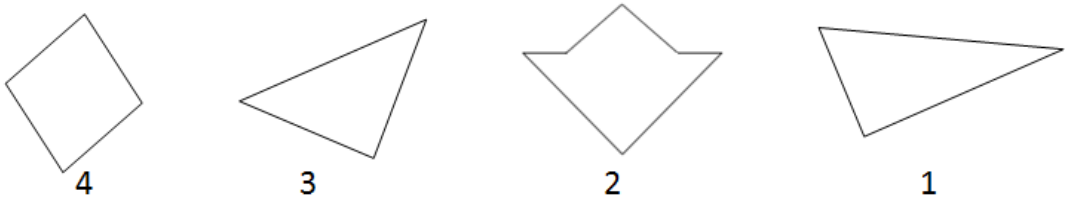
A). الشكل 4 فقط B).الشكل 1 و4 فقط C). الشكل 4 و3 فقط D). كل ماسبق مربعات

المفردة 04: أي شكل من الأشكال التالية يمثل المضلع المنتظم؟



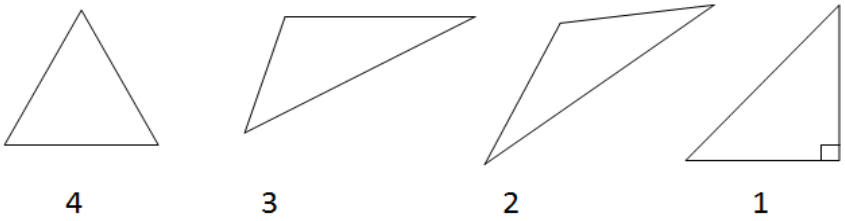
A). 2 و3 فقط B). 1 و3 و4 فقط C). 2 و4 و5 فقط D). كل ما سبق مضلع منتظم

المفردة 05: أي شكل من الأشكال التالية يمثل المثلث؟



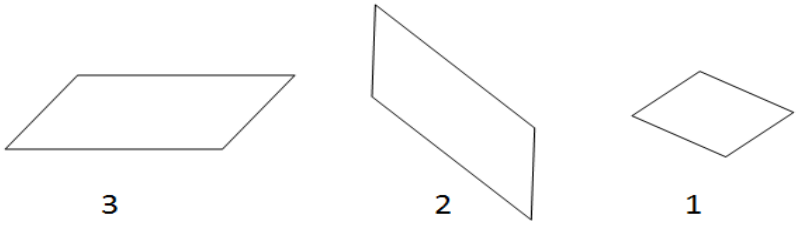
A). 1 و2 و3 فقط B). 2 و4 فقط C). 1 و3 فقط D). جميعها مثلثات

المفردة 06: أي من المثلثات التالية مثلث متساوي أضلاع؟



A). المثلث 4 فقط
B). المثلث 3 و2 فقط
C). المثلث 1 و4 فقط
D). جميعها

المفردة 07: أي شكل من الأشكال التالية يمثل متوازي أضلاع؟



A). الشكل 3 فقط
B). الشكل 1 فقط
C). الشكل 3 و2 فقط
D). كل ماسبق متوازي أضلاع

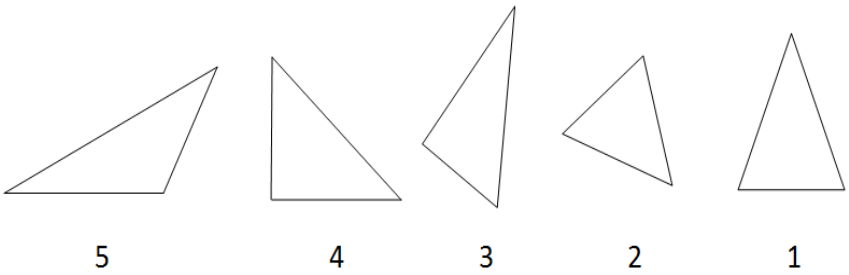
المفردة 08: أي من المثلثات التالية يمثل مثلث متساوي الساقين؟

A). المثلث 1 و2 و3 فقط

B).المثلث 2 و3 و5 فقط

C). المثلث 1 و2 و4 فقط

D). جميعها



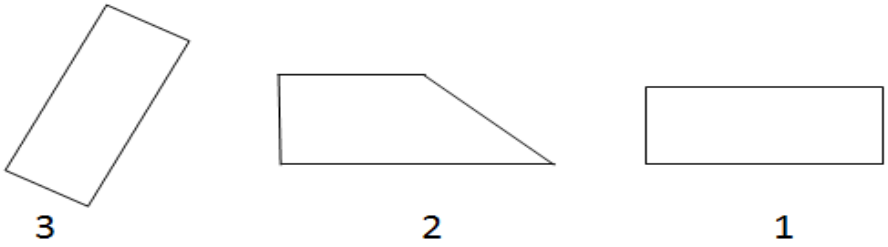
المفردة 09: أي شكل من الأشكال التالية يمثل مستطيل؟

A). الشكل 1 فقط

B). الشكل 2 فقط

C). الشكل 1 و3 فقط

D). الشكل 1 و 2 فقط



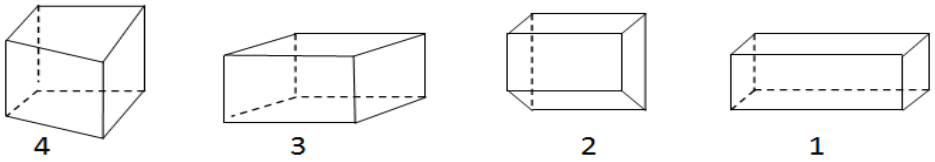
المفردة 10: أي شكل من الأشكال التالية يمثل مكعب؟

A). الشكل 2 فقط

B). الشكل 3 فقط

C). الشكل 3 و2 فقط

D). جميعها



المفردات المقترحة للمستوى الثاني: المستوى الوصفي – التحليلي

المفردة 11: المعين متوازي أظلاع أضلاعه الأربعة متقايسة. أي ممايلي لا ينطبق على المعين؟

A). قطراه متساويان في القياس

B). كل من قطريه ينصف زاويتين متقابلتين فيه

C). قطراه متعامدان

D). الزاويتان المتقابلتان فيه متساويتان في القياس

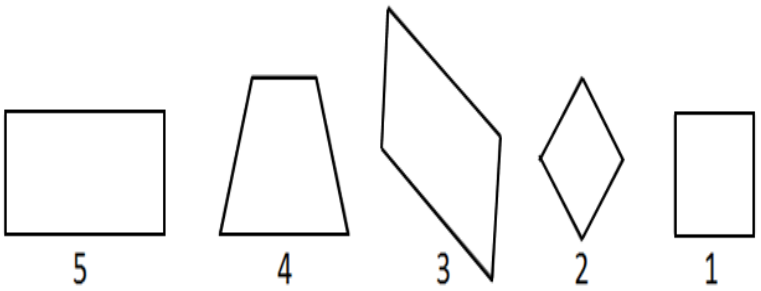
المفردة 12: أي من العبارات التالية **صحيحة** لكل متوازي أضلاع؟

A). القطران متقايسان
B). الزوايا المتقابلة فيه متقايسة
C). القطران متعامدان
D). زواياه متقايسة

المفردة 13: أي مما يلي يمثل أصدق تعريف للمستطيل؟

A). شكل رباعي زواياه متطابقة وقيس كل منها $90°$
B). شكل رباعي فيه كل ضلعين متقابلين متوازيان ومتساويان
C). شكل رباعي فيه كل ضلعين متقابلين متساويان في القياس واحدى زواياه قائمة
D). شكل رباعي فيه كل ضلعين متقابلين متوازيان

المفردة 14: أي من الأشكال التالية تصنف متوازي أضلاع؟



A). الشكل 2 و3 و4 و5 فقط
B). الشكل 1 و2 و3 و4 فقط
C). الشكل 1 و2 و3 و5 فقط
D). جميعها متوازيات أضلاع

المفردة 15: أي من العبارات التالية **صحيحة** لكل مثلث متساوي الأضلاع؟

A). قياس كل زاوية من زوايا المثلث المتساوي الأضلاع يساوي $60°$
B). المثلث الذي أضلاعه الثلاثة متقايسة
C). محاور تماثل المثلث المتساوي الأضلاع هي واسطاته
D). جميع ما ورد أعلاه صحيح في كل مثلث متساوي الأضلاع

المفردة 16: أي من العبارات التالية **صحيحة** لكل شبه منحرف متساوي الساقين؟

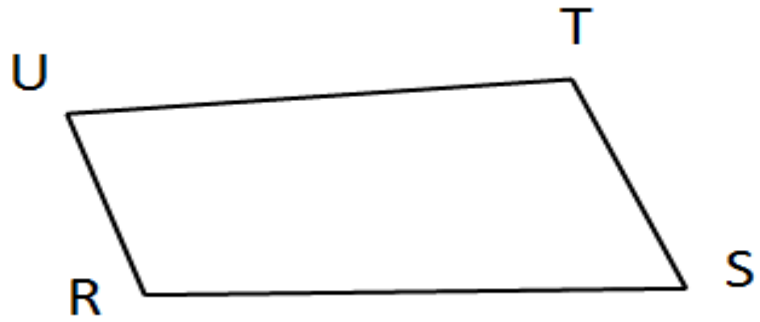
A). قطراه متساويان

B). زاويتا القاعدتين متطابقتان

C). فيه تماثل انعكاسي وخط تماثله يمر في منتصفي قاعدتيه

D). جميع ما ورد أعلاه صحيح في كل شبه منحرف متساوي الساقين

المفردة 17: يعطى الرباعي RSTU. إذا كان (UT) لا يوازي (RS) و (UR) يوازي (ST).
إذن الرباعي RSTU **يجب أن يكون**؟



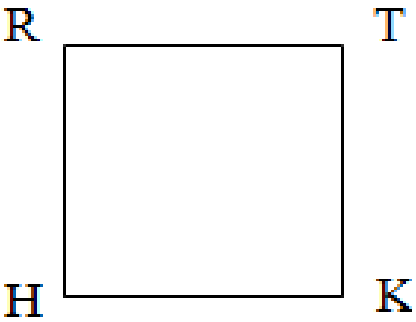
A). شبه منحرف

B). مستطيل

C). معين

D). متوازي الأضلاع

المفردة 18: RTKH مربع، أي من العلاقات التالية **صحيحة** في كل مربع؟



A). RK و KH متساويان

B). TH و RK متعامدان

C). RH و TK متعامدان

D). RH و TH متساويان

المفردة 19: قيس الزاوية المركزية **يساوي**:

A). ضعفي قيس الزاوية المحيطية المشتركة معها بنفس القوس

B). نصف قيس الزاوية المحيطية المشتركة معها بنفس القوس

C). ضعفي قيس أي زاوية محيطية

D). نصف قيس أي زاوية محيطية

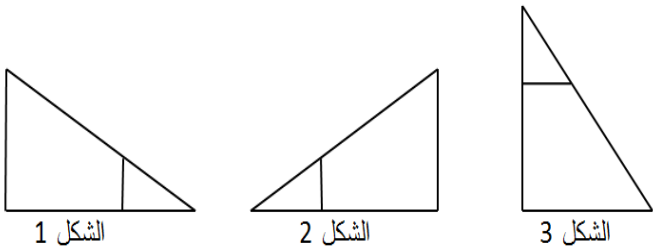
المفردة 20: الشكل (3) هو صورة الشكل (1) بإستعمال **تحويلين نقطيين** على الترتيب (الشكل1؛ ثم الشكل2؛ ثم الشكل3). أي الخيارات التالية **صحيح**؟

A). تناظر ثم إنسحاب

B). تناظر ثم دوران $90°$ في إتجاه عقارب الساعة

C). دوران $180°$ ثم إنسحاب

D). دوران $90°$ عكس إتجاه عقارب الساعة ثم تناظر



المفردات المقترحة للمستوى الثالث: شبه الاستدلالي

المفردة 21: ما هي الخاصية التي تنطبق على جميع المستطيلات ولا تنطبق على بعض متوازيات الأضلاع؟

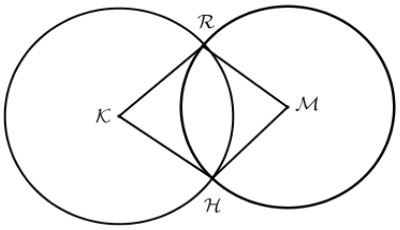
A). الضلعان المتقابلان فيه متساويان في القياس

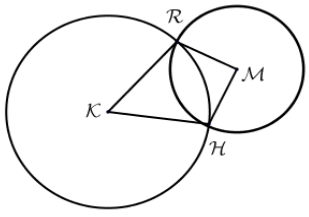
B). الزاويتان المتقابلتان فيه متساويتان في القياس

C). القطران متساويان في القياس

D). الضلعان المتقابلان فيه متوازيان

المفردة 22: النقطتان K ، M مركزا دائرتين تتقاطعان عند R ، H وينتج عنهما شكل رباعي MHKR.





- أي الخيارات التالية **ليس صحيحاً دائما؟**

A). هناك ضلعان في الشكل MHKR متقايسان

B). في الشكل MHKR ، هناك زاويتان على الأقل متقايستان

C). المستقيمان MK و HR متعامدان

D). قيس الزاوية M يساوي قيس الزاوية K

المفردة 23: كل الأشكال في مما يأتي رباعي دائري، **ما عدا؟**

A). المربع

B). المستطيل

C). متوازي الأضلاع

D). شبه المنحرف متساوي الساقين

المفردة 24: فيما يلي جملتان:

الجملة 1: الشكل R هو مستطيل

الجملة 2: الشكل R هو مثلث

– أي من من الخيارات التالية **صحيح؟**

A). إذا كانت الجملة 1 صحيحة، فإن الجملة 2 صحيحة

B). إذا كانت الجملة 1 خاطئة، فإن الجملة 2 صحيحة

C). لا يمكن أن تكون الجملتان 1 و 2 صحيحتين معاً

D). لا يمكن أن تكون الجملتان 1 و 2 خاطئتين معاً

المفردة 25: أي من العبارات التالية **صحيحة؟**

A). كل متوازي أضلاع معين

B). قطرا المعين متعامدان ومتساويان دائما

C). كل معين متوازي أضلاع

D). أطوال أقطار المعين متساوية

المفردة 26: أي من العبارات التالية صحيحة؟

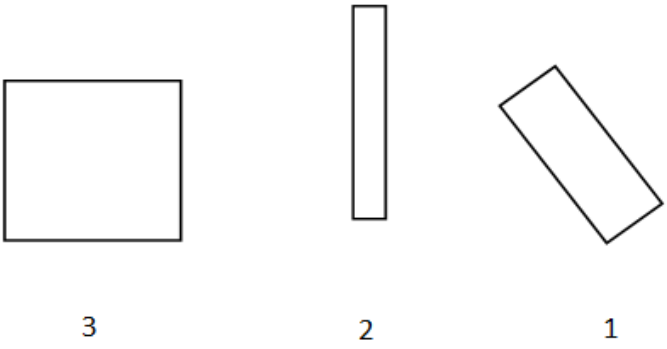
A). جميع خواص المستطيل هي خواص للمربع

B). جميع خواص المستطيل هي خواص لمتوازي الأضلاع

C). جميع خواص المربع هي خواص للمستطيل

D). جميع خواص المربع هي خواص لمتوازي الأضلاع

المفردة 27: أي من الأشكال التالية يمكن اعتباره مستطيلاً؟



A). جميعها

B). 1 فقط

C). 3 و 2 فقط

D). 2 و 1 فقط

المفردة 28: أي من العبارات التالية **صحيحة**؟

A). كل معين قطراه متساويان مربع

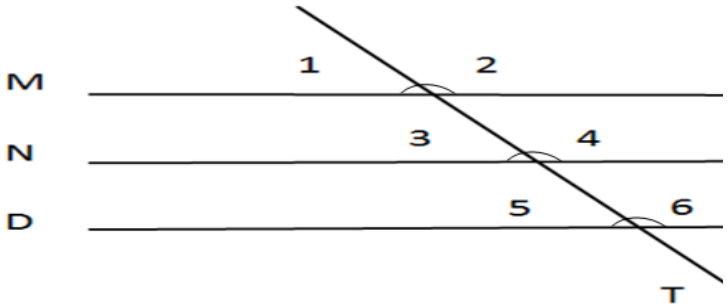
B). كل متوازي أضلاع زواياه قوائم مربع

C). كل متوازي أضلاع قطراه متساويان يكون مربعاً

D). كل مستطيل مربع

المفردة 29: رسم أحمد ثلاثة مستقيمات متوازية (M)، (N)، (D)، ورسم مستقيما (T) قاطعا لهما، تقاطع المستقيم (T) مع المستقيمات (M)، (N)، (D) شكل زوايا وهي مرقمة في الرسم البياني أدناه.

– أي من **التخمين** الذي يتشكل عند أحمد حول شكل الزوايا الناتجة عن هذا التقاطع **صحيح**؟



A). الزوايا 1 و2 و3 متقايسة

B). الزوايا 2 و 4 متكاملتان

C). الزوايا 1 و 3 و 5 متقايسة

D). الزوايا 1 و5 متكاملتان

المفردة 30: أي من العبارات التالية **صحيحة**؟

– الزاويتان المتقابلتان في:

A). الرباعي الدائري متتامتان

B). المعين متكاملتان

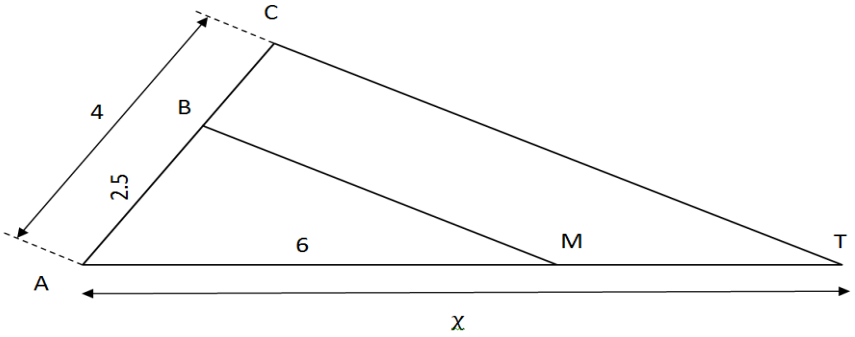
C). المستطيل متتامتان

D). متوازي الأضلاع متقايستان

المفردات المقترحة للمستوى الرابع: الاستدلالي

المفردة 31: في الشكل المقابل القطعة [BM] // [CT]، و [AB] ، [AC] و [AM] ثلاثة قطع أطوالها K ، H ، R على الترتيب بحيث K=2.5cm ، H=4cm ، R=6cm.

– استنتج الطول χ للقطعة [AT] ؟



A). 10.6cm.

B). 9cm .

C). 9.6cm .

D). 10 cm .

المفردة 32: اليك الجملتين التاليتين:

الجملة **I**. في المثلث HGL الزاوية H والزاوية L متتامتان

الجملة **II** . المثلث HGL هو مثلث قائم الزاوية في G

– أي من الخيارات التالية صحيحة؟

A). إذا كانت الجملة I صحيحة، فإن الجملة II خاطئة

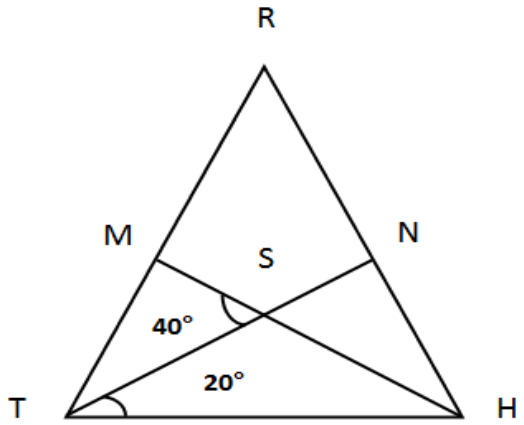
B). إذا كانت الجملة I خاطئة، فإن الجملة II صحيحة

C). إذا كانت الجملة I صحيحة، فإن الجملة II صحيحة

D). الجملتين I وII لا يمكن أن يكون كلا منهما خاطئاً

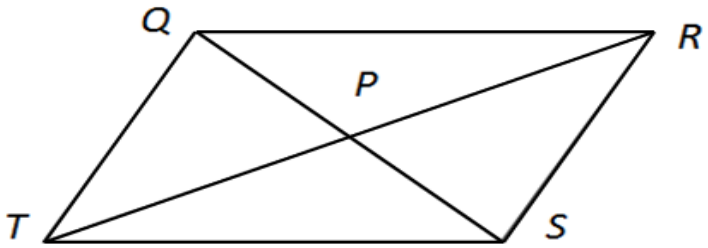
المفردة 33: في الرسم التالي يتقاطع مستقيما الإرتفاع (TN) و (HM) للمثلث RTH عند النقطة S .
والزاوية $M\hat{S}T$ قيسها 40°، والزاوية $S\hat{T}H$ قيسها 20°.

– أي المعلومات **<u>تكفي للبرهان</u>** على أن المثلث RTH هو مثلث متساوي الساقين؟



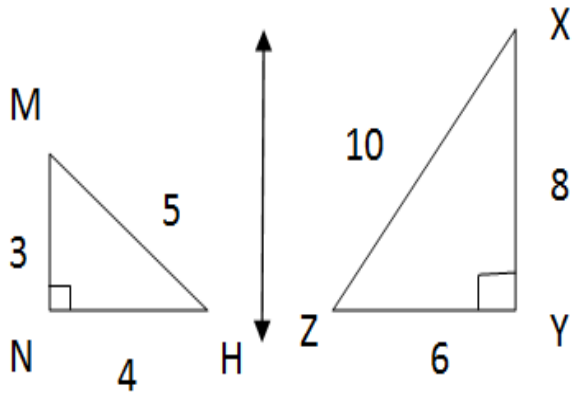
A). اثبات تقايس الأضلاع الثلاثة في المثلث RTH
B). اثبات أن في المثلث RTH زاويتين متقايستين
C). اثبات تقايس الزوايا الثلاثة في المثلث RTH
D). اثبات أن في المثلث RTH زاويتين متكاملتين

المفردة 34: أي من المعلومات **<u>تكفي لإثبات</u>** أن يكون الرباعي (QRST) متوازي أضلاع؟



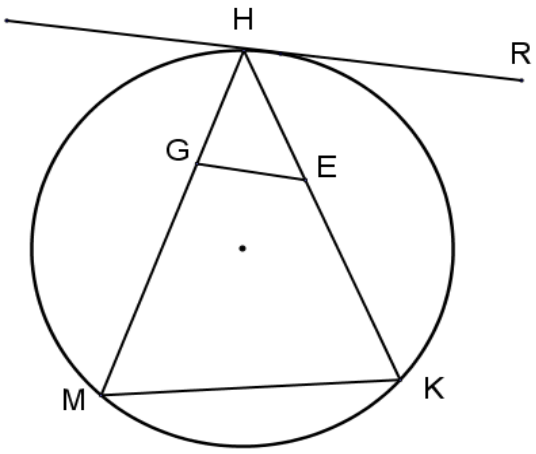
A). $QP = PS$ و $TP = PR$
B). $QR \neq ST$
C). $QR = ST$
D). ضلعان متجاوران متقايسان

المفردة 35: لإثبات أن المثلثين XYZ وMNH متشابهان، يكفي أن تكون، أي من الخيارات التالية صحيحة؟



A). الأضلاع المتناظرة متساوية في القياس
B). الزوايا المتناظرة متساوية في القياس
C). الأضلاع المتناظرة متناسبة
D). الأضلاع والزوايا المتناظرة متساوية في القياس

المفردة 36: إليك المثلث HKM مثلث مرسوم داخل دائرة و (HR) مماس لها. المستقيم (EG) يقطع (HK) في E، و (HM) في G. من المعطيات السابقة **يمكن إثبات** أن الشكل EKMG رباعي دائري **إذا كان**:



A). HE = HG
B). HR // GE
C). KM // EG
D). KM = RH

المفردة 37: اليك الجملتين التاليتين:

الجملة 1 **:** المثلث KMH متقايس الأضلاع رأسه الأساسي K.

الجملة 2 **:** في المثلث KMH، الزاوية $\hat{M}$ و $\hat{H}$ متقايستان.

– أي من الخيارات التالية **صحيح** ؟

A). إذا كانت الجملة 1 صحيحة، فإن الجملة 2 صحيحة

B). إذا كانت الجملة 2 صحيحة، فإن الجملة 1 صحيحة

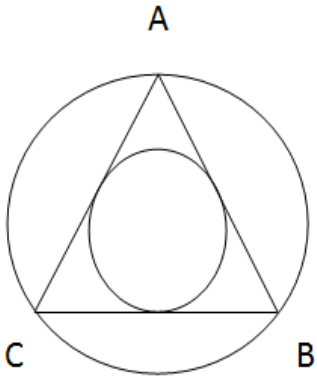
C). إذا كانت الجملة 1 خاطئة، فإن الجملة 2 خاطئة

D). لا يمكن أن تكون الجملتين 1 و2 صحيحتين معاً

المفردة 38: في الشكل المقابل دائرتان متحدتا المركز M، المثلث ABC أضلاعه تمس الدائرة M. لهما نفس المركز.

– برهن على أن المثلث ABC متساوي الأضلاع.

أي العبارات التالية تعتقد أنها غير لازمة للبرهان:



A). أنصاف أقطار الدائرة الواحدة متساوية في الطول

B). المماس لدائرة يكون عموديا على نصف القطر المرسوم من نقطة التماس

C). في الدائرة الواحدة اذا كانت الأوتار على أبعاد متساوية من المركز فإنها تكون متساوية في الطول

D). الأوتار المتساوية في الطول في دائرة على أبعاد متساوية من المركز

المفردة 39: فيما يلي ثلاث خصائص لشكل ما:

الخاصية **F** : له قطران متعامدان.

الخاصية G : هو مربع.

الخاصية **H** : هو معين.

– أي من الخيارات التالية **صحيح**؟

A). F تستلزم G و G تستلزم H

B). F تستلزم H و H تستلزم G

C). G تستلزم H و H تستلزم F

D). H تستلزم F و F تستلزم G

المفردة 40: فيما يلي جملتان:

الجملة 1: إذا كان الشكل مستطيلاً، فإن قطراه ينصف كل منهما الآخر.

الجملة 2: إذا كانت أقطار شكل ما ينصف كل منهما الآخر، فإن الشكل مستطيل.

– أي من من الخيارات التالية **صحيح**؟

A). لإثبات أن الجملة 1 صحيحة، يكفي أن نثبت أن الجملة 2 صحيحة

B). لإثبات أن الجملة 2 صحيحة، يكفي أن نثبت أن الجملة 1 صحيحة

C). لإثبات أن الجملة 2 صحيحة، يكفي أن نجد مستطيلاً واحداً قطراه ينصف كل منهما لآخر

D). لإثبات أن الجملة 2 خاطئة، يكفي إيجاد شكلاً واحداً ليس مستطيلاً قطراه ينصف كل منهما الآخر

انتهى

الملحق رقم(02)

جدول يوضح قائمة بأسماء الأساتذة الخبراء المحكمين للاختبار.

اللقب والاسم	الدرجة العلمية	مكان العمل
1. منصر تيجاني	دكتوراه في الرياضيات	جامعة محمد خيضر بسكرة
2. غمام العيد	دكتوراه في الرياضيات	جامعة حمه لخضر الوادي
3. سعيدة عدوان	ماجستير في الرياضيات	جامعة محمد خيضر بسكرة
4. باية لعجال	ماجستير في الرياضيات	جامعة محمد خيضر بسكرة
5. مداس الطيب	مفتش تربية وطنية لمادة الرياضيات	مديرية التربية الوادي
6. جلالي لزهاري	مفتش تربية وطنية لمادة الرياضيات	مديرية التربية الوادي
7. دليلة بورحلي	أستاذ متوسط مكون في مادة الرياضيات	متوسطة مي زيادة بالمسيلة
8. رشيد بن قدور	أستاذ تعليم ثانوي في مادة الرياضيات	ثانوية حساني لخضر بجامعة
9. آسيا بن زاوي	أستاذ تعليم ثانوي في مادة الرياضيات	ثانوية حساني لخضر بجامعة
10. زوبير كروة	أستاذ تعليم ثانوي في مادة الرياضيات	ثانوية متقنة جامعة– الوادي

الملحق رقم (03)

كراس اختبار مستويات التفكير الهندســـي لفان هيل – في صورته النهائية –

إعــداد البـــاحث

زيـــاد رشيـــد

تعليمات الاجابة على الاختبار

الرجاء عدم فتح كراس الاختبار قبل إعلامك بذلك.

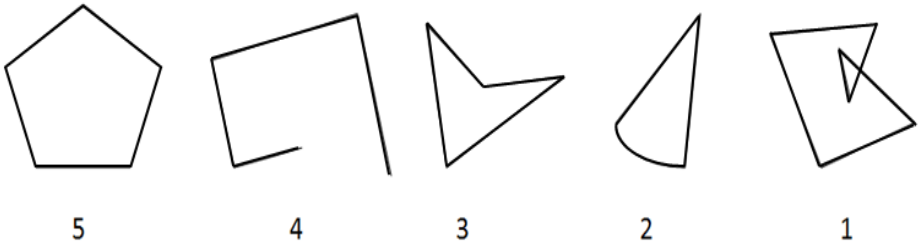
عزيزي / التلميذ، اليك اختبار مكون من 28 مفردة (9 صفحة – بما فيها هذه الصفحة) . قد لا تستطيع الإجابة على كل مفردة في هذا الاختبار، ولكن يرجى بذل أكبر جهد ممكن للإجابة على كل مفردة . يهدف الامتحان إلى بناء اختبار لقياس مستويات التفكير الهندســـي لدى تلامذة المرحلة الثانوية بالجزائر وفقاً لنظرية (فان هيل). لا علاقة لنتيجة هذا الاختبار بتقديرك في المدرسة . تستخدم نتائج هذا الاختبار لغرض البحث العلمي فقط . وقد يؤثر على كيفية تقديم المعلومات في بناء هذا الاختبار لقياس مستويات التفكير الهندســـي في البيئة الجزائرية لاحقاً . نشكرك على تعاونك معنا مسبقاً. معك 5 دقائق من الآن لتعبئة المعلومات في ورقة الاجابة الملحقة مع كراس الامتحان. عندما يتم اعلامك بأن تبدأ بالاجابة:

1. اقرأ كل مفردة بعناية . اقرأ جميع خيارات الإجابة المطروحة أمام كل مفردة.
2. قرر أي إجابة هي تلك التي تعتقد أنها صحيحة (يوجد إجابة واحدة صحيحة لكل مفردة) . ضع دائرة حول حرف (A ،B ،C ، D) ، رمز الإجابة الصحيحة) في ورقة الاجابة.
3. استخدم الفراغ على ورقة الإجابة للرسم. لا تكتب على كراس الاختبار.
4. إذا أردت تغيير أي إجابة، امسح الإجابة الأولى تماماً.
5. لا تخمن في الإجابة.
6. إذا احتجت قلم رصاص آخر، أرفع يدك لطلب ذلك.
7. وقت الإختبار: 60 دقيقة (ساعة)، الرجاء التوقف عن الكتابة ووضع الأقلام جانباً عند الإعلان عن نهاية الوقت.
8. سيتم اعلامك بالوقت في نصف الوقت المحدد (30 دقيقة)، و قبل إنتهاء الوقت بـــ15 دقائق.
9. لا تنس تعبئة المعلومات الشخصية على ورقة الإجابة.

إبدأ، بالتوفيق.

مفردات اختبار مستويات التفكير الهندســـــي

المفردة 01: أي من الأشكال التالية مضلع؟



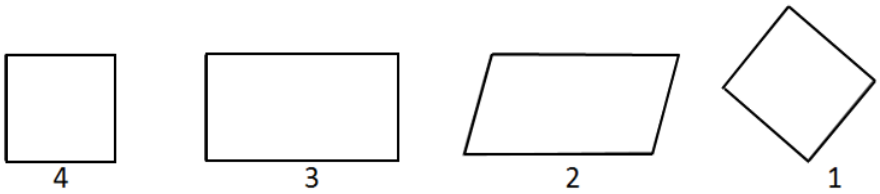
A). 1 و3 و5 فقط

B). 3 و5 فقط

C). 2 و4 فقط

D). جميعها مضلعات

المفردة 02: أي من الأشكال التالية مربع؟



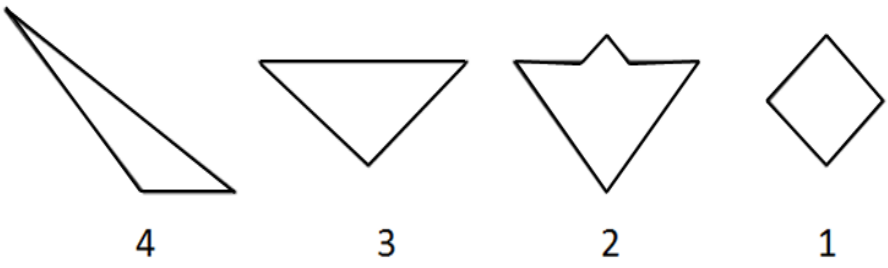
A). 4 فقط

B). 1 و4 فقط

C). 4 و3 فقط

D). جميعها مربعات

المفردة 03: أي من الأشكال التالية مثلث؟



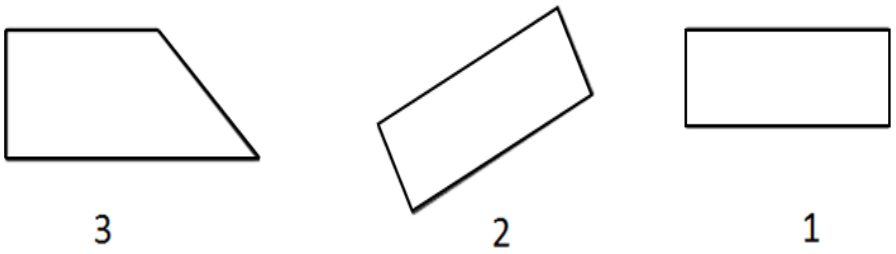
A). ليس أياً منها مثلثاً

B). 3 فقط

C). 4 فقط

D). 3 و 4 فقط

المفردة 04 : أي من الأشكال التالية مستطيل؟



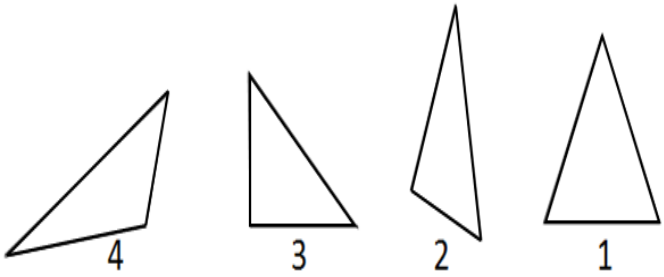
A). 1 فقط

B). 2 فقط

C). 1 و3 فقط

D). 1 و2 فقط

المفردة 05: أي من الاشكال التالية مثلث متساوي الساقين؟



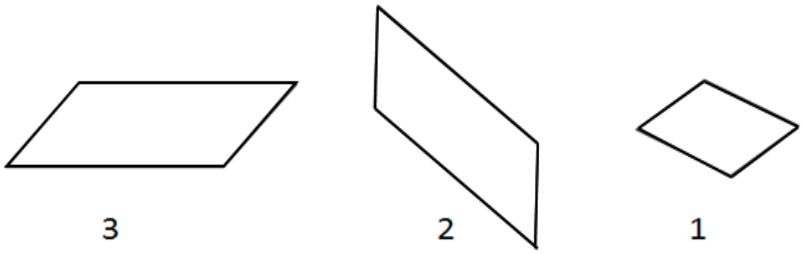
A). المثلث 1 و3 فقط

B). المثلث 3 و4 فقط

C). المثلث 1 و2 فقط

D). جميعها مثلثات متساوية الساقين

المفردة 06: أي من الأشكال التالية متوازي أضلاع؟



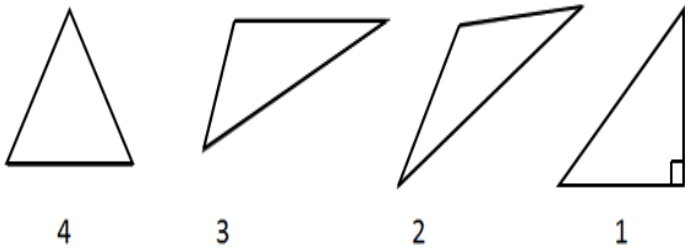
A). الشكل 3 فقط

B). الشكل 1 فقط

C). الشكل 2 و3 فقط

D). جميعها متوازيات أضلاع

المفردة 07: أي من الاشكال التالية مثلث متساوي الاضلاع؟



A). المثلث 4 فقط

B). المثلث 3 و2 فقط

C). المثلث 1 و4 فقط

D). جميعها مثلثات متساوية الأضلاع

المفردة 08: قيس الزاوية المركزية **يساوي**:

A). ضعفي قيس الزاوية المحيطية المشتركة معها بنفس القوس

B). نصف قيس الزاوية المحيطية المشتركة معها بنفس القوس

C). ضعفي قيس أي زاوية محيطية

D). نصف قيس أي زاوية محيطية

المفردة 09: أي من العبارات التالية **صحيحة** لكل متوازي أضلاع؟

A). القطران متقايسان

B). الزوايا المتقابلة فيه متقايسة

C). القطران متعامدان

D). زواياه متقايسة

المفردة 10: أي من العبارات التالية **صحيحة** لكل شبه منحرف متساوي الساقين؟

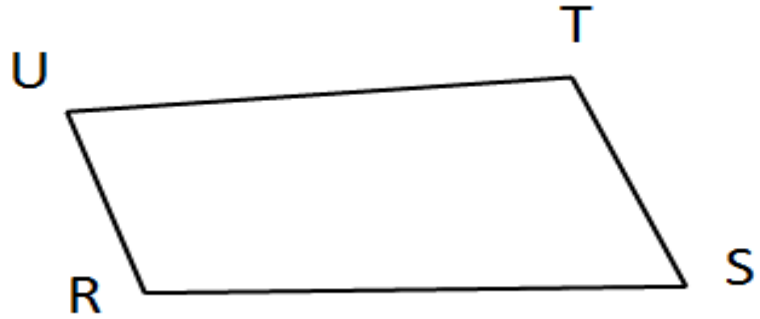
A). قطراه متساويان

B). زاويتا القاعدتين متطابقتان

C). فيه تماثل انعكاسي وخط تماثله يمر في منتصفي قاعدتيه

D). جميع ما ورد أعلاه صحيح في كل شبه منحرف متساوي الساقين

المفردة 11: يعطى الرباعي RSTU. اذا كان (UT) لا يوازي (RS) و (UR) يوازي (ST).
إذن الرباعي RSTU **يجب أن يكون**؟



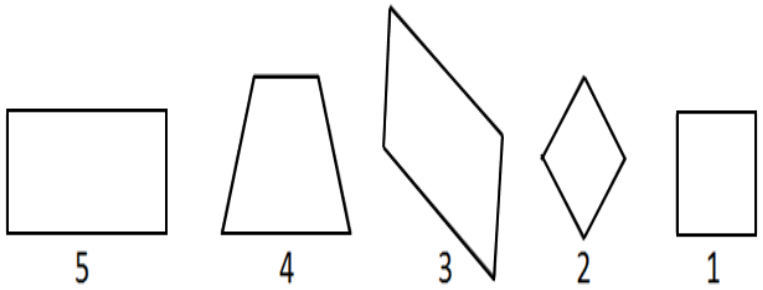
A). شبه منحرف

B). مستطيل

C). معين

D). متوازي الأضلاع

المفردة 12: أي من الأشكال التالية تصنف متوازي أضلاع؟



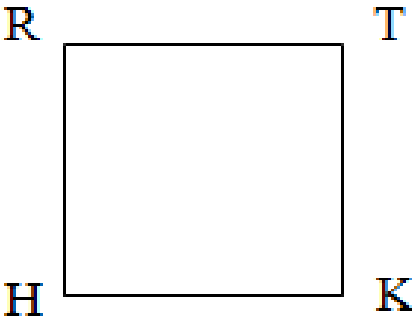
A). الشكل 2 و3 و4 و5 فقط

B). الشكل 1 و2 و3 و4 فقط

C). الشكل 1 و2 و3 و5 فقط

D). جميعها متوازيات أضلاع

المفردة 13: RTKH مربع، أي من العلاقات التالية **صحيحة** في كل مربع؟



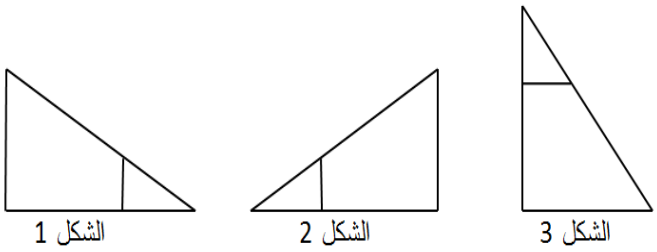
A). RK و KH متساويان

B). TH و RK متعامدان

C). RH و TK متعامدان

D). RH و TH متساويان

المفردة 14: الشكل (3) هو صورة الشكل (1) بإستعمال **تحويلين نقطيين** على الترتيب (الشكل1؛ ثم الشكل2؛ ثم الشكل3). أي الخيارات التالية **صحيح**؟



A). تناظر ثم إنسحاب

B). تناظر ثم دوران $90°$ في إتجاه عقارب الساعة

C). دوران $180°$ ثم إنسحاب

D). دوران $90°$ عكس إتجاه عقارب الساعة ثم تناظر

المفردة 15: ما هي الخاصية التي تنطبق على جميع المستطيلات ولا تنطبق على بعض متوازيات الأضلاع؟

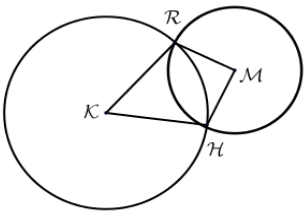
A). الضلعان المتقابلان فيه متساويان في القياس

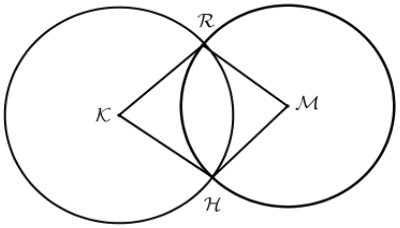
B). الزاويتان المتقابلتان فيه متساويتان في القياس

C). القطران متساويان في القياس

D). الضلعان المتقابلان فيه متوازيان

المفردة 16: النقطتان M ، K مركزا دائرتين تتقاطعان عند R ، H وينتج عنهما شكل رباعي MHKR.





- أي الخيارات التالية **ليس صحيحاً دائما**؟

A). هناك ضلعان في الشكل MHKR متقايسان

B). في الشكل MHKR ، هناك زاويتان على الاقل متقايستان

C). المستقيمان MK و HR متعامدان

D). قيس الزاوية M يساوي قيس الزاوية K

المفردة 17: أي من العبارات التالية **صحيحة**؟

A). كل متوازي أضلاع معين

B). قطرا المعين متعامدان ومتساويان دائما

C). كل معين متوازي أضلاع

D). أطوال أقطار المعين متساوية

المفردة 18: أي من العبارات التالية **صحيحة**؟

- الزاويتان المتقابلتان في:

A). الرباعي الدائري متتامتان

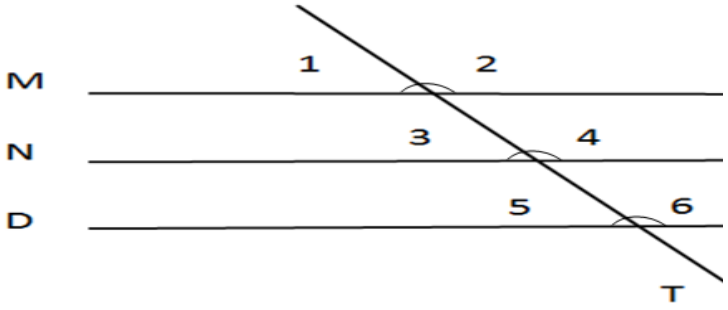
B). المعين متكاملتان

C). المستطيل متتامتان

D). متوازي الاضلاع متقايستان

المفردة 19: رسم أحمد ثلاثة مستقيمات متوازية (M)، (N)، (D)، ورسم مستقيما (T) قاطعا لهما، تقاطع المستقيم (T) مع المستقيمات (M)، (N)، (D) شكل زوايا وهي مرقمة في الرسم البياني أدناه.

- أي من **التخمين** الذي يتشكل عند أحمد حول شكل الزوايا الناتجة عن هذا التقاطع **صحيح**؟



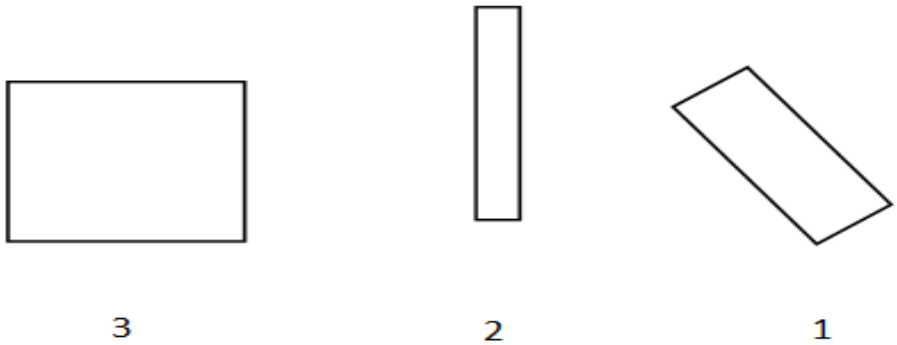
A). الزوايا 1 و2 و3 متقايسة

B). الزوايا 2 و 4 متكاملتان

C). الزوايا 1 و 3 و 5 متقايسة

D). الزوايا 1 و5 متكاملتان

المفردة 20: أي من الأشكال التالية يمكن اعتباره مستطيلاً؟



A). جميعها

B). 1 فقط

C). 3 و 2 فقط

D). 2 و 1 فقط

المفردة 21: فيما يلي جملتان:

الجملة 1: الشكل R هو مستطيل

الجملة 2: الشكل R هو مثلث

– أي من من الخيارات التالية **صحيح**؟

A). إذا كانت الجملة 1 صحيحة، فإن الجملة 2 صحيحة

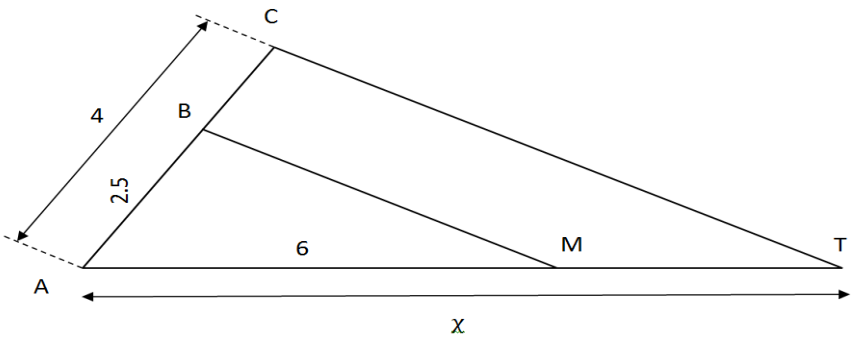
B). إذا كانت الجملة 1 خاطئة، فإن الجملة 2 صحيحة

C). لا يمكن أن تكون الجملتان 1 و 2 صحيحتين معاً

D). لا يمكن أن تكون الجملتان 1 و 2 خاطئتين معاً

المفردة 22: في الشكل المقابل القطعة [BM] // [CT]، و [AB] ، [AC] و [AM] ثلاثة قطع أطوالها K ، H ، R على الترتيب بحيث K=2.5cm ، H=4cm ، R=6cm.

– استنتج الطول χ للقطعة [AT] ؟



A). 10.6cm

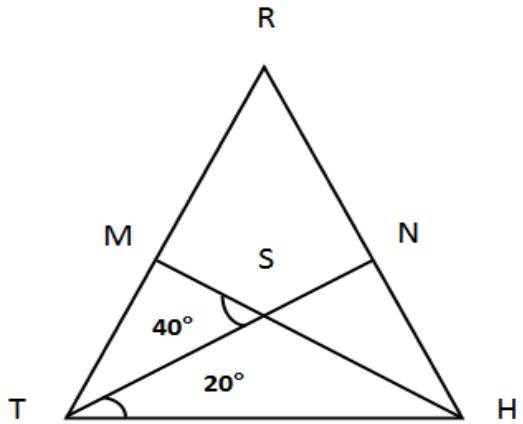
B). 9cm

C). 9.6cm

D). 10 cm

المفردة 23: في الرسم التالي يتقاطع مستقيما الارتفاع (TN) و (HM) للمثلث RTH عند النقطة S . والزاوية $M\hat{S}T$ قيسها 40°، والزاوية $S\hat{T}H$ قيسها 20°.

– أي المعلومات **تكفي للبرهان** على أن المثلث RTH هو مثلث متساوي الساقين؟



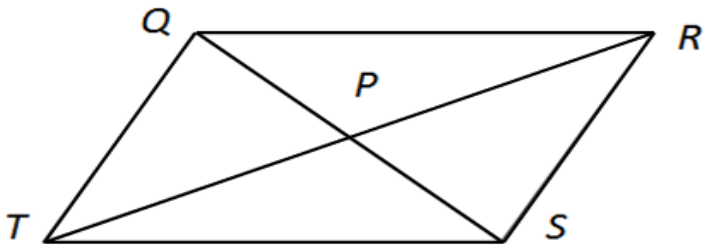
A). إثبات تقايس الأضلاع الثلاثة في المثلث RTH

B). إثبات أن في المثلث RTH زاويتين متقايستين

C). إثبات تقايس الزوايا الثلاثة في المثلث RTH

D). إثبات أن في المثلث RTH زاويتين متكاملتين

المفردة 24: أي من المعلومات **تكفي لإثبات** أن يكون الرباعي (QRST) متوازي أضلاع؟



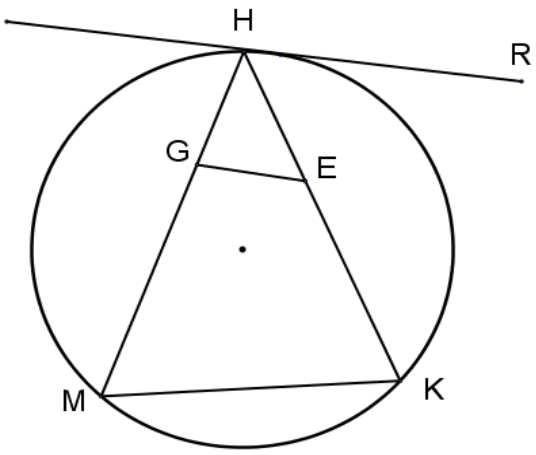
A). $QP = PS$ و $TP = PR$

B). $QR \neq ST$

C). $QR = ST$

D). ضلعان متجاوران متقايسان

المفردة 25: إليك المثلث HKM مثلث مرسوم داخل دائرة و (HR) مماس لها. المستقيم (EG) يقطع (HK) في E، و (HM) في G. من المعطيات السابقة **يمكن إثبات** أن الشكل EKMG رباعي دائري **إذا كان:**



A). $HE = HG$

B). $HR // GE$

C). $KM // EG$

D). $KM = RH$

المفردة 26: اليك الجملتين التاليتين:

الجملة 1 : المثلث KMH متقايس الاضلاع رأسه الأساسي K.

الجملة 2 : في المثلث KMH، الزاوية $\hat{M}$ و $\hat{H}$ متقايستان.

– أي من الخيارات التالية **صحيح** ؟

A). إذا كانت الجملة 1 صحيحة، فإن الجملة 2 صحيحة

B). إذا كانت الجملة 2 صحيحة، فإن الجملة 1 صحيحة

C). إذا كانت الجملة 1 خاطئة، فإن الجملة 2 خاطئة

D). لا يمكن أن تكون الجملتين 1 و2 صحيحتين معاً

المفردة 27: فيما يلي جملتان:

الجملة 1: إذا كان الشكل مستطيلاً، فإن قطراه ينصف كل منهما الآخر.

الجملة 2: إذا كانت أقطار شكل ما ينصف كل منهما الآخر، فإن الشكل مستطيل.

- أي من من الخيارات التالية **صحيح**؟

A). لإثبات أن الجملة 1 صحيحة، يكفي أن نثبت أن الجملة 2 صحيحة

B). لإثبات أن الجملة 2 صحيحة، يكفي أن نثبت أن الجملة 1 صحيحة

C). لإثبات أن الجملة 2 صحيحة، يكفي أن نجد مستطيلاً واحداً قطراه ينصف كل منهما لآخر

D). لإثبات أن الجملة 2 خاطئة، يكفي إيجاد شكلاً واحداً ليس مستطيلاً قطراه ينصف كل منهما الآخر

المفردة 28: فيما يلي ثلاث خصائص لشكل ما:

الخاصية **F** : له قطران متعامدان.

الخاصية G : هو مربع.

الخاصية **H** : هو معين.

- أي من الخيارات التالية **صحيح**؟

A). F تستلزم G و G تستلزم H

B). F تستلزم H و H تستلزم G

C). G تستلزم H و H تستلزم F

D). H تستلزم F و F تستلزم G

الملحق رقم (04)

ورقة الاجابـــــة على الاختبار

- المعلومات الشخصية -

الإسم واللقب : ..... **الجنس :** ..... **تاريخ الميلاد :** ..... **القسم** :.....

تعليمات:

- اقرأ التعليمات الواردة في كراس الاختبار وورقة الاجابة بعناية.
- تأكد من وضع الدائرة في المربع المقابل لرمز الاجابة التي اخترتها.
- تأكد من الاجابة التي تختارها والمكان المناسب للتأشير في ورقة الاجابـــــة.

✓ ضع دائرة حول إجابة واحدة فقط صحيحة أمام كل مفردة في بدائلها الأربعة المقترحة.

المفردة	الخيارات (البدائل)				المفردة	الخيارات (البدائل)			
المفردة 01	D	C	B	A	المفردة 15	D	C	B	A
المفردة 02	D	C	B	A	المفردة 16	D	C	B	A
المفردة 03	D	C	B	A	المفردة 17	D	C	B	A
المفردة 04	D	C	B	A	المفردة 18	D	C	B	A
المفردة 05	D	C	B	A	المفردة 19	D	C	B	A
المفردة 06	D	C	B	A	المفردة 20	D	C	B	A
المفردة 07	D	C	B	A	المفردة 21	D	C	B	A
المفردة 08	D	C	B	A	المفردة 22	D	C	B	A
المفردة 09	D	C	B	A	المفردة 23	D	C	B	A
المفردة 10	D	C	B	A	المفردة 24	D	C	B	A
المفردة 11	D	C	B	A	المفردة 25	D	C	B	A
المفردة 12	D	C	B	A	المفردة 26	D	C	B	A
المفردة 13	D	C	B	A	المفردة 27	D	C	B	A
المفردة 14	D	C	B	A	المفردة 28	D	C	B	A

الملحق رقم (05)

مفتاح تصحيح الاختبار

المفردة	الخيارات (البدائل)				المفردة	الخيارات (البدائل)			
	D	C	B	A		D	C	B	A
المفردة 01			■		المفردة 15		■		
المفردة 02			■		المفردة 16	■			
المفردة 03	■				المفردة 17		■		
المفردة 04	■				المفردة 18	■			
المفردة 05				■	المفردة 19		■		
المفردة 06	■				المفردة 20				■
المفردة 07				■	المفردة 21		■		
المفردة 08				■	المفردة 22		■		
المفردة 09			■		المفردة 23			■	
المفردة 10	■				المفردة 24				■
المفردة 11				■	المفردة 25		■		
المفردة 12		■			المفردة 26				■
المفردة 13			■		المفردة 27			■	
المفردة 14			■		المفردة 28			■	

الملحق رقم (06)

طلب رخصة تربص ميداني للباحث من قسم علم النفس والأرطوفونيا بجامعة وهران2 إلى مديرية التربية لولاية الوادي

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
كلية العلوم الاجتماعية
Faculté des sciences sociales

قسم علم النفس والأرطوفونيا
مصلحة ما بعد التدرج

المرجع: / 2017 التاريخ: 2017/01/15

إلى السيد

مدير التربية لولاية الوادي

موضوع : **طلب رخصة تربص ميداني .**

في إطار تحضير(ها) أطروحة الدكتوراه في علم النفس.

تخصص : **القياس والتقويم**

موضوع الرسالة : " **إستخدام نموذج راش اللوغريتمي أحادي البارامتر ونظرية القياس الكلاسيكية في تحليل وتدريج بنود إختبار مستويات التفكير الهندسي مبني وفق نموذج فان هيل – دراسة سيكومترية مقارنة –** "

نرجو من سيادتكم السماح للطا لب(ة) : **زياد رشيد**

إجراء دراسته (ها) الميدا نية بمؤسستكم وذلك خلال الموسم الدراسي **2017/2016**.

مدة التربص: **من 15 جانفي 2017 إلى غاية نهاية الموسم الدراسي .**

و لكم جزيل الشكر

نائب رئيس القسم المكلف بمصلحة ما بعد التدرج.

د. بلعـــابد عبد القـــادر
نائب رئيس قسم علم النفس و الارطفونيا
- مابعد التدرج -

الملحق رقم (07)

ترخيص مقدم للباحث بالتطبيق الميداني في الثانويات المعنية من طرف مديرية التربية لولاية الوادي

الجمهورية الجزائرية الديمقراطية الشعبية

وزارة التربية الوطنية

مديرية التربية لولاية الوادي
مصلحة التنظيم التربـــوي
مكتب التعليـم الثـانـــوي
الرقم: 435 /2017/1.10

مدير التربية
إلى السادة:
مديري ثانويات :- حساني لخضر - المزارة
- عمراني العابد - تندلة - متقن جامعة

الموضـــوع : تـرخـيـص

المرجـــــع : مراسلة جامعة وهران قسم علم النفس و الارطوفونيا

تحت رقم:05 بتاريخ 2017/01/15

بناء على ما جاء في المرجع المشار إليه أعلاه ، يرخص مدير التربية للطالب **زياد رشيد** قسم علم النفس والارطوفونيا بإجراء دراسة ميدانية ابتداء من 2017/02/01 الى غاية نهاية السنة الدراسية في اطار اعداد رسالة الدكتوراه.

أطلب منكم تسهيل مهمته في إنجاز مهمتكم .

الوادي في: 2017/01/24

مدير التربية

حمادي كمال

الملحق رقم (08)

يوضح بعض مخرجات التحليل الإحصائي وفق نظرية الاستجابة للمفردة (نموذج راش) باستخدام برنامج وينستبس (Winsteps v4.0.0) لمايك لياكر (2017).

مؤشرات تحقق أحادية البعد للاختبار قبل التدريج

```
TABLE 23.1 ZIAD RACHID N=577.sav                ZOU462WS.TXT  Jul 22 2017  7:22
INPUT: 577 PERSON  28 ITEM  REPORTED: 577 PERSON  28 ITEM  2 CATS WINSTEPS 4.0.0
-------------------------------------------------------------------------------

     Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = ITEM information
units
                                          Eigenvalue   Observed   Expected
Total raw variance in observations     =      35.8189 100.0%         100.0%
  Raw variance explained by measures   =       7.8189  21.8%          22.0%
    Raw variance explained by persons  =       3.3385   9.3%           9.4%
    Raw Variance explained by items    =       4.4804  12.5%          12.6%
  Raw unexplained variance (total)     =      28.0000  78.2% 100.0%   78.0%
    Unexplned variance in 1st contrast =       1.6894   4.7%   6.0%

       STANDARDIZED RESIDUAL CONTRAST 1 PLOT

      -3         -2         -1          0          1          2          3
      -+----------+----------+----------+----------+----------+----------+- COUNT
CLUSTER
   .5 +                                 A B                                + 2      1
C     |                                  |                                 |
O  .4 +                                  |                                 +
N     |                                  |          C                      | 1      1
T  .3 +                                  D                                 + 1      1
R     |                                 E|   G     F                       | 3      1
A  .2 +                                  |                                 +
S     |                                  |                                 |
T  .1 +              H                   | I                               + 2      2
      |                                  |                                 |
1  .0 +------------------------------K----|-----J------LM-------------------+ 4      2
      |                                  N   n             m                | 3      2
L -.1 +               k                  |     l                           + 2      2
O     |                i h               |  j                              | 3      2
A -.2 +                                  f    g                            + 2      3
D     |                           e      |                                 | 1      3
I -.3 +                   c       d      |                                 + 2      3
N     |                       b          |                                 | 1      3
G -.4 +                                  |                                 +
      |                                  |         a                       | 1      3
  -.5 +                                  |                                 +
      -+----------+----------+----------+----------+----------+----------+-
      -3         -2         -1          0          1          2          3
                              ITEM MEASURE
 COUNT:              111 11   1   12   14 21212   21 11   1

                             1  3 44132514131211 2 2 1  1 1
PERSON             1  1 51 74845420412530416601139725924111 3 115  5     3
                          T       S        M       S        T
%TILE              0             10 20 40 60 70  80 90            99
```

مؤشرات التمييز والتخمين لمفردات الاختبار قبل تدريجه

TABLE 14.1 ا- الاقسام_لكل -577وزي- الكلي-ين- ZOU992WS.TXT Jun 29 2017 18:37
INPUT: 577 PERSON 28 ITEM REPORTED: 577 PERSON 28 ITEM 2 CATS WINSTEPS 4.0.0

PERSON: REAL SEP.: 1.36 REL.: .65 ... ITEM: REAL SEP.: 9.09 REL.: .99

ITEM STATISTICS: ENTRY ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-AL CORR.	PTMEASUR-AL EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	ESTIM DISCR	ASYMPTOTE LOWER	ASYMPTOTE UPPER	ITEM
1	302	564	.00	.09	1.02	.8	.99	-.3	.31	.33	59.9	63.5	.94	.08	1.00	ITEM1
2	301	575	.05	.09	1.09	3.2	1.09	2.3	.23	.33	56.2	63.4	.59	.06	.88	ITEM2
3	485	573	-1.72	.12	.95	-.6	.89	-1.0	.29	.23	85.0	84.6	1.05	.00	1.00	ITEM3
4	285	576	.18	.09	1.09	3.4	1.09	2.6	.22	.33	59.2	63.6	.56	.12	.92	ITEM4
5	450	568	-1.32	.11	.94	-1.0	.88	-1.4	.34	.26	79.8	79.4	1.09	.00	1.00	ITEM5
6	367	575	-.48	.09	.96	-1.1	.97	-.6	.36	.31	67.5	67.0	1.12	.00	1.00	ITEM6
7	479	575	-1.61	.12	.98	-.3	.92	-.7	.28	.24	83.7	83.3	1.03	.00	1.00	ITEM7
8	177	522	.90	.10	1.06	1.5	1.08	1.6	.26	.34	67.0	70.1	.83	.03	.82	ITEM8
9	299	570	.05	.09	.99	-.4	.98	-.6	.35	.33	63.5	63.4	1.06	.02	1.00	ITEM9
10	171	536	1.02	.10	.99	-.1	1.00	.1	.34	.34	72.2	71.4	1.01	.01	1.00	ITEM10
11	448	573	-1.26	.11	.94	-1.0	.87	-1.6	.35	.27	79.1	78.4	1.10	.09	1.00	ITEM11
12	472	574	-1.53	.11	.93	-1.0	.84	-1.6	.35	.25	82.4	82.3	1.09	.00	1.00	ITEM12
13	360	571	-.45	.09	.92	-2.7	.89	-2.5	.42	.31	70.6	66.5	1.30	.00	1.00	ITEM13
14	243	567	.48	.09	1.01	.3	1.01	.3	.33	.34	65.3	65.5	.97	.00	.96	ITEM14
15	187	562	.94	.10	.93	-1.9	.93	-1.4	.42	.33	73.7	70.4	1.17	.00	1.00	ITEM15
16	251	564	.41	.09	1.02	.8	1.04	1.0	.31	.34	65.1	65.0	.89	.03	.99	ITEM16
17	268	560	.26	.09	.97	-1.2	.97	-1.0	.38	.34	64.8	64.0	1.15	.00	1.00	ITEM17
18	365	559	-.55	.09	1.01	.4	1.02	.4	.29	.31	68.3	67.9	.95	.10	.99	ITEM18
19	410	562	-.95	.10	.98	-.4	.96	-.5	.31	.29	73.3	73.8	1.04	.09	1.00	ITEM19
20	123	576	1.62	.11	.86	-2.4	.79	-2.8	.49	.31	81.3	79.5	1.19	.00	1.00	ITEM20
21	277	562	.18	.09	1.08	3.0	1.14	3.8	.22	.34	61.0	63.7	.56	.10	.90	ITEM21
22	233	557	.53	.09	1.00	.0	1.02	.5	.33	.34	66.1	66.0	.98	.00	.93	ITEM22
23	306	564	-.04	.09	.99	-.5	.99	-.2	.34	.33	64.5	63.6	1.06	.00	1.00	ITEM23
24	318	570	-.11	.09	1.02	.7	1.02	.6	.30	.33	65.6	63.9	.91	.02	.98	ITEM24
25	228	555	.55	.09	1.03	1.1	1.04	1.1	.29	.34	64.5	66.2	.87	.00	.87	ITEM25
26	250	562	.40	.09	1.06	2.0	1.05	1.3	.27	.34	61.2	64.9	.77	.07	.99	ITEM26
27	147	554	1.30	.10	1.00	.0	.96	-.5	.34	.33	75.6	75.3	1.01	.00	1.00	ITEM27
28	159	537	1.15	.10	1.12	2.5	1.22	3.4	.17	.33	71.5	72.9	.74	.06	.83	ITEM28
MEAN	298.6	563.0	.00	.10	1.00	.2	.99	.1			69.6	70.0				
P.SD	104.2	12.8	.90	.01	.06	1.6	.09	1.6			7.8	6.8				

جدول مؤشر الإحصائي Q3 لـ(ين) للإستقلال الموضعي

LARGEST STANDARDIZED RESIDUAL CORRELATIONS
USED TO IDENTIFY DEPENDENT ITEM

CORREL-ATION	ENTRY NUMBER	ITEM	ENTRY NUMBER	ITEM
-.17	10	ITEM10	15	ITEM15
-.16	2	ITEM2	15	ITEM15
-.13	4	ITEM4	19	ITEM19
-.12	4	ITEM4	18	ITEM18
-.12	2	ITEM2	17	ITEM17
-.12	1	ITEM1	14	ITEM14
-.12	6	ITEM6	21	ITEM21
-.12	2	ITEM2	23	ITEM23
-.12	3	ITEM3	16	ITEM16
-.12	6	ITEM6	16	ITEM16
-.12	4	ITEM4	13	ITEM13
-.12	20	ITEM20	26	ITEM26
-.11	15	ITEM15	24	ITEM24
-.11	23	ITEM23	27	ITEM27
-.11	2	ITEM2	19	ITEM19
-.11	18	ITEM18	26	ITEM26
-.11	8	ITEM8	13	ITEM13
-.11	12	ITEM12	15	ITEM15
-.11	9	ITEM9	21	ITEM21
-.11	14	ITEM14	25	ITEM25

مؤشرات الصعوبة و إحصاءات المطابقة لمفردات الاختبار بعد تدريجه

TABLE 10.1 ZIAD RACHID N=566_6نود-ال-ذوف-م.sav ZOU347WS.TXT Jul 8 2017 18:15
INPUT: 566 PERSON 22 ITEM REPORTED: 566 PERSON 22 ITEM 2 CATS WINSTEPS 4.0.0

PERSON: REAL SEP.: 1.30 REL.: .63 ... ITEM: REAL SEP.: 9.13 REL.: .99

ITEM STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-AL CORR.	PTMEASUR-AL EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	ITEM
6	173	511	1.06	.10	1.09	2.0	1.13	2.1	A .27	.37	69.4	71.0	ITEM8
19	310	559	.05	.09	1.05	1.7	1.10	2.4	B .29	.35	62.3	65.0	ITEM24
20	226	544	.69	.09	1.05	1.4	1.10	2.2	C .31	.37	65.1	67.0	ITEM25
21	246	551	.54	.09	1.07	2.2	1.07	1.6	D .29	.36	62.3	66.0	ITEM26
1	295	553	.15	.09	1.04	1.4	1.02	.5	E .32	.36	62.4	64.8	ITEM1
8	172	526	1.15	.10	1.04	.9	1.04	.8	F .33	.37	70.2	71.7	ITEM10
22	146	544	1.46	.10	1.03	.5	1.00	.0	G .34	.36	74.9	75.5	ITEM27
13	247	553	.54	.09	1.01	.4	1.02	.4	H .35	.37	65.3	66.1	ITEM16
11	241	556	.60	.09	1.00	.2	1.01	.3	I .36	.37	67.0	66.4	ITEM14
15	355	548	-.40	.10	1.00	.1	1.01	.2	J .33	.33	68.9	68.3	ITEM18
17	229	546	.68	.09	1.00	.0	1.01	.1	K .37	.37	67.1	67.0	ITEM22
18	303	553	.07	.09	.99	-.3	1.01	.3	k .36	.35	65.2	64.9	ITEM23
5	475	564	-1.58	.12	.99	-.1	.94	-.5	j .27	.25	84.5	84.2	ITEM7
16	404	552	-.84	.10	.99	-.3	.97	-.3	i .32	.31	74.5	74.2	ITEM19
2	475	562	-1.61	.12	.97	-.4	.98	-.1	h .28	.25	84.8	84.5	ITEM3
4	361	564	-.36	.09	.98	-.5	.98	-.4	g .36	.34	66.2	67.9	ITEM6
7	294	559	.18	.09	.98	-.7	.96	-1.0	f .38	.36	67.3	64.7	ITEM9
14	266	549	.37	.09	.97	-1.0	.97	-.8	e .39	.36	67.8	65.1	ITEM17
3	442	557	-1.21	.11	.95	-.9	.87	-1.4	d .35	.28	80.4	79.6	ITEM5
10	465	563	-1.45	.12	.95	-.8	.84	-1.5	c .34	.26	82.4	82.6	ITEM12
9	441	562	-1.16	.11	.94	-1.0	.86	-1.6	b .36	.28	78.6	78.8	ITEM11
12	186	551	1.08	.10	.90	-2.6	.89	-2.0	a .47	.37	76.1	71.1	ITEM15
MEAN	306.9	551.2	.00	.10	1.00	.1	.99	.1			71.0	71.2	
P.SD	103.4	12.2	.93	.01	.04	1.1	.08	1.2			7.1	6.6	

مؤشرات تحقق أحادية البعد للاختبار بعد تدريجه

TABLE 23.0 ZIAD RACHID N=566.sav ZOU442WS.TXT Dec 22 2017 11:22
INPUT: 566 PERSON 22 ITEM REPORTED: 566 PERSON 22 ITEM 2 CATS WINSTEPS 4.0.0

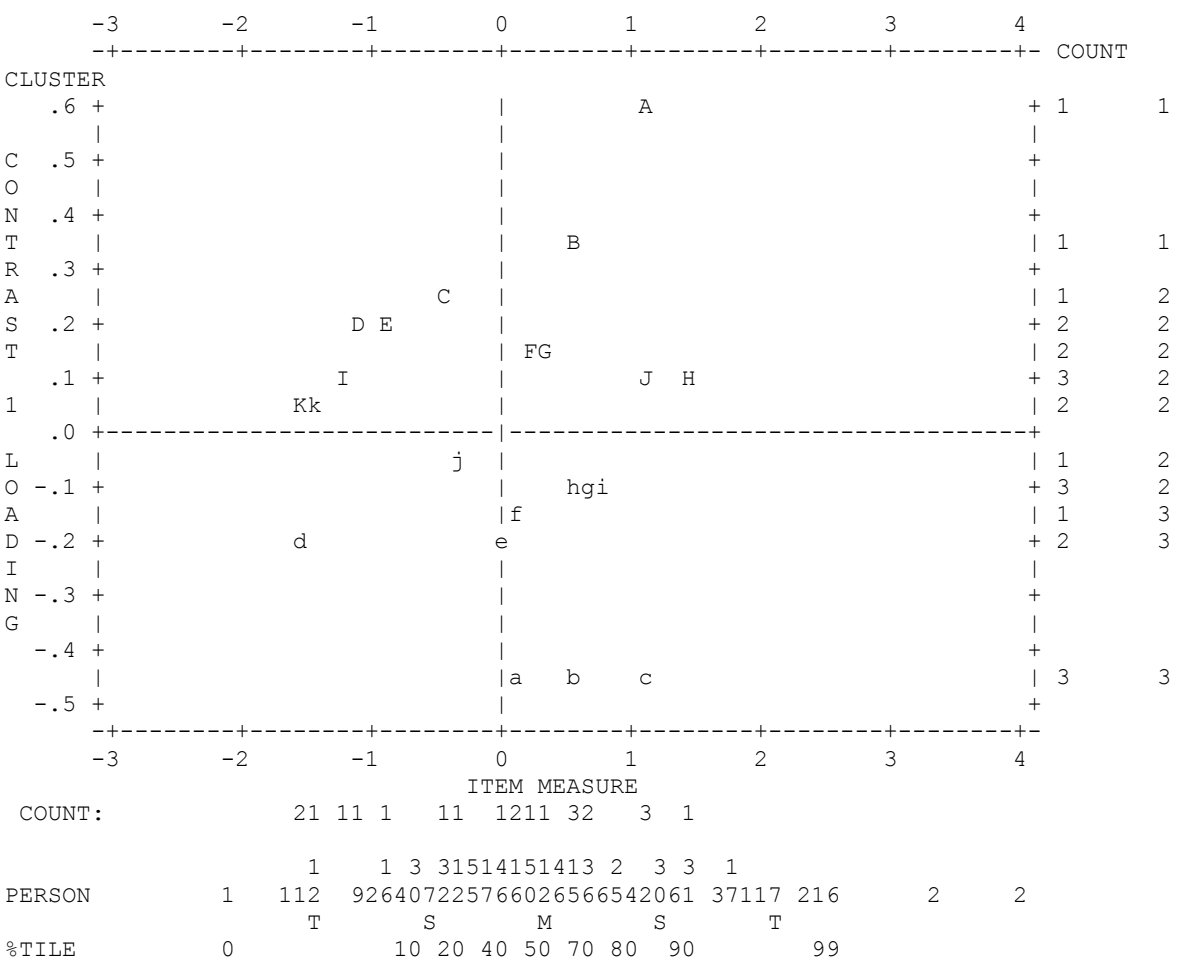
Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = ITEM information units

	Eigenvalue	Observed	Expected
Total raw variance in observations =	28.7690	100.0%	100.0%
Raw variance explained by measures =	**6.7690**	**23.5%**	**23.6%**
Raw variance explained by persons =	3.0125	10.5%	10.5%
Raw Variance explained by items =	3.7565	13.1%	13.1%
Raw unexplained variance (total) =	22.0000	76.5%	100.0% 76.4%
Unexplned variance in 1st contrast =	**1.4738**	**5.1%**	6.7%
Unexplned variance in 2nd contrast =	1.3617	4.7%	6.2%
Unexplned variance in 3rd contrast =	1.2770	4.4%	5.8%
Unexplned variance in 4th contrast =	1.2649	4.4%	5.7%
Unexplned variance in 5th contrast =	1.2062	4.2%	5.5%

Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = ITEM information units

	Eigenvalue	Observed		Expected
Total raw variance in observations =	28.7690	100.0%		100.0%
Raw variance explained by measures =	6.7690	23.5%		23.6%
Raw variance explained by persons =	3.0125	10.5%		10.5%
Raw Variance explained by items =	3.7565	13.1%		13.1%
Raw unexplained variance (total) =	22.0000	76.5%	100.0%	76.4%
Unexplned variance in 1st contrast =	1.4738	5.1%	6.7%	

STANDARDIZED RESIDUAL CONTRAST 1 PLOT



مؤشرات ثبات وفصل الأفراد والمفردات بعد التدريج

TABLE 3.1 ZIAD RACHID N=566_2) 6نـود ال-ذوف- م).s ZOU625WS.TXT Jun 28 2017 17:52
INPUT: 566 PERSON 22 ITEM REPORTED: 566 PERSON 22 ITEM 2 CATS WINSTEPS 4.0.0

SUMMARY OF 564 MEASURED (NON-EXTREME) PERSON

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	11.9	21.4	.30	.50	1.00	.0	.99	.0
P.SD	3.4	1.2	.84	.06	.17	.9	.26	.9
S.SD	3.4	1.2	.84	.06	.17	.9	.26	.9
MAX.	21.0	22.0	3.38	1.04	1.56	2.8	1.89	2.6
MIN.	3.0	14.0	-2.10	.47	.57	-2.7	.53	-2.6

REAL RMSE .52 TRUE SD .66 SEPARATION 1.26 PERSON RELIABILITY .61
MODEL RMSE .51 TRUE SD .67 SEPARATION 1.32 PERSON RELIABILITY .64
S.E. OF PERSON MEAN = .04

MAXIMUM EXTREME SCORE: 2 PERSON .4%

SUMMARY OF 566 MEASURED (EXTREME AND NON-EXTREME) PERSON

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	11.9	21.4	.32	.51				
P.SD	3.5	1.2	.88	.10				
S.SD	3.5	1.2	.88	.10				
MAX.	22.0	22.0	4.63	1.84				
MIN.	3.0	14.0	-2.10	.47				

REAL RMSE .53 TRUE SD .70 SEPARATION 1.30 PERSON RELIABILITY .63
MODEL RMSE .52 TRUE SD .71 SEPARATION 1.37 PERSON RELIABILITY .65
S.E. OF PERSON MEAN = .04

PERSON RAW SCORE-TO-MEASURE CORRELATION = .97
CRONBACH ALPHA (KR-20) PERSON RAW SCORE "TEST" RELIABILITY = .64 SEM = 2.09

SUMMARY OF 22 MEASURED (NON-EXTREME) ITEM

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	306.9	551.2	.00	.10	1.00	.1	.99	.1
P.SD	103.4	12.2	.93	.01	.04	1.1	.08	1.2
S.SD	105.9	12.5	.95	.01	.05	1.2	.08	1.2
MAX.	475.0	564.0	1.46	.12	1.09	2.2	1.13	2.4
MIN.	146.0	511.0	-1.61	.09	.90	-2.6	.84	-2.0

REAL RMSE .10 TRUE SD .92 SEPARATION 9.13 ITEM RELIABILITY .99
MODEL RMSE .10 TRUE SD .92 SEPARATION 9.20 ITEM RELIABILITY .99
S.E. OF ITEM MEAN = .20

ITEM RAW SCORE-TO-MEASURE CORRELATION = -1.00
Global statistics: please see Table 44.
UMEAN=.0000 USCALE=1.0000

الملحق رقم (09)

تواصل الباحث بالبريد الإلكترونـــي مع الدكتور يوسيسكين (Usiskin) مدير مشروع جامعة شيكاغو (1982).

ziad rachid <ziadpsy@gmail.com> 9 décembre 2015 à 16:46
À : z-usiskin@uchicago.edu

Dr Usiskin

I'm Ziad Rachid PhD student specialization Measurement and Evaluation at the University of Oran in Algeria, ask for help about **VAN HIELE GEOMETRY TEST** which is the objective of the PhD through the use of Rasch model and investigate both the Van Hiele theory for geometric development and an associated test. in the Algerian environment and compare it with the classical theory. I want information about consistency coefficient Cronbach's alpha coefficients and validity in the original or previous studies of the study.

Accept My sincere greetings and thanks

Zalman Usiskin <z-usiskin@uchicago.edu> 10 décembre 2015 à 01:09
À : ziad rachid <ziadpsy@gmail.com>
Cc : Sharon Senk <senk@math.msu.edu>

Dear Ziad:

We appreciate your interest in the van Hiele Geometry Test.

I do not know how much you know about the test. The test itself, and its creation, are described in detail at <http://ucsmp.uchicago.edu/resources/van-hiele/>. From that site you can download the report of the test. Reading the report is important because this is not a test in which the number of correct responses on the 25 questions constitutes the scale on which a van Hiele level for a student was identified. Furthermore, the test was not designed to test students but to test the van Hiele theory itself. Thus any psychometric properties that are not sufficiently established in the results may be due to any of three factors: a weakness of the test, a weakness in the van Hiele theory, and the brevity of the instrument. Furthermore, both the van Hiele theory and most Rasch analyses are based on an assumption that learning a subject is essentially one-dimensional. If the van Hiele theory is incorrect, then Rasch modeling would not fit well.

Well over 100 dissertations and papers have been written about studies where the test was used. A search of Dissertation Abstracts will locate many of these. The only reports that I know that tried to analyze the test's psychometric properties are in the May 1990 issue of the *Journal for Research in Mathematics Educaiton*. The articles are by Mark Wilson and Mary Crowley. In the same issue, Sharon Senk and I wrote a response: Evaluating a Test of van Hiele Levels: A Response to Crowley and Wilson, *Journal for Research in Mathematics Education* 21:3 (May 1990), 242-245. ERIC EJ 411 002.

If you wish to use the test, then you need to follow the instructions found on the website identified above.

I hope this information is useful to you and wish you the best in your doctoral work.

Sincerely,

Zalman Usiskin
Professor Emeritus of Education
Director, University of Chicago School Mathematics Project
The University of Chicago
1225 East 60th Street
Chicago, IL 60637 USA

Re: van Hiele question Boîte de réception x

Zalman Usiskin <z-usiskin@uchicago.edu> 07/08/2016
À moi

anglais › arabe Afficher le message traduit Ne pas traduire automatiquement cette langue : anglais x

Zalman Usiskin
z-usiskin@uchicago.edu
Afficher les détails

Dear Ziad:

I am sorry to have taken so long to respond to your e-mail.

Our work dealing with the van Hiele theory is described at <http://ucsmp.uchicago.edu/resources/van-hiele/>. There you will see that in the early 1980s, we developed a test (the van Hiele Geometry Test) in an attempt to determine the van Hiele levels of secondary school students in the United States. From that site, you can download the report "Van Hiele Levels and Achievement in Secondary School Geometry". That report contains a copy of the test, how to determine van Hiele levels from the responses students give to the items, and other information necessary to use the test.

On that site, you can also see the information we need if you would like permission to use the van Hiele Geometry Test or use a translation of it. The test has been translated into at least 7 languages, but neither Arabic nor French is among them.

I hope this information is helpful and would be happy to try to answer any questions you might have about the test or the theory.

Zalman Usiskin
Professor Emeritus of Education
Director, University of Chicago School Mathematics Project
The University of Chicago
1427 East 60th Street
Chicago, IL 60637 USA

الملحق رقم (10)

يوضح الحوار بين الباحث و مايك ليناكر(Mike Linacre) مطور برنامج (Winsteps) ومجموعة من الباحثين المهتمين في موقع التواصل الإجتماعي الفايس بوك.

، الساعة 06:56 ص2017 أبريل.17
Rachid Ziad

What is the acceptable value Item reliability and person reliability ?

Carlos Fernando Collares
Item reliability: 0.98; person reliability; 0.90 or 0.95 for high-stakes tests.

Trevor Bond
I am sorry; I cannot be so definitive. "Acceptable" means: acceptable for your purposes. you can defend your decisions. Reliability is over-used and under-powered. Separation and strata for items/persons are much more important than reliability

Carlos Fernando Collares
I agree. It depends on your purposes.

Rachid Ziad
thank you

Vijai K Pandey
Item reliability 0.98, person reliability 0.65 to 0.90 depending upon the construct, type of reliability and purpose of measurement.

Carlos Fernando Collares
I would say that person reliability should be ideally above 0.95 for high-stakes examinations.

Hynek Cígler
Item reliability is the indicator of test calibration stability (estimation of item difficulty parameters).
However, there couldn't be any cut-off for item reliability. Item reliability depends on the variance of item difficulty parameters. If all the items are at similar difficulty, item reliability could be low and this wouldn't mean that the measurement is bad.

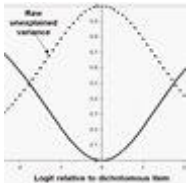
Mike Linacre
Generally, if item reliability is low, then your person sample size is not large enough to produce a robust item difficulty hierarchy. Look at your items. How many different levels of performance do they represent? Let's say "n". Then required minimum item reliability = n*n / (1 + n*n), similarly for the required minimum person reliability, but higher person reliability usually requires more items in the test.

Rachid Ziad

What are the accepted values in Variance explained by measures and unexplained variance and the first contrast When used Principal Contrast Analysis with winsteps ?

Mike Linacre

Rachid, expected values of variance explained for different person and item S.D.s are shown at www.rasch.org/rmt/rmt201a.htm Figure 4. The percent of variance explained varies from 20% to 80% for typical situations.

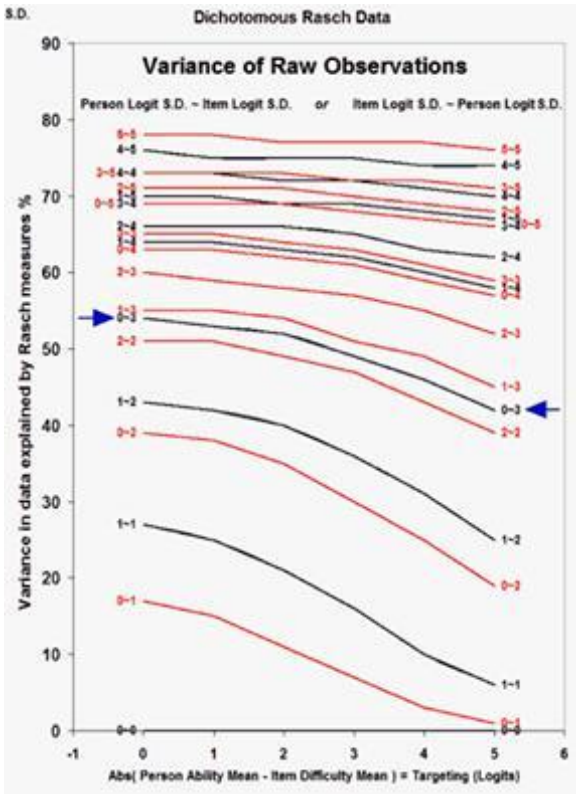


Data Variance Explained by Rasch Measures

RASCH.ORG

Mike Linacre

Here is Figure 4: www.rasch.org/rmt/gifs/rmt221ab.gif



RASCH.ORG