نوقشـــت رســـالة الطالبة ســـــمية ســويد

بعنـــــــوان :

أثـــر التفاعل بين طول الاختبار وخصائص العينة على الخصائص السيكومترية ودرجات القطع لاختبار الشخصية المتعدد الأوجه مينيسوتا M-M-P.I-2

وأجيزت يـــوم الاثنيـــــــن الواقع في ٢٠١٦/١/١١ من قبل السادة أعضـــاء لجنة الحكم التالية أسماؤهم :

الاســـــــــم	الصفـــــة	التوقيـــــع
د. إيمـــــان عـــز	عضواً مشرفاً	
د. عزيـــــزة رحمـــة	عضـــواً	
د. اعتـــدال عبد الله	عضـــواً	

تـــم إجـــراء التعديلات المطلوبة وأصـــبحت الرســـالة صـــالحة لمـــنح درجـــة الماجستير في القياس والتقويم التربوي والنفسي ــ قسم القياس والتقويم .

الجمهورية العربية السورية

جامعة دمشق

كلية التربية

قسم القياس والتقويم النفسي والتربوي

أثر التفاعل بين طول الاختبار وخصائص العينة على الخصائص السيكومترية ودرجات القطع لاختبار الشخصية المتعدد الأوجه مينيسوتا M.M.P.I – 2

بحث مقدم لنيل درجة الماجستير في القياس والتقويم النفسي والتربوي

إعداد الطالبة

سمية حسين سويد

إشراف الدكتورة

إيمان عز

الأستاذ المساعد في قسم القياس والتقويم النفسي والتربوي– كلية التربية– جامعة دمشق

العام الدراسي

2015-2016م

1436-1437هـ

شكر وتقدير

بعد خروج هذا البحث إلى حيز الوجود وبفضل وتوفيق من الله تعالى أتقدم بجزيل الشكر والامتنان والعرفان بالجميل إلى الأستاذة الفاضلة الدكتورة إيمان عز، التي أخذت بيدي لإنجاز هذا البحث، ومنحتني كل الاهتمام والتوجيه والمساعدة الممكنة، جزاها الله عني كل خير وتوفيق.

كما أتقدم بالشكر والامتنان للدكتور عدنان عمورة الذي منحني من وقته وفضله الكثير، في الإضاءة على القوانين الإحصائية المستخدمة في البحث، للوصول لفهمٍ أشمل ورؤيةٍ أوضح لهذه القوانين.

وأتقدم بكل الشكر والامتنان والتقدير إلى جميع أساتذتي الكرام في جامعة دمشق وإلى كل من قدم لي العون والمساعدة لإنجاز هذا البحث، وأخص بالشكر: الأستاذ الدكتور رمضان درويش، الدكتور ياسر جاموس، الدكتورة عزيزة رحمة، الدكتورة رنا قوشحة، الدكتور سالم سالم، الدكتور زكوان قريض، أطال الله بعمرهم وجزاهم عني كل خير.

كما أتقدم بجزيل الشكر لأعضاء لجنة التحكيم المحترمين، على تفضلهم بقبول مناقشة هذا البحث، وعلى الوقت والجهد الذي بذلوه في تقييمها.

ولا يفوتني أن أتوجه بالشكر والامتنان للكادر الطبي وإدارة مشفى ابن رشد للأمراض النفسية.

وكذلك لإدارة مركز الشهيد زهير حبي، وأخص بالشكر الدكتورة أمل شكو لمساعدتها لي في التطبيق على العينات المرضية، حيث بذلت معي من وقتها وجهدها وزودتني بنصائحها القيّمة، جزاها الله عني كل خير.

فهرس المحتويات

	رقم الصفحة
الفصل الأول: موضوع البحث ومنهجه:	1- 10
1- مقدمة.	2
2- مشكلة البحث.	2- 5
3- أهمية البحث.	5
4- أهداف البحث.	5- 6
5- أسئلة البحث.	6
6- منهج البحث.	7
7- مجتمع البحث وعينته.	7
8- أدوات البحث.	7
9- حدود البحث.	7- 8
10- مصطلحات البحث العلمية وتعريفاته الإجرائية.	8- 10
الفصل الثاني: بعض الدراسات السابقة:	رقم الصفحة
	11- 25
1- المقدمة.	12
2- الدراسات المحلية.	12- 13
3- الدراسات العربية.	14- 20
4- الدراسات الأجنبية.	20- 24
5- تعقيب على ما سبق من دراسات.	24- 25
الفصل الثالث: الإطار النظري:	رقم الصفحة
	26- 93
أولاً: قياس الشخصية:	29- 37
1-مفهوم الشخصية في علم النفس.	29- 32
2- نماذج من استخبارات الشخصية.	32
3- المشكلات التي تواجه اختبارات الشخصية وقياسها.	32- 34
4- رائز مينيسوتا المتعدد الأوجه للشخصية.	34- 36
5- القرار التشخيصي وأهميته في ضوء مقاييس الشخصية.	36- 37

ثانياً: العينات الإحصائية:	38- 50
1- المجتمع الإحصائي والعينة.	38
2- شروط العينة.	39
3- عشوائية المعاينة.	39- 40
4- طرق المعاينة.	40- 44
5- تقدير حجم العينة.	44- 45
6- معايير تحديد حجم العينة.	45- 46
7- استراتيجيات تحديد حجم العينة.	46- 48
8- حجم العينة وقوة الاختبار.	48- 49
9- مصادر الخطأ في العينات.	49- 50
ثالثاً: الخصائص السيكومترية للاختبار والعوامل المؤثرة فيها:	51- 72
1- مفهوم الثبات وأهميته.	51- 52
2- درجة الثبات والدرجة الحقيقية.	52- 53
3- مصادر أخطاء القياس التي تؤثر في الثبات.	53- 54
4- الانحراف المعياري والخطأ المعياري للقياس.	55- 57
5- أنواع معاملات ثبات درجات الاختبار.	57- 60
6- العلاقة بين طول الاختبار وثباته.	60- 61
7- عوامل تؤثر في معامل ثبات الاختبار.	61- 62
8- مفهوم الصدق وأهميته.	62- 63
9- أنواع معاملات صدق الاختبار.	64- 68
10- العوامل المؤثرة على صدق الاختبار.	68- 69
11- العلاقة بين الصدق والثبات.	69- 70
12- مشكلة ثبات اختبارات الشخصية.	70- 72
رابعاً: ما وراء التحليل الاحصائي وقياس الشخصية:	73- 93
1- تعريف وإضاءة على منهج ما وراء التحليل الاحصائي.	73
2- مراحل منهج ما وراء التحليل الاحصائي.	74- 75

3- بعض الاختبارات الاحصائية المستخدمة في منهج ما وراء التحليل الاحصائي والموظفة في البحث الحالي: 75- 79

- الاختبار الاحصائي v لدراسة دلالة الفرق بين عدة معاملات ارتباط بيرسون 75
- الاختبار الاحصائي M لدراسة دلالة الفرق بين عدة معاملات ألفا كرونباخ. 76- 78
- حجم التأثير (Effect size) (قوة الارتباط) 78- 79
- إحصاء Z فيشر. 79

4- القوانين الاحصائية التي يستند إليها القانون الاحصائي M: 79- 81

- مقياس كاي مربع بدرجات حرية (N – 1) 79- 80
- درجات الحرية. 80
- إحصاء فيشر الخاص بالاختبار. 81
- قانون ألفا كرونباخ. 81
- قانون التباين العام للاختبار. 81

5- بعض مشكلات ومحدودات منهج ما وراء التحليل الإحصائي. 81- 82

6- درجات القطع (Cut - off scores) 82- 93

- مسمياتها وتعريفاتها. 82
- طرق تحديد درجة القطع. 83- 91
- الأسس التقويمية لطرق تحديد درجة القطع. 92- 93

الفصل الرابع: خطّة البحث وإجراءاته: **رقم الصفحة**

94- 112

1- مقدمة. 95

2- منهج البحث. 95

3- مجتمع البحث وعينته. 95- 100

4- أدوات البحث. 101- 106

5- إجراءات التطبيق. 106- 109

6- المعالجات الإحصائية. 110- 112

الفصل الخامس: نتائج البحث وتفسيرها:	رقم الصفحة
	113-174
1- الإجابة عن السؤالين الأول والثاني.	114- 122
2- الإجابة عن السؤالين الثالث والرابع.	122- 129
3- الإجابة عن السؤالين الخامس والسادس.	129- 135
4- الإجابة عن السؤال السابع.	135- 143
5- الإجابة عن السؤال الثامن.	143- 151
6- الإجابة عن السؤال التاسع.	151- 165
7- الإجابة عن السؤال العاشر.	165- 167
8- الإجابة عن السؤال الحادي عشر.	167
9- مناقشة النتائج.	167- 174
10- مقترحات البحث.	174
ملخص البحث باللغة العربية	**رقم الصفحة**
	175- 176
المراجع	**رقم الصفحة**
	177- 182
الملاحق	**رقم الصفحة**
	183- 184
ملخص البحث باللغة الانكليزية	**رقم الصفحة**
	I - III

فهرس الجداول

رقم الجدول	عنوان الجدول	رقم الصفحة
1	طرق المعاينة الاحتمالية	41-43
2	طرق المعاينة غير الاحتمالية	43-44
3	عدد أفراد العينة من وجهه نظر الباحثين	48
4	أنواع معاملات ثبات درجات الاختبار	58- 59
5	تفسير قوة العلاقة لمعاملات الثبات	71
6	أنواع الطرق التحكيمية لتحديد درجة القطع	84- 86
7	أنواع الطرق التي تعتمد جزئياً على التحكيم وتسترشد ببيانات تجريبية لتحديد درجة القطع	87- 88
8	أنواع الطرق التي تعتمد على البيانات التجريبية وتسترشد بالتحكيم	89- 91
9	توزع أفراد عينة المضطربين نفسياً	98
10	الخصائص الديموغرافية لعينة البحث	98
11	توزع أفراد عينة طلبة السنة الرابعة حسب الكليات	99
12	توزع أفراد عينة طلبة الماجستير حسب الاختصاصات	100
13	عدد بنود المقاييس العيادية في النموذجين المُستخدمين في الدراسة	107
14	الإحصاءات الوصفية للعينات ذات الأحجام المختلفة	109
15	قيم معاملات ثبات الإعادة للمقاييس العيادية الطويلة والقصيرة لأحجام العينات المختلفة	114
16	نتائج الاختبار الإحصائي (V) لاختبار الفروق في معاملات ثبات الإعادة بين المقاييس الطويلة والقصيرة عند أحجام العينات المختلفة	115
17	قيم الاختبار (Z) لفحص الفروق في معاملات الثبات بين المقاييس الطويلة والقصيرة عند أحجام العينات المختلفة	116-117
18	نتائج الاختبار الإحصائي (V) لفحص الفروق في معاملات ثبات الإعادة بين كل عينتين في المقاييس الطويلة	118
19	نتائج الاختبار الاحصائي (V) لفحص الفروق في معاملات ثبات الإعادة بين كل عينتين في المقاييس القصيرة	119
20	نتائج الاختبار الإحصائي (Z) لفحص الفروق في معاملات ثبات الإعادة بين كل عينتين من العينات الثلاث في المقاييس الطويلة	120-121

21	قيم الاختبار الإحصائي (Z) لفحص الفروق في معاملات ارتباط ثبات الإعادة بين كل عينتين من العينات الثلاث في المقاييس القصيرة	122-121
22	قيم معاملات ثبات ألفا كرونباخ للمقاييس العيادية الطويلة والقصيرة وفق أحجام العينات المختلفة	122
23	نتائج الاختبار الإحصائي (M) لاختبار دلالة الفروق في معاملات ثبات ألفا بين المقاييس الطويلة والقصيرة عند حجم العينة (284)	123
24	نتائج الاختبار الاحصائي (M) لاختبار دلالة الفروق في معاملات ثبات ألفا بين المقاييس الطويلة والقصيرة عند حجم العينة (567)	124
25	نتائج الاختبار الاحصائي (M) لاختبار دلالة الفروق في معاملات ثبات ألفا بين المقاييس الطويلة والقصيرة عند حجم العينة (851)	125
26	نتائج الاختبار الإحصائي (M) لاختبار دلالة الفروق في معاملات ثبات ألفا بين المقاييس الطويلة والقصيرة عند حجم العينة (1134)	126
27	نتائج الاختبار الإحصائي (M) لاختبار دلالة الفروق بين معاملات ثبات ألفا للمقاييس الطويلة عند أحجام العينات المختلفة	127
28	نتائج الاختبار الإحصائي (M) لاختبار دلالة الفروق بين معاملات ثبات ألفا للمقاييس القصيرة عند أحجام العينات المختلفة	128
29	قيم معاملات الصدق المحكي للمقاييس العيادية الطويلة والقصيرة مع مثيلاتها في قائمة شطب الأعراض وفقاً لأحجام العينات المختلفة	129
30	نتائج الاختبار الاحصائي (V) لفحص الفروق بين معاملات الصدق المحكي للمقاييس الطويلة والقصيرة عند أحجام العينات المختلفة	130
31	قيم الاختبار (Z) لفحص الفروق في معاملات الارتباط المحكي للمقاييس الطويلة والقصيرة عند أحجام العينات المختلفة	131
32	نتائج الاختبار الاحصائي (V) لفحص الفروق في معاملات الصدق المحكي بين كل عينتين في المقاييس الطويلة	132
33	نتائج الاختبار الإحصائي (V) لفحص الفروق في معاملات الصدق المحكي بين كل عينتين في المقاييس القصيرة	133
34	نتائج الاختبار الإحصائي (Z) لفحص الفروق في معاملات الصدق المحكي بين كل عينتين من العينات الثلاث في المقاييس الطويلة	134

35	نتائج الاختبار الإحصائي (Z) لفحص الفروق في معاملات الصدق المحكي بين كل عينتين من العينات الثلاث في المقاييس القصيرة	135
36	نتائج تحليل التباين الأحادي لمتوسطات درجات مقياس الاكتئاب لكلٍ من عينتي الأسوياء والمضطربين نفسياً، لنوعي بنود الاختبار الطويل والقصير	136
37	نتائج المقارنات البعدية بين متوسطات درجات مقياس الاكتئاب لكلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير	137
38	نتائج تحليل التباين الأحادي لمتوسطات درجات مقياس البارانويا لكلٍ من عينتي الأسوياء والمضطربين نفسياً، لنوعي بنود الاختبار الطويل والقصير	138
39	نتائج المقارنات البعدية بين متوسطات درجات مقياس البارنويا لكلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير	139
40	نتائج تحليل التباين الأحادي لمتوسطات درجات مقياس الفصام لكلٍ من عينتي الأسوياء والمضطربين نفسياً، لنوعي بنود الاختبار الطويل والقصير	140
41	نتائج المقارنات البعدية بين متوسطات درجات مقياس الفصام لكلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير	141
42	نتائج تحليل التباين الأحادي لمتوسطات درجات مقياس الهوس الخفيف لكلٍ من عينتي الأسوياء والمضطربين نفسياً، لنوعي بنود الاختبار الطويل والقصير	142
43	نتائج المقارنات البعدية بين متوسطات درجات مقياس الهوس الخفيف لكلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير	143
44	نتائج تحليل التباين الأحادي لمتوسطات درجات مقياس الاكتئاب لكلٍ من عينتي طلبة الماجستير والطلبة الجامعيين لنوعي بنود الاختبار الطويل والقصير	144
45	نتائج المقارنات البعدية بين متوسطات درجات مقياس الاكتئاب لكلٍ من عينتي طلبة الماجستير والطلبة الجامعيين، لنوعي بنود الاختبار الطويل والقصير	145
46	نتائج تحليل التباين الأحادي لمتوسطات درجات مقياس البارانويا لكلٍ من عينتي طلبة الماجستير والطلبة الجامعيين لنوعي بنود الاختبار الطويل والقصير	146
47	نتائج المقارنات البعدية بين متوسطات درجات مقياس البارانويا لكلٍ من عينتي طلبة الماجستير والطلبة الجامعيين، لنوعي بنود الاختبار الطويل والقصير	147
48	نتائج تحليل التباين الأحادي لمتوسطات درجات مقياس الفصام لكلٍ من عينتي طلبة الماجستير والطلبة الجامعيين لنوعي بنود الاختبار الطويل والقصير	148

49	نتائج المقارنات البعدية بين متوسطات درجات مقياس الفصام لكلٍ من عينتي طلبة الماجستير والطلبة الجامعيين، لنوعي بنود الاختبار الطويل والقصير	149
50	نتائج تحليل التباين الأحادي لمتوسطات درجات مقياس الهوس الخفيف لكلٍ من عينتي طلبة الماجستير والطلبة الجامعيين لنوعي بنود الاختبار الطويل والقصير	150
51	نتائج المقارنات البعدية بين متوسطات درجات مقياس الهوس الخفيف لكلٍ من عينتي طلبة الماجستير والطلبة الجامعيين، لنوعي بنود الاختبار الطويل والقصير	151
52	درجات القطع الخام لمقياس الاكتئاب العيادي، لكلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير	152
53	درجات القطع التائية لمقياس الاكتئاب العيادي، لكلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير	152
54	درجات القطع الخام للمقياس البارانويا العيادي، لكلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير	153
55	درجات القطع التائية لمقياس البارانويا العيادي، لكلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير	153
56	درجات القطع الخام للمقياس الفصام العيادي، لكلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير	153
57	درجات القطع التائية لمقياس الفصام العيادي، لكلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير	154
58	درجات القطع الخام لمقياس الهوس العيادي، لكلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير	154
59	درجات القطع التائية لمقياس الهوس العيادي، لكلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير	154
60	الفروقات بين متوسطات درجات القطع لعينتي الأسوياء والمضطربين نفسياً، لكلٍ من المقاييس الطويلة والقصيرة	155
61	الإحصاءات الوصفية لدرجات كل من عينتي الأسوياء والمضطربين نفسياً، على المقاييس العيادية الأربعة، لنوعي بنود الاختبار الطويل والقصير	156

فهرس الاشكال والرسوم البيانية

رقم الشكل	عنوان الشكل أو الرسم البياني	رقم الصفحة
1	توزُّع قيم أفراد العينة ضمن انحرافين معياريين، عند مستوى الثقة 95%	46
2	توزع كاي تربيع	80
3	المدرج التكراري لدرجات مقياس الاكتئاب الطويل لعينة الأسوياء	157
4	المدرج التكراري لدرجات مقياس الاكتئاب القصير لعينة الأسوياء	157
5	المدرج التكراري لدرجات مقياس الاكتئاب الطويل لعينة المضطربين نفسياً	158
6	المدرج التكراري لدرجات مقياس الاكتئاب القصير لعينة المضطربين نفسياً	158
7	المدرج التكراري لدرجات مقياس البارانويا الطويل لعينة الأسوياء	159
8	المدرج التكراري لدرجات مقياس البارانويا القصير لعينة الأسوياء	159
9	المدرج التكراري لدرجات مقياس البارانويا الطويل لعينة المضطربين نفسياً	160
10	المدرج التكراري لدرجات مقياس البارانويا القصير لعينة المضطربين نفسياً	160
11	المدرج التكراري لدرجات مقياس الفصام الطويل لعينة الأسوياء	161
12	المدرج التكراري لدرجات مقياس الفصام القصير لعينة الأسوياء	161
13	المدرج التكراري لدرجات مقياس الفصام الطويل لعينة المضطربين نفسياً	162
14	المدرج التكراري لدرجات مقياس الفصام القصير لعينة المضطربين نفسياً	162
15	المدرج التكراري لدرجات مقياس الهوس الطويل لعينة الأسوياء	163
16	المدرج التكراري لدرجات مقياس الهوس القصير لعينة الأسوياء	163
17	المدرج التكراري لدرجات مقياس الهوس الطويل لعينة المضطربين نفسياً	164
18	المدرج التكراري لدرجات مقياس الهوس القصير لعينة المضطربين نفسياً	164

فهرس ملاحق البحث

رقم الملحق	عنوان الملحق	رقم الصفحة
1	**تحديد حجم العينة (من 500- >100000)**	183
2	**تحديد حجم العينة (من 100- 450)**	184

الفصل الأول

موضوع البحث ومنهجه

أولاً: المقدمة.

ثانياً: مشكلة البحث.

ثالثاً: أهمية البحث.

رابعاً: أهداف البحث.

خامساً: أسئلة البحث.

سادساً: منهج البحث.

سابعاً: مجتمع البحث وعينته.

ثامناً: أدوات البحث.

تاسعاً: حدود البحث.

عاشراً: مصطلحات البحث العلمية وتعريفاته الإجرائية.

الفصل الأول

موضوع البحث ومنهجه

أولاً: المقدمة:

يُعد مفهوما الصدق والثبات من أهم المفاهيم التي ترتكز إليها نظرية القياس النفسي، ومما يعزز أهمية هذين المفهومين أنهما يُعدان شرطين أساسيين من شروط صلاحية الاختبار، أياً كان نوعه، ومن المعلوم أن الاختبار النفسي يوضع في الأساس لتلبية غرضٍ أو مجموعةٍ من الأغراض، ليكون الأساس في اتخاذ عددٍ من القرارات المهمة، ولكي يؤدي الاختبار دوره على النحو الأمثل لا بد أن يقدّم معلومات واقعية وصحيحة حول السمة أو الخاصية موضع القياس، أي أن يكون صادقاً، وأن يكون حساساً للفروق الفردية بين الأفراد في هذه السمة أو الخاصية، أي أن يكون ثابتاً، ومن المعلوم أيضاً أن هناك عوامل عديدة تؤثر في دقة معاملات صدق وثبات أدوات القياس، بعضها يتعلق بإدارة الاختبار كالعوامل البيئية، وبعضها يتعلق باستجابات المفحوصين كالقلق والاضطراب الناتجين عن الاختبار، والبعض الثالث يتعلق بالاختبار نفسه مثل طول الاختبار وصياغة بنوده وترتيبها، والبعض الآخر يتعلق بخصائص عينة البحث التي يطبق عليها الاختبار والتي تتحدد بعدة عوامل منها: العمر والجنس والصحة النفسية ومستوى التعليم، إضافةً إلى حجم العينة، حيث من الصعب الاعتماد على نسبة معينة مثل 10% حتى نضمن تمثيل العينة للمجتمع الأصلي، وبالتالي لابد من اللجوء إلى الأساليب الإحصائية لتحديد الحد الأدنى لحجم العينة، ودراسة أثره على الخصائص السيكومترية بشكل علمي.

لذا فإنَّ البحث الحالي يسعى لتقصي أثر عاملين من العوامل المؤثرة في الخصائص السيكومترية ودرجات القطع، وهما طول الاختبار وخصائص العينة، من خلال استخدام رائز الشخصية المتعدد الأوجه مينيسوتا MMPI-2.

ثانياً: مشكلة البحث:

يُمثل القياس مظهراً بارزاً من مظاهر التقدم العلمي في المجالات المختلفة، ويُعد أمراً بالغ الأهمية في جميع العلوم دون استثناء من أجل فهم الظواهر المتعلقة بهذه العلوم، وتفسيرها والتنبؤ بالعلاقات القائمة بين متغيراتها، ومحاولة ضبطها والتحكم بها.

والقياس كما يعرفه ميخائيل (ميخائيل، 1997) هو عملية التقدير الموضوعي الكمي، وإعطاء قيم عددية للسمة أو الخاصية المقيسة، استناداً إلى قواعد محددة، تتم من خلالها المطابقة بين السمة أو الخاصية المقيسة والرقم الدال على درجة وجودها، حيث تتم هذه العملية باستخدام أدوات قياس مختلفة تتنوع وتتعدد بحسب الهدف من استخدامها، ويمكن تعريف أداة القياس على أنَّها مجموعة من البنود أو الأسئلة أو المواقف التي تمثل القدرة أو السمة أو الخاصية المطلوب

قياسها، وعلى هذا فإنه يمكن القول بأن هذه الأداة تمثل عينة من مكونات هذه القدرة أو السمة أو الخاصية، وكلما كانت هذه العينة قادرة على تمثيل المجتمع الأصلي الذي أُخذت منه، كانت هذه الأداة جيدة وصالحة ويمكن الاعتماد على نتائجها (عبد الرحمن، 1998، ص159)، وتعدّ مشكلة اختيار أداة البحث من أهم المشكلات التي تواجه الباحث في العلوم النفسية والتربوية، حيث يتوقف عليها كل قياس وكل نتيجة يخرج بها، ولا تُعد أداة القياس صالحة للاستعمال إلا إذا توافرت فيها شروط معينة هي بمثابة أهداف يحاول مُصمِّم الاختبار تحقيقها أثناء تصميمه للاختبار، أهمها: الصدق والثبات والموضوعية أي الخصائص السيكومترية (أبو علام ،2005، ص355).

ولهذه الخصائص أهمية بالغة في البحوث النفسية والتربوية على وجه الخصوص لأن القياس في هذين المجالين غير مباشر، ولذلك يجب التأكد من أن ما تقيسه أدوات البحث يمكن الثقة فيه والاعتماد عليه في جمع البيانات، كما ترتبط عملية القياس بغاياتٍ عديدة غالباً ما تتعلق بالتصنيف والتنبؤ والانتقاء والترفيع حسب نوع أداة القياس والغاية منها، والتي تتم اعتماداً على قواعد محددة تحديداً مسبقاً لها مسميات متنوعة منها درجات القطع أو درجات التمكن أو مستويات الأداء، حيث يتم اعتماد هذه الدرجات بهدف اتخاذ القرارات، وكلما زادت أهمية القرارات التي تُبنى اعتماداً عليها، كانت الحاجة أكبر إلى تحديد هذه الدرجات بعناية فائقة، ودراسة العوامل التي تؤثر فيها، بهدف الوصول إلى درجات قطع يمكن الوثوق بمصداقيتها وموضوعيتها (التميمي، 1999، ص ص 3، 4).

ومن المعلوم أن الخصائص السيكومترية لأداة القياس وكذلك درجات القطع تتأثر بعدد كبير من العوامل، منها ما يتعلق بالاختبار ذاته مثل طول الاختبار، والذي يشير إلى عدد وحدات الاختبار التي تمثل السمة أو الخاصية المقيسة، حيث يؤثر طول الاختبار في الصدق بالدرجة الأولى، فإذا كان الاختبار قصيراً وتضمن عدداً ضئيلاً من البنود يضعف تمثيله للسمة المقيسة وبالتالي يضعف صدقه (ميخائيل، 2008، ص 176)، ويؤثر أيضاً بدوره في الثبات وذلك لأنَّ العدد الأكبر من البنود يؤدي إلى الحصول على عينة أكبر من السلوك المقيس، وبالتالي يؤدي إلى رفع معامل الثبات (مجيد، 2007، ص 120)، فالاختبارات الأكثر طولاً تعطي بشكل عام درجات أكثر ثباتاً، ذلك لأن الأخطاء العشوائية الموجبة والسالبة في الاختبار الأطول ستحظى بفرصة أكبر أن يلغي بعضها بعضاً (النبهان، 2004، ص256). ومع ذلك لابد من الإشارة إلى أن الزيادة في طول الاختبار له محددات معينة، منها الوقت اللازم لتطبيق الاختبار، والجهد والتعب والملل الذي يحدث عند المفحوص نتيجة زيادة طول الاختبار، ومنها أيضاً أنه لا يمكن إضافة أسئلة بجودة الأسئلة الأصلية وكفاءتها، أي أن الزيادة في طول الاختبار تكون مفيدة لأجل الثبات عند نقطة معينة تعود بعدها لتنقص من الثبات (رحمة، 1999، ص50)، وهذا ما

يؤكده جونز هوروكز (Johns, 1965) حيث يشير إلى أنه إذا كانت الاختبارات طويلة جداً فإن عوامل إضافية لمادة البحث سوف تجهده، وتقلل التركيز في دراسته، والأنشطة غير المرغوبة للمهمة موضع البحث سوف تتدخل لجعل الأهداف الأساسية موضع شك (Horrocks,1965,P. 71).

ومن العوامل المؤثرة في الخصائص السيكومترية ودرجات القطع ما يتعلق بعينة البحث مثل حجم العينة، والذي يشير إلى عدد الأفراد الذين طُبّق عليهم الاختبار واعتُمد عليهم في استخراج مؤشرات الصدق والثبات، حيث يغلب على العديد من الدراسات والبحوث ذات الصلة بمجال القياس النفسي والعقلي استخدام عينات ذات أحجام مختلفة ولأغراض متباينة ومتعددة، ويندر أن يكون غرض تلك الدراسات دراسة حجم العينة بالذات وأثره على دقة القياس، وبشكل عام كلما كبر حجم العينة انعكس ذلك بصورة إيجابية على النتائج وأسهم في زيادة الثقة بالعينة والنتائج المتحصلة عنها، ولكن بالمقابل يجب الأخذ بعين الاعتبار أن زيادة حجم العينة يتطلب المزيد من الوقت والجهد والنفقات، مما يتعارض مع الفكرة الأساسية الموجهة لأسلوب استخدام العينة. وبطبيعة الحال يجب أن يتناسب حجم العينة مع حجم الأصل أو المجتمع الأصلي الذي ستسحب منه، ومما يجب تأكيده أيضاً أن القرار حول حجم العينة ونسبة تمثيلها للمجتمع الأصلي لا يتحدد فقط بكم المجتمع، بل يتحدد أيضاً بطبيعته وصفاته الأساسية، كما يتحدد بأغراض الدراسة ومستوى الدقة المرجوة منها (ميخائيل، 2001، ص 104). وقد تناولت دراسات عديدة أثر واحد أو أكثر من هذه العوامل في الخصائص السيكومترية للمقياس أو في درجات قطعه، حيث أشارت نتائج دراسة (السكت، 2013) إلى أن الخصائص السيكومترية لأدوات القياس تتأثر بعدد البنود وحجم العينة وهي ليست مستقلة عنها، وإذا ازداد عدد البنود وحجم العينة تزداد قيم معاملات الصدق والثبات وبفروق دالة إحصائياً عند مستوى الدلالة (0,05).

كما أكدت نتائج دراسة (الأحمدي، 2008) على وجود حجم تأثير دال لطول الاختبار وحجم العينة في دقة واتساق النتائج، وأظهرت نتائج دراسة هيلمان (Hellman, 2006) أن المقياس الأصلي الأطول يتميز باتساق داخلي عالٍ، ويحقق ثباتاً أعلى من النسخة المختصرة، وأشارت نتائج دراسة (ذيب، 2013) إلى تساوي نسب درجات قطع الاختبار الطويل والمتوسط والقصير وفق طريقتي نيدلسكاي والمجموعات المحكية حيث بلغت هذه النسبة (64.3)، كما بينت أن الاختبار الطويل حقق أعلى نسبة لدرجة القطع وفقاً لطريقة المجموعات المتضادة، حيث بلغت هذه النسبة (72.6)، في حين أظهرت نتائج دراسة (التميمي، 1999) عدم وجود علاقة دالة إحصائياً بين طول الاختبار ودرجات قطعه، أي لا يترتب على زيادة عدد المفردات تدني أو ارتفاع في نسب درجات القطع. إلا أن أياً من تلك الدراسات لم تجمع بين المتغيرات الأربعة

والمتمثلة بأثر التفاعل بين طول الاختبار وحجم العينة في الخصائص السيكومترية ودرجات القطع، لذا عمدت الباحثة في الدراسة الحالية للجمع بين هذه المتغيرات باستخدام رائز مينيسوتا المتعدد الأوجه للشخصية، والذي ساعد في استخدام هذا الرائز عوامل عديدة منها: أنه اختبار عالمي، يطبق على الأسوياء وغير الأسوياء، معير وخصائصه السيكومترية مدروسة على البيئة السورية، حيث تمّ تقنينه من قبل الدكتورة إيمان عز عام 2012 (عز، قيد النشر)، على عينة من طلبة السنة الثالثة في جامعة دمشق بلغ عدد أفرادها 4376 طالباً وطالبة، والعامل الأهم وجود النماذج البحثية في أساس الدراسات التي دارت حول هذا الاختبار، والتي قامت ببناء نماذج مقاييس عيادية من بنود الرائز نفسه ولكل جانب من الجوانب التي يقيسها، ولكن بعدد بنود مختلف، مما سمح بدراسة أثر الاختلاف في طول هذه المقاييس على كل من الخصائص السيكومترية ودرجات القطع، أهم هذه النماذج البحثية هو النموذج الذي وضعه هاريس ولنجوس والذي اعتمدته الدراسة الحالية.

وبناءً على ذلك فإن مشكلة البحث تتمحور حول الإجابة عن السؤال الأساسي الآتي:

- **ما أثر التفاعل بين طول الاختبار وخصائص العينة على الخصائص السيكومترية ودرجات القطع لاختبار الشخصية المتعدد الأوجه مينيسوتا MMPI-2؟**

ثالثاً: أهمية البحث:

تنطلق أهمية البحث الحالي من النقاط الآتية:

1- بيان العلاقة بين طول الاختبار وخصائص العينة في تأثيرهما في الخصائص السيكومترية ودرجات القطع للاختبار، مما قد يساعد معدّي الاختبارات والمقاييس في اتخاذ القرار المناسب لتحديد عدد البنود وحجم العينة الملائم لتحقيق الخصائص السيكومترية.

2- قلة الدراسات والأبحاث التي اعتمدت عاملين أحدهما يرتبط بالاختبار ذاته والآخر يرتبط بالعينة، في دراسة أثرهما في الخصائص السيكومترية من ناحية، ودرجات القطع من ناحية ثانية.

3- البحث الأول محلياً – في حدود علم الباحثة- الذي يدرس أثر التفاعل لعاملين في الخصائص السيكومترية ودرجة القطع.

رابعاً: أهداف البحث:

يتمثل الهدف الأساسي للبحث بدراسة أثر التفاعل بين طول الاختبار وخصائص العينة في الخصائص السيكومترية ودرجات القطع لاختبار الشخصية المتعدد الأوجه مينيسوتا، ويتفرع عنه الهدفان الآتيان:

1- دراسة أثر طول الاختبار في قيم معاملات صدقه وثباته، ودرجات القطع الناتجة عنه.

2- دراسة أثر خصائص العينة في قيم معاملات صدق الاختبار وثباته، ودرجات القطع الناتجة عنه.

خامساً: أسئلة البحث:

ينطلق تحقيق أهداف البحث من الأسئلة الآتية:

1- هل توجد فروق ذات دلالة إحصائية بين معاملات ثبات درجات الاختبار بالإعادة باختلاف طول الاختبار؟

2- هل توجد فروق ذات دلالة إحصائية بين معاملات ثبات درجات الاختبار بالإعادة باختلاف حجم العينة؟

3 - هل توجد فروق ذات دلالة إحصائية بين معاملات ثبات الاتساق الداخلي باختلاف طول الاختبار؟

4- هل توجد فروق ذات دلالة إحصائية بين معاملات ثبات الاتساق الداخلي باختلاف حجم العينة؟

5- هل توجد فروق ذات دلالة إحصائية بين معاملات الصدق المحكي باختلاف طول الاختبار بدلالة محك قائمة الأعراض؟

6- هل توجد فروق ذات دلالة إحصائية بين معاملات الصدق المحكي باختلاف حجم العينة بدلالة محك قائمة الأعراض؟

7- هل توجد فروق ذات دلالة إحصائية بين معاملات صدق الاختبار باستخدام طريقة المجموعات المتضادة /بين عينتي الأسوياء، والمضطربين نفسياً/ باختلاف طول الاختبار؟

8- هل توجد فروق ذات دلالة إحصائية بين متوسطات الدرجات على المقاييس العيادية استناداً لمتغير المرحلة الدراسية /مرحلة التعليم الجامعي، ومرحلة التعليم ما بعد الجامعي – الماجستير/ باختلاف طول الاختبار؟

9- هل توجد فروق ذات دلالة إحصائية بين درجات قطع المقاييس العيادية لكلٍ من عينتي الأسوياء والمضطربين نفسياً، باختلاف طول الاختبار؟

10- ما أثر التفاعل بين طول الاختبار وخصائص العينة على الخصائص السيكومترية للاختبار؟

11- ما أثر التفاعل بين طول الاختبار وخصائص العينة على درجات القطع؟

سادساً: منهج البحث:

بعد الاطلاع على بعض الدراسات والبحوث السابقة، ومراجعة المناهج البحثية، توصلت الباحثة إلى أنَّ المنهج الملائم للبحث الحالي هو المنهج الوصفي التحليلي المقارن، وهو المنهج الذي يعتمد على دراسة الظاهرة كما توجد في الواقع ويهتم بوصفها وصفاً دقيقاً ويُعبر عنها تعبيراً كيفياً وكمياً، فالتعبير الكيفي يصف لنا الظاهرة ويبين خصائصها، بينما التعبير الكمي يعطينا وصفاً رقمياً لمقدار الظاهرة أو حجمها ودرجات ارتباطها مع الظواهر المختلفة الأخرى (عبيدات، 1993، ص 219).

سابعاً: مجتمع البحث وعينته:

تألَّف مجتمع البحث من ثلاثة مجتمعات أصلية، هي مجتمع طلبة مرحلة التعليم الجامعي ومجتمع طلبة مرحلة التعليم ما بعد الجامعي من جامعة دمشق، ومجتمع المضطربين نفسياً من مشفى ابن رشد للأمراض النفسية، ومركز الشهيد زهير حبي / العيادة النفسية /.

ومن هنا تكونت عينة البحث من ثلاث مجموعات أساسية هي على النحو الآتي:

1- مجموعة طلبة مرحلة التعليم الجامعي، والتي بلغ عدد أفرادها /1134/ طالباً وطالبة.

2- مجموعة طلبة مرحلة التعليم ما بعد الجامعي/ الماجستير/، والتي بلغ عدد أفرادها /206/ طلاب وطالبات.

3- مجموعة المضطربين نفسياً المُصابين بواحدة من الاضطرابات الآتية: الفصام، الاكتئاب، توهم المرض، الانحراف السيكوباتي، والتي بلغ عدد أفرادها /150/ حالة.

ثامناً: أدوات البحث:

تم استخدام الأدوات الآتية:

1- رائز الشخصية المتعدد الأوجه مينيسوتا MMPI-2: قامت الدكتورة إيمان عز بتقنينه والتحقق من صدقه وثباته في الجمهورية العربية السورية عام 2012 (عز، قيد النشر).

2- قائمة الشطب للأعراض (SCL-90-R) : قامت الدكتورة إيمان عز بإعدادها للبيئة السورية، كما تمَّ دراسة خصائصها السيكومترية، واستخدامها كمحك لاختبار روتر لتكملة الجمل الناقصة في رسالة الدكتوراه للطالبة هيفاء البقاعي عام 2007 (البقاعي، 2007).

تاسعاً: حدود البحث:

الحدود المكانية: جامعة دمشق، مركز الشهيد زهير حبي /العيادة النفسية/، مشفى ابن رشد للأمراض النفسية.

الحدود الزمانية: تمَّ تطبيق البحث خلال العام الدراسي 2013/2012 م.

الحدود البشرية: – عينة من طلبة السنة الرابعة من جامعة دمشق.

– عينة من طلبة الماجستير من جامعة دمشق.

– عينة من المضطربين نفسياً المُصابين بواحدة من الاضطرابات الآتية:

الفصام، الاكتئاب، توهم المرض، الانحراف السيكوباتي.

عاشراً: مصطلحات البحث العلمية وتعريفاته الإجرائية:

- **اختبار مينيسوتا المتعدد الأوجه للشخصية MMPI:**

هو أحد أهم الاختبارات المُصمّمة لقياس السمات الأساسية للشخصية، أعده هاثاواي وماكنلي عام (1943) متضمناً (550) بنداً ثم أضيف له (17) بنداً مكرراً في كراسة الأسئلة، يضم عدداً من المقاييس التشخيصية أو العيادية مع أربعة مقاييس للصدق، تمّ تعييره على البيئة السورية من قبل الدكتور أحمد عنبر عام 1990 (عنبر،1990).

- **اختبار مينيسوتا المتعدد الأوجه للشخصية 2 –MMPI:**

هو النسخة الثانية لاستخبار مينيسوتا المتعدد الأوجه للشخصية، ظهرت عام 1989 كاستدراك للنقائص والعيوب التي عانت منها النسخة الأولى، وتتألف من عشرة مقاييس عيادية، تمّ تقنين هذا الرائز ودراسة خصائصه السيكومترية في الجمهورية العربية السورية من قبل الدكتورة إيمان عز عام 2012 (عز، قيد النشر).

- **قائمة الأعراض: /Symptom Checklist/90:**

تعرف اختصاراً ب (SCL–90)، وهي من أكثر القوائم استخداماً في العيادات والمشافي النفسية الأمريكية، تتألف من /90/ بنداً، يُجاب عن كل منها بأحد الخيارات التالية: "أبداً، قليلاً، أحياناً، كثيراً، دائماً " ضمن مدى من الدرجات يتراوح ما بين /0–4/، تصلح للتطبيق الفردي أو الجماعي على الأفراد الذين تجاوزوا عمر/13/ سنة، وتُصحّح إما يدوياً أو عن طريق الحاسوب، وغالباً ما يستغرق التطبيق مدة تتراوح بين /12–15/ دقيقة (البقاعي،2007، ص93).

- **طول الاختبار Test Length :**

اصطلاحاً: هو عينة من الوحدات التي تمثل القدرة أو السمة التي يقيسها الاختبار (عبد الرحمن، 1993، ص176).

إجرائياً: تُعرفه الباحثة إجرائياً بأنَّه: عدد البنود المكونة لكل مقياس من المقاييس العيادية، في كلٍ من النموذج الأصلي الذي أعدّه هاثاواي وماكنلي، والنموذج البحثي الذي أعدّه هاريس ولنجوس.

خصائص العينة Sample characteristics:

إجرائياً:

المقصود بخصائص العينة إجرائياً: تمثيل العينة لتشمل بعض المتغيرات المراد دراسة أثرها في الخصائص السيكومترية ودرجات القطع، وهي على النحو التالي:

- متغيّر المرحلة الدراسية: مرحلة التعليم الجامعي /طلبة السنة الرابعة / و مرحلة ما بعد التعليم الجامعي /طلبة الماجستير/.
- متغيّر الحالة الصحية: عينة الأسوياء وعينة المضطربين نفسياً / وتشمل حالات الفصام، الاكتئاب، توهم المرض، الانحراف السيكوباتي/.
- متغيّر حجم العينة: حيث جرى سحب واختيار أحجام العينات على أساس حسابها كنسب مئوية من عدد بنود رائز مينيسوتا MMPI-2 ككل، أي (567) بنداً، ذلك لأن رائز مينيسوتا وإن كان يتضمن مقاييس عيادية عديدة، إلا أنه اختبار كلي ويطبق غالباً ككل، وكانت أحجام العينات المدروسة على النحو الآتي: 50% و 75% و 100% و 150% و 200% من عدد بنود الاختبار.

- **حجم العينة Sample Size:**

اصطلاحاً: هو عدد وحدات المعاينة التي يجري اختبارها ويرمز له عادة بالرمز (n) (الزكري، 2005، ص10).

إجرائياً: عدد الأفراد الذين طبق عليهم الاختبار واعتمد عليهم في استخراج مؤشرات الصدق والثبات من ناحية، ومقدار تمثيلها للمجتمع الأصلي من ناحية أخرى.

وقد جرى سحب واختيار أحجام العينات في البحث الحالي على أساس حسابها كنسب مئوية من عدد البنود الكلي للرائز، وكانت أحجام العينات المدروسة على النحو الآتي: 50% و 75% و 100% و 150% و 200% من عدد بنود الاختبار.

- **الخصائص السيكومترية Psychometric Characteristics:**

اصطلاحاً: يُقصد بالخصائص السيكومترية للاختبار الخصائص المتعلقة بالصدق والثبات والقابلية للاستعمال، والتي يتم حسابها بعد تجريب الاختبار على عينة ممثلة للمجتمع. وتعتمد جودة الاختبار على مدى توافر بيانات مناسبة لهذه الخصائص (مراد، 2002، ص350).

إجرائياً: يقصد بالخصائص السيكومترية إجرائياً:

– قيم معاملات ثبات الاختبار بالإعادة، وقيم معاملات ثبات الاتساق الداخلي للمقاييس العيادية المُستخدمة في الدراسة لدى كل من نموذجي هاثاواي وماكنلي، وهاريس ولنجوس. وفق أحجام العينات المختلفة.

– قيم معاملات الصدق المحكي بدلالة محك قائمة الأعراض، وفق أحجام العينات المختلفة.

– قيم معاملات صدق المجموعات المتضادة /أسوياء ومضطربون نفسياً/.

- **درجة القطع Cut-Off Score:**

اصطلاحاً: هي نقطة على متصل درجات الاختبار، تستخدم لتصنيف الأفراد إلى فئات تعكس الأداء لمختلف الجوانب المُراد قياسها (مجيد،2007، ص186).

إجرائياً: هي الدرجة التائية التي تفصل بين السواء والاضطراب.

وتمّ حساب درجتي القطع العليا والدنيا لرائز مينيسوتا MMPI-2 في الدراسة الحالية باستخدام المعادلتين الآتيتين:

- درجة القطع العليا = متوسط الدرجات + 2 × الانحراف معياري.
- درجة القطع الدنيا = متوسط الدرجات – 2 × الانحراف معياري.

الفصل الثاني
دراساتٌ سابقة

أولاً: مقدمة.

ثانياً: الدراسات المحلية.

ثالثاً: الدراسات العربية.

رابعاً: الدراسات الأجنبية.

خامساً: تعقيب على ما سبق من دراسات.

الفصل الثاني

دراساتٌ سابقة

أولاً: مقدمة:

نظراً لعدم حصول الباحثة على دراساتٍ تربط بين متغيرات البحث الأربعة، والمتمثلة بأثر التفاعل بين طول الاختبار وخصائص العينة في الخصائص السيكومترية ودرجات القطع، سوف يتم استعراض بعض الدراسات التي دارت حول واحدٍ أو أكثر من هذه المتغيرات في الفصل الحالي، للاستفادة منها ومعرفة ما توصلت إليه من نتائج، وقد تمّ تقسيم هذه الدارسات إلى دراساتٍ محلية، ودراساتٍ عربية، وأخرى أجنبية، وسيتم عرضها مرتبةً من الأحدث إلى الأقدم.

ثانياً: الدراسات المحلية:

1. **دراسة (السكت،2013) بعنوان (أثر حجم العينة وعدد البنود على الخصائص السيكومترية لمقياس أيزنك للشخصية EPQ – R).**

أهداف الدراسة: هدفت الدراسة إلى معرفة أثر حجم العينة في الخصائص السيكومترية لمقياس أيزنك للشخصية EPQ – R، كما هدفت أيضاً إلى معرفة أثر عدد البنود في الخصائص السيكومترية لمقياس أيزنك للشخصية EPQ – R.

عينة الدراسة: تكونت من عينة من طلبة جامعة دمشق للعام الدراسي 2010-2011م.

منهج الدراسة: استخدمت الباحثة المنهج الوصفي التحليلي.

أداة الدراسة: مقياس أيزنك للشخصية، وهو من تأليف هانز أيزنك، أعدّه وترجمه للعربية الأستاذ الدكتور امطانيوس ميخائيل بصوره الثلاث: المختصرة (A)، والقصيرة (S)، والطويلة (R).

نتائج الدراسة: توصلت الدراسة إلى النتائج الآتية:

- الخصائص السيكومترية لأدوات القياس تتأثر بعدد البنود وحجم العينة وهي ليست مستقلة عنها، وإذا ازداد عدد البنود وحجم العينة تزداد قيم معاملات الصدق والثبات وبفروق دالة إحصائياً عند مستوى الدلالة (0,05).
- أعطت الصورة المختصرة (A) معاملات ثبات أدنى من الصورتين الطويلة (R) والقصيرة (S)، ولم تعط ارتباطات بينية عندما حسب الصدق عن طريق الترابطات البينية، ولم يسفر التحليل العاملي لها عن أي نتائج وقد يعزى ذلك إلى عدد البنود غير الكافي.
- أعطت الصورة القصيرة (S) معاملات ثبات وصدق جيدة وقدمت توافقاً في البنية العاملية والبنية الأساسية للمقياس من حيث أنه يقيس أربعة أبعاد وهي العصابية والانبساطية والذهانية والكذب، وهذا ما يقدم مؤشراً هاماً لصدقها العاملي، وأعطت أيضاً فروقاً ذات دلالة

إحصائية بين عيناتها الثلاث (36 – 48 – 96) وأعلى قيمة للقوة الإحصائية وحجم التأثير.

- أبدت الصورة الطويلة (R) ذات الـ 100 بند وحجم العينة (100) و(200) أعلى قيم لمعاملات ثبات ألفا كرونباخ، وكذلك فيما يخص الصدق بطريقة الارتباطات البينية، والسبب في ذلك أنه كلما زاد عدد البنود زاد تمثيل السمة المقيسة و زاد بالتالي الصدق، مما يرجح قوتها وصلاحيتها للاستخدام من حيث تقديمها لمعلومات دقيقة عن المفحوصين.

2. **دراسة (العلي، 2011) بعنوان: (تدريج اختبار مينيسوتا المتعدد الأوجه للشخصية MMPI-2 باستخدام نموذج راش).**

أهداف الدراسة: هدفت الدراسة إلى التأكد من تحقيق متطلبات القياس الموضوعي من خلال دراسة مدى فعالية نموذج راش في تدريج اختبارات الشخصية، كما هدفت إلى اختبار أثر متغيرين في نتائج تدريج اختبار مينيسوتا المتعدد الأوجه للشخصية عند استخدام نموذج راش أحادي البعد وهما (قوة الارتباط بين البند والبعد، حجم العينة) باستخدام محكات الدقة المتمثلة في الخطأ المعياري والثبات ودالة المعلومات.

عينة الدراسة: تكونت عينة الدراسة من (3000) طالب وطالبة من طلاب السنة الثالثة في جامعة دمشق.

منهج الدراسة: استخدمت الباحثة المنهج التكاملي.

أداة الدراسة: اختبار مينيسوتا المتعدد الأوجه للشخصية الصورة المراجعة، حيث تمّ تدريجه باستخدام نموذج راش، وتم استخدام البرنامج الحاسوبي بايلوج Bilog الذي اعتُمد عليه في تحليل البيانات.

نتائج الدراسة: توصلت الدراسة إلى النتائج الآتية:

- استبعاد عدد من المفردات غير الملائمة وفق مؤشر الملاءمة، علماً أن حذف المفردات لم يتم على أساس تشخيصي إنما وفق مؤشرات إحصاءات الملاءمة.
- الحصول على مقياس مدرج بوحدة اللوجيت لقياس الاضطرابات النفسية العشرة الرئيسية.
- استبعاد عدد من المفردات ذات الارتباط السالب بالبعد المكون لها.
- أفضل حجم عينة من بين أحجام العينة المستخدمة في تدريج اختبار مينيسوتا باستخدام نموذج راش وفقاً لمحكات الدقة الثلاثة هو حجم العينة (300– 700).
- وكذلك خلصت نتائج الدراسة إلى أن الثبات أعلى ما يكون عندما تكون العينة متوسطة الحجم (700) مقارنة بالعينتين (300–2000)، ما عدا مقياس الفصام حيث كان الثبات أعلى ما يكون في العينة (300) ثم (2000) ثم (700).

ثالثاً: الدراسات العربية:

1. **دراسة (ذيب، 2013) بعنوان: أثر طول الاختبار على تحديد درجة القطع لاختبار تحصيلي محي المرجع في مادة علم نفس الشخصية) في الجامعة العراقية/ كلية التربية.**

أهداف الدراسة: هدفت الدراسة إلى تعرُّف أثر طول الاختبار في تحديد درجة القطع لاختبار تحصيلي محكي المرجع وفق أربع طرق لتحديد لهذه الدرجة وهي:

- طريقة أنجوف.
- طريقة ندلسكاي.
- طريقة المجموعات المتضادة.
- طريقة المجموعات المحكية.

عينة الدراسة: تكونت عينة الدراسة من (195) طالب وطالبة، (92) ذكراً و (103) إناث، من طلبة قسم العلوم التربوية والنفسية من كلية التربية بالجامعة المستنصرية.

منهج الدراسة: استخدمت الباحثة المنهج الوصفي التحليلي المقارن، لأنها ركزت على مقارنة درجات القطع للاختبار الناتجة عن تطبيق طرق مختلفة لتحديد درجة القطع في ضوء متغير طول الاختبار.

أداة الدراسة: اختبار تحصيلي محكي المرجع في مادة علم نفس الشخصية، جميع أسئلته من نوع الاختيار من متعدد، وتشمل كل مفردة على أربعة خيارات.

نتائج الدراسة: توصلت الدراسة إلى النتائج الآتية:

– تساوي نسب درجات قطع الاختبار الطويل والمتوسط والقصير وفق طريقتي نيدلسكاي والمجموعات المحكية وبلغت هذه النسبة (64.3).

– الاختبار الطويل حقق أعلى نسبة لدرجة القطع وفقاً لطريقة المجموعات المتضادة، حيث بلغت هذه النسبة (72.6).

- تقارب نسبتي الطلبة المتقنين إلى حدٍ ما، والطلبة غير المتقنين عند تطبيق طريقة المجموعات المحكية على الاختبارات المختلفة مقارنةً بالطرق الأخرى.
- حققت طريقتا نيدلسكاي والمجموعات المحكية أدنى درجات قطع عند تطبيقهما على الاختبارات المختلفة (طويل، متوسط، قصير).

2. **دراسة (الشريم وعودة، 2010) بعنوان: (تطوير أسلوب قائم على الأهداف السلوكية لتحديد درجة القطع: دراسة مقارنة مع أسلوب أنجوف) المجلة الأردنية في العلوم التربوية.**

هدف الدراسة: هدفت الدراسة إلى تطوير أسلوب قائم على الأهداف لتقدير درجة القطع، وتقييم هذا الأسلوب بمقارنته مع أسلوب أنجوف القائم على فقرات الاختبار.

عينة الدراسة: تكونت عينة الدراسة من قسمين: عينة المحكمين وعينة الطلبة.

تكونت عينة المحكمين من (30) محكماً ومحكمة تم تقسيمهم إلى ثلاث مجموعات على النحو الآتي:

- مجموعة التأليف: مكونة من عشرة محكمين ممن شاركوا في تأليف منهاج الرياضيات في وزارة التربية والتعليم.
- مجموعة الخبراء: مكونة من عشرة محكمين ممن يحملون مؤهل دكتوراه أو ماجستير في أساليب تدريس الرياضيات.
- مجموعة المعلمين: وهم المعلمون الذين يحملون درجة البكالوريوس في الرياضيات.

أما عينة الطلبة فقد تكونت من (171) طالباً تم اختيارهم من المدارس التابعة لإحدى مديريات التربية والتعليم في الأردن.

منهج الدراسة: المنهج الوصفي التحليلي.

أداة الدراسة: تم تطوير قائمة بالأهداف السلوكية التفصيلية لوحدة تحليل المقادير الجبرية من منهاج الرياضيات، ويتبع كل هدف مجموعة من الفقرات التي تقيس هذا الهدف، وقد تم عرض قائمة الأهداف وقائمة الفقرات على المحكمين لتقدير درجة القطع لكل هدف، كما تم تطوير صورتين لاختبار محكي المرجع في الرياضيات يتكون كل منهما من ثلاثين فقرة من نوع الاختيار من متعدد تغطي القائمتين حسب درجة انسجام الفقرة بالهدف.

نتائج الدراسة: توصلت الدراسة إلى النتائج الآتية:

- درجة القطع للأهداف كانت (0.56) في جولتي التحكيم.
- كانت للصورة الأولى حسب أسلوب أنجوف (0.67) في الجولة الأولى و (0.66) في الجولة الثانية، وللصورة الثانية (0.63) في الجولتين.
- كانت جميع معاملات الارتباط بين تقديرات المحكمين في جولتي التحكيم مرتفعة ودالة إحصائياً عند مستوى الدلالة (0.01).
- لا يوجد فرق دال إحصائياً بين معاملي الثبات للنموذجين.
- النموذج القائم على الأهداف يعد مناسباً لتقدير درجة القطع، وللتنبؤ بالتحصيل المدرسي للطلبة، لغايات التشخيص في مدى تحقق الأهداف.

3. دراسة (الشريم وسوالمة، 2006) بعنوان (تحديد درجة القطع لاختبار محكي المرجع في الرياضيات باستخدام نموذجي أنجوف وندلسكاي) المجلة الأردنية في العلوم التربوية.

هدف الدراسة: هدفت هذه الدراسة إلى مقارنة نموذجي أنجوف وندلسكاي لتقدير درجة القطع لاختبار محكي في الرياضيات.

عينة الدراسة: تكونت عينة المحكمين من 80 محكماً ومحكمة، 76 منهم من معلمي مادة الرياضيات ممن يدرّسون الصف العاشر الأساسي ولا تقل الخبرة العملية لكل منهم عن سنتين،

وأربعة منهم مشرفون تربويون لمبحث الرياضيات، وتكونت عينة الطلبة من 120 طالباً من طلبة الصف العاشر.

منهج الدراسة: المنهج الوصفي التحليلي.

أداة الدراسة: اختبار محكي في مقرر معادلة الخط المستقيم للصف العاشر الأساسي، مكون من ثلاثين فقرة من نوع الاختيار من متعدد لكلٍ منها أربعة بدائل واحد منها فقط يمثل الإجابة الصحيحة.

نتائج الدراسة: توصلت الدراسة إلى النتائج الآتية:

- درجة القطع للاختبار باستخدام طريقة أنجوف تتراوح بين 0.62 و 0.68.
- درجة القطع للاختبار باستخدام طريقة ندلسكاي تتراوح بين 0.49 و 0.57.
- لا يوجد فرق دال إحصائياً عند مستوى الدلالة (0.05) بين معاملي ثبات النموذجين.
- لا يوجد فرق دال إحصائياً بين معاملي الارتباط لصعوبة الفقرات وتقديرات المحكمين لها في النموذجين عندما لا يُزوَّد المحكمون بمعاملات صعوبة الفقرات، بينما يوجد فرق دال إحصائياً بين معاملي الارتباط عند معرفة المحكمين لصعوبة الفقرات، أي أن تقديرات المحكمين لدرجة القطع تتأثر في حال وجود مؤشرات عن معاملات الصعوبة للفقرات.
- يعد نموذج أنجوف هو الأسهل من حيث الاستخدام وبساطة الإجراءات.
- يعد نموذجا أنجوف وندلسكاي الأكثر ملاءمةً للتنبؤ بالتحصيل المدرسي.

4. **دراسة (الأحمدي، 2008) بعنوان: أثر اختلاف طول الاختبار وحجم العينة على دقة واتساق تقدير بارامترات المفردة والفرد في نظرية الاستجابة المفردة) في المملكة العربية السعودية.**

أهداف الدراسة: هدفت الدراسة إلى المقارنة بين دقة واتساق بارامترات المفردة والفرد في نماذج نظرية الاستجابة للمفردة IRT اللوغارتمي الأحادي البارامتر، واللوغارتمي الأحادي البارمتر المعدل، واللوغارتمي الثلاثي البارامتر المعدل، واللوغارتمي الثلاثي غير البارامتري، مع العينات الصغيرة والاختبارات القصيرة، وذلك للكشف عن طرائق تطوير الاختبارات القصيرة ذات المفردات من نوع الاختيار من متعدد وتطوير الدراسات والبحوث القائمة على العينات الصغيرة بتحديد الأداء لكل نموذج من النماذج الأربعة، من حيث دقته واتساقه في تقدير بارامترات المفردة والفرد فيها.

أداة الدراسة: اختبارات تحصيلية من نوع الاختيار من متعدد.

وقد قامت الباحثة بقياس دقة واتساق تقدير بارامترات المفردة باستخدام الإحصاءات الآتية: معامل الارتباط، مؤشر التحيز، مؤشر الكفاءة، ثم تم فحص التأثيرات الرئيسية لنموذج طول الاختبار وحجم العينة ولتفاعله في التحيز والكفاءة لتقدير البارامترات باستخدام تحليل التباين وفق النموذج

العاملي (2،3،4) لفحص التحيز والكفاءة، وجرى استعمال قياسات حجم التأثير size effect ، بعدّها بديلاً لاختبارات الدلالة.

نتائج الدراسة: توصلت الدراسة إلى النتائج الآتية:

- وجود حجم تأثير دال لطول الاختبار وحجم العينة على دقة واتساق النتائج.
- كان حجم هذا التأثير كبيراً بالنسبة لطول الاختبار، في حين لم يكن كبيراً بالنسبة لحجم العينة.
- لم يكن هناك حجم تأثير كبير لتفاعل (النموذج × طول الاختبار × حجم العينة) في دقة تقدير بارامتر تمييز المفردة واتساقها.
- أشارت نتائج دقة تقدير بارامتر صعوبة المفردة واتساقها إلى أن النموذج اللوغارتمي الأحادي البارامتر له قيم معاملات ارتباط أعلى قليلاً من النموذج اللوغارتمي الثلاثي البارامتر المعدل، يليه النموذج اللوغارتمي الأحادي البارامتر المعدل، وأخيراً النموذج الثلاثي غير البارامتري عند كل ظرف من ظروف الاختبار، بينما النموذج اللوغارتمي الأحادي البارامتر المعدل كان له أقل تحيز.
- كما أن كلاً من النماذج اللوغارتمي الأحادي البارامتر المعدل واللوغارتمي الثلاثي البارامتر المعدل تميل على نحو طفيف لبخس التقدير، في حين أن النموذج اللوغارتمي الأحادي البارامتر يميل للمغالاة على نحو طفيف في التقدير، والنماذج الأربعة تتفاوت فيما بينها على نحو طفيف في الكفاءة.
- ولم يكن هناك حجم تأثير كبير لطول الاختبار في الدقة والاتساق، وكان هناك حجم تأثير كبير لحجم العينة على الدقة والاتساق، وكان هناك حجم تأثير كبير لتفاعل (حجم العينة × طول الاختبار × النموذج) في الدقة.
- ولم يكن هناك حجم تأثير كبير لتفاعل (النموذج × حجم العينة × طول الاختبار) في اتساق تقدير بارامتر صعوبة المفردة.

5. دراسة (عوض،2004) بعنوان: (دراسة أثر عدد المفردات وموقع المفردة من الاختبار على قيمة معامل ثبات المقياس) في مصر.

أهداف الدراسة: هدفت هذه الدراسة إلى الكشف عن أثر مجموعة من العوامل والمتغيرات في القيمة العددية لمعامل الثبات وهذه العوامل والمتغيرات هي:

1. عدد بنود الاختبار.
2. موقع البند من الاختبار

عينة الدراسة: تكونت عينة الدراسة الأساسية من (110) طلاب وطالبات من طلاب الدبلوم الخاص بالتربية السنة التكميلية، أما عينة الدراسة الأولية فتكونت من (156) طالباً وطالبة من طلاب الدبلوم العام في التربية (نظام العاملين).

أداة الدراسة: قام الباحث بإعداد أربعة اختبارات تحصيلية من نمط الاختيار من متعدد في مادة القياس والتقويم.

نتائج الدراسة: توصلت الدراسة إلى النتائج الآتية:

وجود فروق ذات دلالة إحصائية عند مستوى الدلالة (0،05)، في قيمة معامل الثبات المحسوبة لكل من هاتين الصورتين لصالح الصورة الأطول والترتيب التصاعدي.

6.دراسة (التميمي،1999) بعنوان: (أثر كل من نوع المحكم وطول الاختبار على تحديد درجة القطع لاختبار محكي المرجع يقيس الكفايات الرياضية في العمليات الحسابية على الأعداد بالصف السادس الابتدائي بمدينة جدة) في المملكة العربية السعودية.

هدف الدراسة: هدفت الدراسة لمعرفة أثر متغيري الدراسة (نوع المحكم – طول الاختبار) في عملية تحديد درجة القطع وفق الطرائق المستخدمة في الدراسة وهي نيدلسكاي – أنجوف – المجموعات المتضادة – المجموعات المحكية – عند تطبيقها على اختبار موحد وعينة دراسية واحدة.

عينة الدراسة: تكونت عينة الدراسة من طلاب السادس الابتدائي بمدارس محافظة جدة الحكومية في الفصل الثاني للعام 1416 – 1417 هجري.

أداة الدراسة: اختبار محكي المرجع يقيس الكفايات الرياضية المختلفة في العمليات الحسابية على الأعداد بالمرحلة الابتدائية.

نتائج الدراسة: توصلت الدراسة إلى النتائج الآتية:

- توجد فروق ذات دلالة إحصائية عند مستوى الدلالة(0.05) بين تقديرات المحكمين لدرجات قطع الاختبار وفق طريقة نيدلسكاي باختلاف نوع المحكم .
- لا توجد فروق ذات دلالة إحصائية بين تقديرات المحكمين لدرجات قطع الاختبار وفق طريقة أنجوف باختلاف نوع المحكم .
- حققت تقديرات مجموعتي المعلمين والمشرفين أعلى نسبة لدرجة قطع الاختبار وفق طريقة أنجوف، بينما حققتا أدنى نسبة وفق طريقة نيدلسكاي.
- تساوت تقديرات مجموعتي المعلمين والمشرفين لدرجات قطع الاختبار وفق طريقة نيدلسكاي.
- تكررت نسبة درجة القطع (57.1) بصورة كبيرة بين النسب المحددة لدرجة قطع الاختبارات الثلاثة الناتجة عن تطبيق الطرق المختلفة باستثناء طريقة أنجوف.

- لا توجد علاقة بين طول الاختبار ودرجات قطع الاختبار. أي لا يترتب على زيادة عدد المفردات تدنٍّ أو ارتفاع في نسب درجات القطع، باستثناء طريقة المجموعات المتضادة.
- يقل مجموع خطأي التصنيف الأول والثاني عند زيادة طول الاختبار وفق طرق تحديد درجات القطع المختلفة .
- تساوي نسبة درجة قطع الاختبارات الثلاثة /57.1 / الناتجة عن تطبيق طريقة نيدلسكاي والمجموعات المحكية .
- تقارب نسبتي الطلاب المتقنين وغير المتقنين في الاختبارات الثلاثة الناتجة عن تطبيق طريقة المجموعات المحكية.
- حققت طريقتا المجموعات المحكية و نيدلسكاي أدنى نسبة لدرجة قطع الاختبار وذلك في نتائج الاختبارات الثلاثة.

7. دراسة (علام، 1991) بعنوان: (مقارنة بعض طرق تحديد مستويات الأداء في اختبار محكي المرجع).

هدف الدراسة: هدفت الدراسة إلى الموازنة بين أربع طرق لتحديد مستويات الأداء في الاختبار الذي أعدّه الباحث، اثنتان منهما من الطرق المطلقة التي تعتمد على التحكيم، وهما طريقتا أنجوف وندلسكاي، والطريقتان الأخريان تعدان من الطرق النسبية أي التي تعتمد على تحكيم الأداء وجمع بيانات فعلية تتعلق بهذا الأداء وهما طريقتا المجموعة الحدية، والمجموعة المتناقضة.

عينة الدراسة: تكونت عينة المحكمين من (26) محكماً ومحكمة من ثلاث مجموعات متباينة على النحو الآتي:

- المجموعة الأولى: اشتملت على 14 معلماً من معلمي المدارس الثانوية الحاصلين على دبلوم في التربية.
- المجموعة الثانية: اشتملت على ستة مدرسين مساعدين حاصلين على درجة الدكتوراه في علم النفس ولديهم خبرة لا تقل عن ثلاثة أعوام في التدريس.
- المجموعة الثالثة: اشتملت على ستة من أساتذة علم النفس التعليمي.

أما بالنسبة لعينة الطلاب فقد بلغت 382 طالباً من طلاب كلية التربية ممن يدرسون مقرراً سيكولوجياً في الفروق الفردية.

منهج الدراسة: المنهج الوصفي التحليلي.

أداة الدراسة: اختبار محكي المرجع يقيس الكفايات الأساسية في مقرر سيكولوجي في الفروق الفردية لطلبة التربية.

نتائج الدراسة: توصلت الدراسة إلى النتائج الآتية:

- أدى استخدام الطرائق الأربع المختلفة في الدراسة إلى الحصول على درجات قطع مختلفة نسبياً.
- كانت متوسطات درجات القطع 40، 44، 43، 41، لكل من طريقة ندلسكاي والمجموعة الحدية وأنجوف والمجموعات المتناقضة على التوالي، أي أن أعلى المتوسطات لدرجات القطع التي تم الحصول عليها باستخدام الطرق الأربع كان لطريقة المجموعة الحدية التي تعد من الطرق النسبية، في حين أن أدنى هذه المتوسطات كان لطريقة ندلسكاي التي تعد من الطرق المطلقة أي التي تعتمد فقط على التحكيم دون الاسترشاد ببيانات تجريبية.

رابعاً: الدراسات الأجنبية:

1. دراسة سكوت (Scott, 2009) بعنوان: (دراسة محاكاة استخدمت حجم العينة لمعرفة وظائف البند المختلفة باستخدام مقاييس قصيرة) في اسكتلندا.

"A simulation study provided sample size guidance for differential item functioning (DIF) studies using short scale"

هدف الدراسة: هدفت الدراسة إلى محاكاة البيانات التي استخرجت من تطبيق مقياس (HRQOL)(مقياس نوعية صحة الحياة)، ويتألف هذا المقياس من أربعة بدائل إجابة، واستخدمت هذه البيانات لدراسة تأثير كل من حجم العينة وطول المقياس في القوة الإحصائية ونسبة الخطأ من النوع الأول.

وقد استخدم العلماء في دراستهم تحليلات (DIF) لفحص نزاهة البنود أو دقتها لتقويم الأداة التي صممت لمقياس نوعية صحة الحياة (HRQOL).

أداة الدراسة: استخدم الباحث مقياس(HRQOL) (مقياس نوعية صحة الحياة).

نتائج الدراسة: توصلت الدراسة إلى النتائج الآتية:

– عندما لا تستخدم تحليلات (DIF) تكون نسبة الخطأ من النوع الأول أقرب إلى (5%)، وقد لوحظ ذلك في المقياس (HRQOL) المؤلف من بديلي الإجابة.

– كان حجم العينة الأنسب /300/ ليلائم قوة إحصائية أكبر من 80%، وبالتالي كانت القوة عالية، أما المقياس الطويل فكان حجم العينة /200/ كان ملائماً وكافياً.

2. دراسة لوزانو (Lozano, 2008) بعنوان: (تأثير عدد بدائل الإجابة وحجم العينة على ثبات وصدق مقاييس التصنيف) في اسبانيا

"Effect of the number of response and sample size on the reliability and validity of rating scales"

هدف الدراسة: هدفت الدراسة إلى تحديد عدد بدائل الإجابة الأمثل التي تحسّن كلاً من الصدق والثبات.

منهج الدراسة: استخدم الباحثون المنهج الوصفي.

أداة الدراسة: مقاييس التصنيف المُعتمدة، حيث استخدم الباحثون البرنامج الحاسوبي (Prelis) و (طريقة مونتي كارلو)، وطريقة (جورسكو وسور بوم 1993) لاستخراج النتائج.

نتائج الدراسة: توصلت الدراسة إلى النتائج الآتية:

– كلما ازداد عدد بدائل الإجابة تحسن كل من الصدق والثبات.

– العدد الأمثل للبدائل بين (4) و(7)، وعندما يكون عدد بدائل الإجابة أقل من (4) بدائل ينقص كل من الصدق والثبات، وعندما يكون عدد البدائل أكثر من (7) تزداد صعوبة تفسير الخصائص السيكومترية للمقياس.

– أفضل نتائج الثبات تحققت عند حجم العينة (50)، وبدأت بالتناقص عند حجم العينة (200)، و(300)، و(500). وعند عدد بدائل الإجابة (3).

3. دراسة هيلمان (Hellman, 2006) بعنوان: (دراسة لتأثيرات متوسط العمر وعدد البنود في نتائج الثبات) في أوكلاهوما.

" A reliability generalization study on the survey of perceived organizational support: The effects of mean age and number of items on score reliability"

هدف الدراسة: هدفت الدراسة إلى توظيف تقنيات الثبات المعمم (generalization reliability – RG)، كما هدفت إلى التحري عن العوامل المؤثرة في معاملات الثبات عند تطبيق مقياس (SPOS)، وهذه العوامل هي: عدد البنود، وحجم العينة، ونسبة الإناث، ونسبة الأقلية، ومتوسط عمر العينة القيمة المحتملة التي من الممكن أن تحسب لدراسة تغير درجات الثبات.

أداة الدراسة: استخدم الباحث مقياس (SPOS) يتألف من (36) بنداً وهو أداة قياس بطريقة التقرير الذاتي والإجابة عنه مؤلفة من (7) بدائل إجابة بطريقة ليكرت، والنسخة المختصرة مؤلفة من (16) بنداً، والقصيرة مؤلفة من (8) بنود.

نتائج الدراسة: توصلت الدراسة إلى النتائج الآتية:

- هنالك ثلاثة متغيرات ذات ارتباطات هامة إحصائياً، هذه المتغيرات هي عدد البنود المستخدمة، ومتوسط عمر العينة، وحجم العينة في كل الدراسات التي استخدمت أشكالاً مختلفة من مقياس (SPOS).
- استخدام النسخة المختصرة من مقياس (SPOS) ذي البعد الواحد بعدد بنود أقل يكون ملائماً لأسباب اقتصادية، أما مقياس (SPOS) الأصلي الأطول ذو البعد الواحد فيتميز باتساق داخلي عالٍ أكثر، ويحقق ثباتاً أعلى من النسخة المختصرة التي ربما ينقص

استخدامها معامل الثبات، خصوصاً إذا كان معدل الارتباطات الداخلية بين البنود المختارة ضعيفاً، لذلك لابد من حساب ألفا كوظيفة أساسية لمعرفة معدل ارتباط البنود بالموضوع المقيس.

- عندما أُدخل عدد البنود مع حجم العينة إلى الأدوات، أنتجت علاقة ارتباطية ذات أهمية إحصائية.

4. دراسة (Woher, 1991) بعنوان: (مقارنة تجريبية لطرق تحديد درجة القطع المتصلة بالمحتوى والخاصة بتحديد صدق المحك المتصل)

" **An empirical comparison of cut off score methods for content related and criterion related validity settings).**

هدف الدراسة: هدفت الدراسة إلى تقييم مقارنة للطرائق المختلفة لتحديد درجات القطع المتصلة بالمحتوى والخاصة بتحديد صدق المحك المتصل.

أداة الدراسة: صمم الباحث اختبار لقياس الكفاية في مقرر علم النفس العام يحتوي على (30) بنداً من الخيارات المتعددة.

عينة الدراسة: تكونت عينة الدراسة من (372) طالباً من طلاب قسم علم النفس، قسمت العينة إلى مجموعتين:

المجموعة الأولى: عبارة عن (121) طالباً، استخدمت هذه المجموعة كعينة لصدق أو لنمو طرق تحديد درجة القطع المبنية على ما يلي: الانحدار وأساس التقدير والمجموعات المتضادة.

المجموعة الثانية: عبارة عن (251) طالباً استخدمت هذه المجموعة كعينة للمقارنة الفعلية.

نتائج الدراسة: توصلت الدراسة إلى النتائج الآتية:

- هنالك اختلافات بين طرائق تحديد درجات القطع وفق الطرق المختلفة.
- وكان متوسط مستويات أداء المحك للأفراد الذين اجتازوا كل درجات القطع متشابهةً جداً لذلك يجب أن يكون اختيار طريقة تحديد درجة القطع متسقاً مع استخدام الاختبار، ويكون تفسيره معروفاً بوضوح.
- عند اختيار الطريقة يجب أن تكون مصممة بشكل جيد.

5. دراسة بوني (Bunny- Pozehl ,1990) بعنوان: (تطبيق نماذج الاستجابة للمفردة في القياس مرجعي المحك: دراسة آثار اختيار النموذج وأحجام العينة وطول الاختبار على الثبات ودقة التقدير)

"Application of items response theory to criterion – referenced measurement: An investigation of the effects of model choice"

هدف الدراسة: هدفت الدراسة إلى بحث آثار اختيار النموذج وأحجام العينة وطول الاختبار على الثبات ودقة التقدير.

أداة الدراسة: – اختبار الصحة العقلية.

– تم تطبيق خمس طرق لانتقاء البند، وهي النموذج أحادي وثنائي وثلاثي البارامتر، إضافة إلى النموذج أحادي البارامتر المعدل والنموذج ثنائي البارامتر المعدل.

نتائج الدراسة: توصلت الدراسة إلى النتائج الآتية:

بينت النتائج أن الثبات قد ارتفع في الحالات التي يكون عندها طول الاختبار /100/ بند عنه في حالة طول الاختبار /150/ بنداً ، كما أشارت النتائج إلى أن النموذجين ثنائي وثلاثي البارامتر أفضل ملاءمة للبيانات المعطاة في حالة تمام العينة بالكامل مع ثلاثة أطوال مختلفة للاختبار، كما بينت أن أكثر النماذج ملاءمة، بل الأفضل في حالة العينات كبيرة الحجم هي النموذج ثلاثي البارامتر، بينما كان النموذج أحادي البارامتر وكذلك أحادي البارامتر المعدل الأفضل وبدرجة متساوية تقريباً في حالة العينات صغيرة الحجم.

6.دراسة هالبن وسيمون (Halpin and Sigmon, 1983) بعنوان (تطبيقات لتحديد صدق درجات القطع باستخدام ثلاث مجموعات مختلفة من المحكمين وفق ثلاث طرق تحكيمية).

" minimum competency standard set by three divergent groups of raters using three judgmental procedures: implications for validity"

أهداف الدراسة: هدفت هذه الدراسة إلى الإجابة عن التساؤلات التالية:

أ- هل ينتج عن استخدام كل من الطرائق التالية إيبل و نيدلسكي و أنجوف درجات قطع مختلفة في حال استخدام اختبار مقنن وموحد؟

ب- هل تؤثر كفاية المحكمين المختلفة في حال استخدامهم طرق موحدة في معايير الاختبار، أي هل ينتج عن اختلاف كفاياتهم درجات قطع مختلفة وذلك لاختبار موحد؟

ت- هل هناك تفاعل بين مجموعة المحكمين أي هل تختلف درجات القطع المحددة بطرق مختلفة باختلاف مجموعات المحكمين؟

ث- ما مدى صدق التطبيقات لنتائج هذه الدراسة؟

عينة الدراسة: تكونت عينة الدراسة من خمسة عشر محكماً، وزعوا إلى ثلاث مجموعات كل مجموعة تحتوي خمسة أعضاء.

تشتمل المجموعة الأولى على خمسة من طلاب درجة الدكتوراه في تعليم اللغة الانكليزية، وتشتمل المجموعة الثانية على خمسة من معلمي اللغة الإنكليزية في مدارس المرحلة الثانوية أربعة منهم حاصلون على درجة الماجستير في تعليم اللغة الانكليزية و الخامس منهم حاصل على درجة البكالوريوس، وتشتمل المجموعة الثالثة على خمسة أعضاء من الكلية لديهم خبرة وتدريبات في تعليم اللغة الإنكليزية، وبعضهم حاصل على درجة الدكتوراه في الفلسفة والبعض الآخر حاصل على درجة دكتوراه في التربية.

وكانت عينة الطلاب في هذه الدراسة مؤلفة من /173/ طالباً من طلاب جامعة ميسوري.

أداة الدراسة : طبق الباحثان اختبار جامعة ميسوري للغة الإنجليزية ويشتمل على /90/ سؤالاً من الاختيارات المتعددة، ويهدف إلى تقييم الكفاية في إملاء الكلمات والترقيم وتكوين الجمل وبناء القواعد وبناء الفقرات.

نتائج الدراسة: توصلت الدراسة إلى النتائج الآتية:

أ- فروقاً في درجات القطع الناتجة بالطرق المختلفة، وكانت متوسطات درجات القطع الناتجة كما يلي: طريقة ايبل (62,61) طريقة نيدلسكاي (42,63) طريقة أنجوف (56,02).

ب-عندما تضم تلك الطرق المستخدمة فإن تحديد درجات القطع لا يختلف برغم أن مجموعة المحكمين يمثلون مستويات مختلفة من الكفاية.

ج-عند عمليات المقارنة بين المجموعات الأخرى من المحكمين في هذه الدراسة فإن عينة مدرسي الثانوية العامة حددت أدنى درجات باستخدام طريقة نيدلسكاي، وحددت أيضاً أعلى درجات عند استخدامها لطريقة أنجوف.

خامساً: تعقيب على ما سبق من دراسات:

معظم الدراسات السابقة التي عادت إليها الباحثة تناولت متغيراً واحداً أو متغيرين من متغيرات البحث، ولم يتم الحصول على أية دراسة تفسر العلاقة بين متغيرات البحث الأربعة والمتمثلة بطول الاختبار وخصائص العينة والخصائص السيكومترية ودرجات القطع، بعض هذه الدراسات تناول أثر بعض العوامل على الخصائص السيكومترية للاختبار، مثل دراسة السكت (2013) التي تناولت أثر عدد البنود و حجم العينة في الخصائص السيكومترية، ودراسة الأحمدي (2008) التي تناولت أثر طول الاختبار وحجم العينة في دقة واتساق تقدير بارامترات المفردة، ودراسة عوض (2004) التي تناولت أثر عدد المفردات في قيمة معامل ثبات المقياس، ودراسة لوزانو (2008) التي تناولت تأثير عدد بدائل الإجابة وحجم العينة في ثبات وصدق مقاييس

التصنيف، ودراسة هيلمان(2006) التي تناولت أثر عدد البنود في نتائج الثبات، ودراسة بوني (1990) التي تناولت أثر حجم العينة وطول الاختبار في الثبات ودقة التقدير.

وبعضها الآخر تناول أثر بعض العوامل في درجات قطع الاختبار مثل دراسة ذيب (2013) التي تناولت أثر طول الاختبار في تحديد درجة القطع، ودراسة التميمي (1990) التي تناولت أثر كل من نوع المحكم وطول الاختبار في تحديد درجة القطع، في حين جمعت الدراسة الحالية بين المتغيرات الأربعة، بهدف الوصول إلى معرفة أثر كل من طول الاختبار وخصائص العينة في الخصائص السيكومترية ودرجات القطع، باستخدام رائز الشخصية المتعدد الأوجه مينيسوتا MMPI-2.

الفصل الثالث: الجانب النظري

الإحصاء والخصائص السيكومترية

أولاً: قياس الشخصية

1. مفهوم الشخصية في علم النفس.
2. نماذج من استخبارات الشخصية.
3. المشكلات التي تواجه اختبارات الشخصية وقياسها.
4. رائز مينيسوتا المتعدد الأوجه للشخصية.
5. القرار التشخيصي وأهميته في ضوء مقاييس الشخصية.

ثانياً: العينات الإحصائية:

1. المجتمع الإحصائي والعينة.
2. شروط العينة.
3. عشوائية المعاينة.
4. طرق المعاينة.
5. تقدير حجم العينة.
6. معايير تحديد حجم العينة.
7. استراتيجيات تحديد حجم العينة.
8. حجم العينة وقوة الاختبار.

9. مصادر الخطأ في العينات.

ثالثاً: الخصائص السيكومترية للاختبار والعوامل المؤثرة فيها:

1. مفهوم الثبات وأهميته.
2. درجة الثبات والدرجة الحقيقية.
3. مصادر أخطاء القياس التي تؤثر في الثبات.
4. الانحراف المعياري والخطأ المعياري للقياس.
5. أنواع معاملات ثبات درجات الاختبار.
6. العلاقة بين طول الاختبار وثباته.
7. عوامل أخرى تؤثر في معامل ثبات الاختبار.
8. مفهوم الصدق وأهميته.
9. أنواع معاملات صدق الاختبار.
10. العوامل المؤثرة في صدق الاختبار.
11. العلاقة بين الصدق والثبات.
12. مشكلة ثبات اختبارات الشخصية.

رابعاً: منهج ما وراء التحليل الإحصائي و قياس الشخصية:

1. تعريف وإضاءة على منهج ما وراء التحليل الإحصائي.
2. مراحل منهج ما وراء التحليل الإحصائي.
3. بعض الاختبارات الإحصائية المستخدمة في منهج ما وراء

التحليل الإحصائي والموظفة في البحث الحالي:

- **الاختبار الإحصائي V لدراسة دلالة الفرق بين عدة معاملات ارتباط بيرسون.**
- **الاختبار الإحصائي M لدراسة دلالة الفرق بين عدة معاملات ألفا كرونباخ.**
- **حجم التأثير/ قوة الارتباط (Effect size).**
- **إحصاء Z فيشر.**

4. القوانين الإحصائية التي يستند إليها القانون الإحصائي M:

- **مقياس كاي مربع بدرجات حرية (N - 1).**
- **درجات الحرية.**
- **إحصاء فيشر الخاص بالاختبار.**
- **قانون ألفا كرونباخ.**
- **قانون التباين العام للاختبار.**

5. بعض مشكلات ومحدودات منهج ما وراء التحليل الإحصائي.

6. درجات القطع (Cut - off scores)

- **مسمياتها وتعريفاتها.**
- **طرق تحديدها وميزاتها وعيوبها.**
- **الأسس التقويمية لطرق تحديدها.**

الفصل الثالث

الجانب النظري

يقف الفصل الحالي عند مفهوم الشخصية في علم النفس ونماذج من اختبارات الشخصية، ويتناول موضوع العينات الاحصائية، وطرق المعاينة واستراتيجيات تحديد حجم العينة، وكذلك الخصائص السيكومترية للاختبار والعوامل المؤثرة فيها، كما يتضمن إضاءة على منهج ما وراء التحليل الإحصائي، ومراحله، وبعض قوانينه الأساسية المستخدمة في الدراسة الحالية، وأخيراً تعريف درجة القطع ومسمياتها وطرق تحديدها.

أولاً: قياس الشخصية:

لابد في البداية من الحديث عن مفهوم الشخصية في علم النفس، وبعض النماذج من اختبارات الشخصية، وعرض بعض المشكلات التي تواجه هذه اختبارات الشخصية، وتسليط الضوء على رائز مينيسوتا المتعدد الأوجه للشخصية MMPI-2 والنماذج البحثية التي أعدت حوله والمقاييس العيادية التي أُجريَت عليها الدراسة الحالية، وكذلك الحديث عن القرار التشخيصي وأهميته في قياس الشخصية، وبناءً على ذلك سوف يتناول هذا المحور موضوع الشخصية.

1- مفهوم الشخصية في علم النفس:

لدراسة الشخصية أهميّة ومكانة كبرى بين عدد من العلوم التطبيقية والإنسانية التي ترتبط على نحوٍ مباشر أو غير مباشر بالإنسان كفرد وكعضو في المجتمع في آنٍ واحد، والقياس في ميدان الشخصية أمر جوهري وأساسي لما يحققه من تقدم وتطور في دراسة هذا الموضع مؤثراً ومتأثراً بالشخصية، يستند إليها كما تستند إليه (عز،1990، ص4). ونظراً لما لها في علم النفس من مكانة اقترح بعض علماء النفس أن يطلق عليها "علم الشخصية " Person - ology إشارةً إلى إمكان قيامها تخصصاً قائماً بذاته. (عبد الخالق، 1994، ص4)

يدرس علم النفس الشخصية من ناحية مكوناتها وأبعادها الأساسية ونموها وتطورها ومحدداتها الوراثية والبيئية وطرق قياسها واضطراباتها، كل ذلك على أساس نظريات متعددة، كثيراً ما تكون متباينة، وإن كان الهدف بينها مشتركاً، وهو التنبؤ بالسلوك الإنساني في مختلف المواقف والأوقات (الأنصاري، 2000، ص 29).

حيث احتلت الشخصية منذ نهايات القرن التاسع عشر مكانة هامة في الدراسات النفسية، ويصدق هذا القول في حال دراسة الشخصية السوية كما يصدق في حال دراسة الشخصية

المضطربة، وقد ساعد على تأكيد هذه المكانة عوامل عديدة من بينها النظر إلى السلوك على أنه محصّلة لشخصية تعمل من حيث هي وحدة متكاملة، وفيها ما تنطوي عليه من عناصر ومركبات ودوافع وقدرات، إلا أن هذا الاهتمام العظيم بالشخصية لا يسلم من الاختلاف في المنحى الذي تأخذه الأبحاث والدراسات المختلفة حولها (ملحم، 2001، ص 302).

ويستند قياس الشخصية إلى عدة افتراضات أهمها:

- أن الأفراد يختلفون فيما بينهم، وأن دراسة الشخصية تتطلب إمكانية تقدير وقياس هذه الفروق.
- أن معظم السمات السلوكية ثابتة للقياس وتتدرج في استمرارية تخضع له.
- الشخصية تتميز بشيء من الثبات ولكنه ليس ثباتا مطلقاً.

وليس غريباً أن يرى البعض أن القياس النفسي هو قياس الشخصية بكل جوانبها، وهم في هذا يستندون إلى فهمهم أو نظريتهم للشخصية باعتبارها تتضمن كل الجوانب العقلية أي المعرفية، وكذلك الجوانب غير العقلية التي تتضمن الجوانب المزاجية والاجتماعية أيضاً.

وهناك ميل واضح لاستخدام لفظة مقياس أو استفتاء أو استخبار أو قائمة أو سجل في قياس الشخصية بدلاً من لفظة اختبار، حيث توحي التسمية الأخيرة بوجود إجابات صحيحة وأخرى خاطئة، كما يحدث في الاختبارات المعرفية أو العقلية كاختبارات الذكاء والقدرات، بينما يطلب من المُختَبر في قياس الشخصية أو أحد مكوناتها أن يجيب عن الأسئلة أو العبارات أو القضايا أو المواقف الموجودة أمامه بما يشعر به، وبما يفضله وبما يكرهه وأساليب سلوكه واتجاهاته، وعليه لا توجد إجابات صحيحة وأخرى خاطئة بل كل إجابة لها قيمتها ودلالتها (مجيد، 2007، ص ص 347 – 348).

وكلمة شخصية هي إحدى الكلمات الشائعة والدارجة على ألسنة الناس في بلدان العالم المختلفة وتشير هذه الكلمة بالمعنى الشائع إلى ماهية الشخص الإنساني وما يميزه عن غيره من الكائنات البشرية، وفي حين يرى جيلفورد Guilford أن الشخصية تنطوي على جميع خصائص الفرد المستقرة نسبياً، وتتضمن المظهر الجسماني، والذكاء، والاستعدادات، والميول، والاتجاهات والقيم والأمراض النفسية، يرى غيره أن مفهوم الشخصية يقتصر على السمات السلوكية غير المعرفية التي تكشف عن مفهوم الذات، والخلق، والنسق الفردي، وبخاصة المكونات الانفعالية والدافعية لسلوك الفرد، وبذلك يُستبعد الذكاء والاستعدادات والمظهر الجسماني، غير أن التوجهات الحديثة اقترحت ضرورة أخذ الجوانب المعرفية بعين الاعتبار من أجل التوصل إلى نظرية متكاملة للشخصية (علام، 2002، ص 581).

هذا ويشير ألبورت Allport إلى أن مصطلح الشخصية Personality المستعمل في اللغات الأوربية الحديثة يعود في أصله إلى الكلمة اللاتينية Personalitas، وتعني هذه الكلمة القناع أو

الغطاء الذي يتستّر وراءه الشخص للظهور بمظهر معين، فكأن الشخصية بهذا المعنى هي ما يُظهره الفرد للآخرين وليس ما يعبر عن حقيقته أو جوهره أو عالمه الداخلي (الأنصاري، 1999، ص ص 16- 17)، ويتفق هذا المعنى إلى حد بعيد مع بعض تعريفات الشخصية التي تلحّ على المظهر الخارجي للشخص وتأثيره في الآخرين، ولا تعير وزناً كبيراً للنواحي الأخرى الداخلية والخفيّة للشخصية، إلا أن التعريفات الحديثة للشخصية لا تقف عند هذا المعنى فحسب بل تذهب أبعد من ذلك بكثير، بعد أن أصبح موضوع الشخصية أحد أهم الموضوعات وأكثرها حساسية وإثارة الجدل في العلوم الإنسانية بمجالاتها المتنوعة، ولاسيما في العلوم النفسية، وقد أحصى ألبورت ما يزيد على خمسين تعريفاً للشخصية تناولتها من زوايا مختلفة: نفسية واجتماعية وقانونية وأنثروبولوجية وغيرها، مما يشير إلى تعقد موضوع الشخصية من جهة، وتشعب الاتجاهات التي تصدّت له وغناها من جهة أخرى، حيث انطلقت طائفة واسعة من التعريفات النفسية للشخصية، وهي التعريفات التي يطلق عليها عادةً اسم التعريفات السلوكية أو شبه السلوكية للشخصية وتنطلق من الصيغة السلوكية الشهيرة: مثير– استجابة، ويستعير بعضها العنصر الأول من هذه الصيغة ليلصقه بالشخصية، وينظر إلى الشخصية بوصفها مثيراً، ويستعير بعضها الآخر العنصر الثاني من هذه الصيغة ويلصقه بالشخصية لتصبح الشخصية من وجهة نظره استجابة وليست مثيراً، والفئتان متفقتان على الرغم من التعارض القائم بينهما، على أن الشخصية هي ما يظهر للخارج ولا شيء سواه.

وبخلاف التعريفات السلوكية أو شبه السلوكية للشخصية، التي ألحت على ما يظهر في الخارج على حساب ما يجري في الداخل، انطلقت طائفة أخرى واسعة من التعريفات والنظريات من ضرورة إعطاء الأولوية للعالم الداخلي للإنسان أو ما يجري في الداخل، وأعطت الوزن الأكبر لإدراكات الفرد، ومشاعره الذاتية، ودوافعه الخاصة، وقيمه التي قد لا تظهر في سلوكه الخارجي، كما قد يكون من الصعب إخضاعها للملاحظة والقياس، وترفض هذه الطائفة من التعريفات التفسير الآلي أو الميكانيكي للسلوك الذي تطرحه الصيغة السلوكية التقليدية الشهيرة : مثير– استجابة (ميخائيل، 2006، ص ص 21- 22).

وعموماً فإن علماء النفس يمكن أن يتفقوا على عدد من الخصائص العامة لتعريف مقبول ومتسع للشخصية، على الرغم من أن استنادهم إلى تعريفات توجّه دراساتهم يمكن أن تعدّل أو تتغاضى عن إحدى هذه الخصائص أو بعضها، ولعل التعريف الآتي يقدم إطاراً عاماً حول مفهوم الشخصية. فالشخصية تشير إلى تكامل جميع خصائص الفرد المستقرة نسبياً في تنظيم فريد يحدد محاولاته للتكيف مع بيئته الدائمة التغيير، ويمكن تعديله في ضوء هذه المحاولات. ويتضمن هذا التعريف العناصر الآتية:

- تعدد السمات والإمكانات والقدرات التي تشتمل عليها الشخصية.
- انتظام هذه السمات والإمكانات والقدرات في نسق متكامل.
- تفرُّد الشخصية، حيث تميّز بين الفرد وغيره من الأفراد.
- تأثير الشخصية في العلاقات التفاعلية بين الفرد وبيئته، وبغيره من الأفراد.
- اعتبار الشخصية من خصائص الفرد المستقرة نسبياً عبر مدة زمنية طويلة (علام، 2002، ص582).

2- نماذج من استخبارات الشخصية:

تزدحم خزانة الاختبارات النفسية بعددٍ كبير من استخبارات الشخصية، نشرت باللغة الإنجليزية حتى أواخر القرن العقد التاسع من القرن الماضي، يعاني بعضها من عيوب ونواقص كثيرة. ويشير جنثر وجنثر Gunther and Gunther إلى قائمة من الاستخبارات تتضمن أهم عشرة استخبارات رُتبت اعتماداً على عدد البحوث التي أُجريَت حولها أو بواسطتها، وهذه الاستخبارات هي التالية (مرتبة تنازلياً مع ذكر أسماء مؤلفيها):

1- استخبار مينيسوتا المتعدد الأوجه للشخصية MMPI (هاثاواي، ماكنلي).
2- قائمة إدواردز للتفضيل الشخصي EPPS (إدواردز).
3- استخبار عوامل الشخصية الستة عشر 16PF(كاتل وزملاؤه).
4- استخبار كاليفورنيا النفسي أو استخبار كاليفورنيا للشخصية CPI (جف).
5- دراسة القيم Study of values (ألبورت، فرنون، لندزي).
6- استخبار آيزنك للشخصية EPI (آيزنك، آيزنك).
7- قائمة التوجيه الشخصي Personal Orientation (شوستروم).
8- مقياس تنسي لمفهوم الذات Self-Concept (فيتس).
9- قائمة رصد الصفات Adjective check list (جف، هيلبرون).
10- قائمة الشخصية الشاملة OPI (هيست، يونج) (ميخائيل، 2006، ص 63).

3- المشكلات التي تواجه اختبارات الشخصية وقياسها:

يختلف قياس الشخصية عن قياس الذكاء أو غيره من الجوانب المعرفية في أن قياس الشخصية والجوانب غير المعرفية يكون أكثر تعقيداً، حيث تُعاني اختبارات الشخصية المشكلات نفسها التي تعاني منها الاختبارات السيكولوجية بصورةٍ عامة، غير أنه يُضاف إليها وجود مشكلات خاصة بها، ولعل من أهم هذه المشكلات هي مسألة التزييف Packing في الإجابة، إضافةً إلى مشكلة التّغيير الذي يحدث في السلوك المراد قياسه باختبارات الشخصية، وهذا يؤدي بالنتيجة إلى ضُعف ثبات اختبارات الشخصية، وقد يعود سبب ذلك إلى أن الاستجابة على هذه الاختبارات تتأثر بالظروف الواقعة بدرجة أعلى من تأثر المجال العقلي بها.

فالفرد الذي يُطَبّق عليه اختبار الشخصية قد لا تكون واضحة لديه السمات أو الخصائص التي يجري قياسها، أو كيفية تفسير الدرجات، كما أن الاستجابات تكون عرضة لتحيّزات مختلفة، وكذلك ربما تؤثر عملية القياس أو الموقف الاختباري في سلوك الفرد في المُختَبر، مما قد يؤدي إلى تغيير استجابته في الاختبار، وبالتالي جعل صدق الاختبار موضع شك.

في حين أن تحيُّز استجابات الفرد، أو اتجاهه العقلي تُعد من المشكلات التي لا توجد، أو على الأقل توجد بدرجة محدودة في قياس الجوانب المعرفية، والمقصود بالاتجاه العقلي في الاستجابة: ميل أو نزعة الفرد للاستجابة في اتجاه معين، بغض النظر عن محتوى المفردة أو الفقرة الاختبارية، فعندما يتم قياس الانبساط مثلاً يجب قياس ما إذا كان الفرد الذي يُطبَّق عليه المقياس يسلك بطريقة تعبر عن الانبساط، وليس مجرّد الموافقة أو عدم الموافقة على عبارات يتضمنها هذا المقياس حول سلوك الانبساط (علام، 2006، ص217).

ومن المشكلات الأخرى التي تواجه اختبارات قياس الشخصية هي مشكلة الوصول إلى تعريف عام وشامل للشخصية، حيث أن هناك تعريفات تقريباً متفق عليها في بعض المجالات، إلا أنه وكما يبدو من الصعب الوصول إلى تعريف عام وثابت للشخصية، لأن تعريف الشخصية يرتبط بالنظريات المتعددة التي تفسرها، والتي لكل منها مفهوم معين للشخصية، وهذا بطبيعة الحال ينعكس على كيفية قياس الشخصية وما يراد قياسه.

وفي بعض الأحيان قد تظهر بعض الأغراض الأخرى التي ليست من الأغراض المحدّدة التي بُني من أجلها الاختبار وفق نظرية معيّنة، مما يتطلب في هذه الحال إيجاد صدقه البنائي مرة أخرى. وهناك مشكلة أخرى تكاد أن تكون من أصعب مشكلات قياس الشخصية، وهي تتعلق بعينات السلوك المُقاس، حيث عادةً في قياس الشخصية يتم أخذ عينات من السلوك التي تُعتبر ممثلة للمجتمع الأصلي للسلوك المقيس، لكنها في الواقع قد لا تكون كذلك، وبالتالي سيكون الحكم أو القياس في هذه الحالة خاطئاً ولا يعبر حقيقةً عن السلوك المراد قياسه.

بيد أن هذه المشكلات لا تعني الابتعاد عن قياس الشخصية وعدم إعداد اختبارات أو مقاييس لها، إلا أنها تفرض على الباحث أن يكون دقيقاً عند إعداد الاختبار، وأن يأخذ هذه المشكلات بعين الاعتبار كي يضع الحلول لتجاوزها أو التقليل من آثارها، ومن هذه الحلول والمعالجات:

- أن يبتعد الباحث عن الذاتية في إعداد الاختبار، وعند تفسير النتائج، وينطلق الباحث من طبيعة المفاهيم في المجتمع التي يبني عليها الاختبار.
- أن لا يكون في الاختبار مجال للغش أو التأثير المباشر لنوع معين من الإجابة.
- أن يكون مجتمع الدراسة وتفاصيله واضحة أمام الباحث وهذا يرتبط بالتطور النظري له.
- أن يتم اختيار العينة بشكل دقيق ومُمثل، ويُفضل أخذ عينات متنوعة.

- اختيار المقاييس التي تقترب من الحياة الواقعية وتبتعد عن الأسئلة المباشرة في القياس (مجيد، 2007، ص367).

4- رائز مينيسوتا المتعدد الأوجه للشخصية – النسخة الثانية MMPI-2:

يُعد رائز مينيسوتا المتعدد الأوجه للشخصية Minnesota Multiphasic Personality والذي يُعرف اختصاراً بMMPI-2 من أكثر الاختبارات النفسية استخداماً في العالم، أعدّ هذا الرائز كل من ستارك هاثاواي Starke Hathaway وشارنلي ماكنلي Charnley Mckinley، وظهر بصورته الأولى عام 1943متضمناً /550/ بنداً، وأضيف لها فيما بعد /16/ بنداً أو عبارة مكررة في كراسة الأسئلة، بحيث تضمن بصورته الثانية /567/ بنداً أخذت شكل العبارات التقريرية، الإجابة عنها بصح أو خطأ، أُعِدّ هذا الرائز في الأساس لغرض التشخيص التفريقي، والكشف عن بعض الاضطرابات النفسية، ومظاهر سوء التوافق الشخصي والاجتماعي لدى غير الأسوياء بصورة خاصة، لكن وبفضل تراكم البحوث والخبرات العيادية حوله، تحول فيما بعد إلى رائز للسمات الأساسية للشخصية (ميخائيل، 2006، ص77).

يتضمن هذا الرائز /81/ مقياساً هي:

- (3) مقاييس للصدق Validity Scales.
- (3) مؤشرات الصدق الإضافية Additional Validiyt Indicators
- (10) مقاييس عيادية Clinical Scales.
- (31) مقياساً فرعياً للمقاييس العيادية Clinical Subscales.
- (15) مقاييس المحتوى Content Scales.
- (13) مقاييس تكميلية Supplementary Scales : (4) تقليدية Traditional و(9) إضافية Additional
- (6) مقاييس (مجموعات) للبنود الحرجة Critical Item Sets. (عز، قيد النشر، ص ص 4- 5).

وهنا لابد من الإشارة إلى النماذج البحثية في أساس الدراسات التي دارت حول هذا الرائز، والتي قامت ببناء نماذج مقاييس عيادية من بنود الرائز نفسها لكل جانب من الجوانب التي يقيسها، ولكن بعدد بنود مختلف، مما سمح بدراسة أثر الاختلاف في طول هذه المقاييس على كلٍ من الخصائص السيكومترية ودرجات القطع العليا والدنيا، أهم هذه النماذج البحثية النموذج الذي وضعه كلٍ من هاريس ولنجوس Harris and Lingoes، والذي اعتمدته الدراسة الحالية، حيث قامت الباحثة بحساب الفرق في عدد البنود بين المقاييس العيادية في هذا

النموذج والمقاييس العيادية في النموذج الأصلي للرائز الذي أعدّه هاثاواي وماكنلي، واكتفت بإجراء الدراسة على تلك المقاييس التي أظهرت فروقات أكبر في عدد البنود بين هذين النموذجين، لأنها تُمكِّن من دراسة أثر الفرق في عدد البنود على الخصائص السيكومترية ودرجات القطع، وفيما يلي وصف لهذه المقاييس التي أُجريت عليها الدراسة الحالية:

- **مقياس الاكتئاب Depression (D):**

يرتبط بالتفكير الانتحاري، ويعتبر مؤشراً لاستياء الفرد وعدم الرضا عن الحياة. ويتضمن مشاعر تتعلق بالتثبيط وعدم الاهتمام، وفقدان الأمل، وانخفاض الدافعية، والحساسية الزائدة والمشكلات البدنية، وأمراضاً أو شكاوى تتضمن التأخر النفسحركي (المرجع السابق، ص6).

عدد بنود هذا المقياس في النموذج الأصلي الذي أعدّه كل من هاثاواي وماكنلي 57 بنداً، وعدد بنوده في النموذج البحثي المقترح من قبل هاريس ولنجوس49 بنداً، وبهذا يكون الفرق في عدد البنود بين هذين النموذجين ثمانية بنود، أي ما يعادل 14.03% من عدد بنود المقياس الأساسي.

- **مقياس البارانويا (الزّور) Paranoia (Pa):**

وضع لتحديد المرضى الذين يظهرون أعراضاً زورية تتعلق بمرجعية الأفكار، والشك، والشعور بالاضطهاد (المرجع السابق، ص6).

عدد بنود هذا المقياس في النموذج الأصلي الذي أعدّه كل من هاثاواي وماكنلي 40 بنداً، وعدد بنوده في النموذج البحثي المقترح من قبل هاريس ولنجوس34 بنداً، وبهذا يكون الفرق في عدد البنود بين هذين النموذجين ستة بنود، أي ما يعادل 15 % من عدد بنود المقياس الأساسي.

- **مقياس الفصام Schizophrenia (Sc):**

وضع لتحديد المرضى الذين يعانون من أعراض الفصام من حيث محتوى عمليات التفكير الغريبة، والإدراكات الخاصة، والاغتراب الاجتماعي، واضطرابات المزاج والسلوك، وصعوبات في التركيز (المرجع السابق، ص6).

عدد بنود هذا المقياس في النموذج الأصلي الذي أعدّه كل من هاثاواي وماكنلي 78 بنداً، وعدد بنوده في النموذج البحثي المقترح من قبل هاريس ولنجوس68 بنداً، وبهذا يكون الفرق في عدد البنود بين هذين النموذجين عشرة بنود، أي ما يعادل 12.82% من عدد بنود المقياس الأساسي.

- **مقياس الهوس الخفيف Hypomania (Ma):**

وضع لتحديد المرضى الذين يظهرون أعراضاً للاستثارة الهوسية كالتكلف، والهياج، وطيران الأفكار، والمزاج المرتفع، والنشاط الزائد سلوكياً ومعرفياً (المرجع السابق، ص6).

عدد بنود هذا المقياس في النموذج الأصلي الذي أعدّه كل من هاثاواي وماكنلي 46 بنداً، وعدد بنوده في النموذج البحثي المقترح من قبل هاريس ولنجوس34 بنداً، وبهذا يكون الفرق في عدد البنود لهذا المقياس بين هذين النموذجين اثني عشر بنداً، أي ما يعادل 26.08% من عدد بنود المقياس الأساسي.

والجدول ذو الرقم /13/ يوضح الفرق في عدد بنود المقاييس العيادية بين هذين النموذجين، كما يوضح النسب المئوية للفرق في عدد البنود.

5- القرار التشخيصي وأهميته في ضوء مقاييس الشخصية:

تعتمد الدراسة العلمية للسلوك الإنساني على أساليب القياس والتقويم التربوي والنفسي، كما ينبغي أن تستند القرارات المتعلقة بالأفراد مثل قرارات الانتقاء والتصنيف والفحص والتشخيص والتسكين والترفيع والتوجيه والإرشاد إلى هذه الأساليب، والتي يجب ألا يشوبها الحدس أو الذاتية أو التحيز.

ويتضمن مفهوم القياس عملية جمع المعلومات وتنظيمها، وكذلك نتائج هذه العملية، أما مفهوم التشخيص فهو وضع تعريف أو مُسمّى لاضطراب أو مرض باستخدام أحد أنظمة التصنيف المُتفق عليها، والقياس الإكلينيكي هو العملية التي يستخدمها الإكلينيكي في تجميع واستنتاج المعلومات التي يحتاجها للتشخيص، محدداً الأسباب ومخططاً للعلاج والتنبؤ بمآل الاضطراب، وبمجرد تجميع هذه المعلومات يستخدم الباحثون نظم تصنيف معينة لتنظيم وتقييم معرفتهم، والتي تقوم على مجموعة من المحكات أو الخصائص المطابقة لأعراض المرض. ويُعد قياس الشخصية أمراً مهماً للغاية ولاسيما في مجال التشخيص والفحص الإكلينيكي، حيث إن سلوك الفرد في الماضي والحاضر يتنبأ بسلوكه في المستقبل (عبد الرحمن، 2000، ص ص 127- 128).

كما ترتبط التطبيقات النفسية للمقاييس أو الروائز بشكلٍ كبير بعمليتي التنبؤ والتشخيص، ودورهما في تحديد الاضطراب الذي يمكن أن يكون لدى فرد ما، وما سيؤول إليه حال الفرد مستقبلاً، سواء في ميدان الاضطرابات النفسية أو في ميدان الشخصية السوية. (عز، 1990، ص5).

حيث تُستخدم مقاييس الشخصية في غرضين رئيسين أحدهما نظري بحثي، والآخر عملي تطبيقي، فالغرض النظري البحثي يمكّن من تحديد متغيرات الشخصية وتكويناتها الفرضية، عن طريق قياس هذه المتغيرات، ومحاولة الربط بينها وبين أنماط السلوك الملاحظ، أما الغرض العملي التطبيقي فيمكّن من اتخاذ القرارات والتنبؤ بسلوك الأفراد في مواقف عملية متنوعة، حيث تُستخدم مقاييس الشخصية في أغراض التشخيص الإكلينيكي في المستشفيات والعيادات النفسية، وذلك لتحديد مدى وطبيعة الاضطراب في الشخصية الذي يعانيه المريض، كما تُستخدم مقاييس

الشخصية في مجال الإرشاد والعلاج النفسي لاضطرابات الشخصية بأساليب سيكولوجية، حيث إنّ كليهما يعتمد على التفاعل اللفظي بين المريض والمختص النفسي، بغرض مساعدة المريض على تغيير سلوكه لكي يتيسر له التكييف، لذلك تُستخدم مقاييس الشخصية قبل وأثناء وبعد العلاج، حيث تصاحب التشخيص الإكلينيكي (علام، 2002، ص668). بالإضافة إلى استخدامها في مجال انتقاء الأفراد، ولاسيما في المجالات الصناعية والعسكرية، في ترقية الأفراد أو قبولهم في برامج تدريبية معينة.

ويميز كلاينمونتز Kleinmuntz عام 1961بين طريقتين للنظر إلى الغربلة عندما تطبق الروائز النفسية هما: طريقة التشخيص وطريقة التنبؤ. الطريقة الأولى تهدف إلى تعرّف حالات المرض أو الشذوذ من خلال الصفات الحاسمة التي تدل على الاضطراب الانفعالي في موقف من المواقف، أما الطريقة الثانية فتهدف للتنبؤ بظهور بعض هذه الصفات في المستقبل عند من نطبق عليه الرائز في الوقت الحالي، وتعدّ آنستازي Anastasi التشخيص والتنبؤ مفهومين متقاربين في معناهما في ميدان الروائز النفسية (عنبر، 1990، ص31).

وعلى الرغم من العيوب المتعددة التي تحيط بروائز الشخصية فإنّها لم تمنع البتّة من شيوع استخدامها بين الباحثين، وإقبالهم حتى في أقل حدود الاستخدام، باعتبارها أداةً لجمع المعلومات من الأفراد تساعد بدرجة ما في عملية الوصف والتشخيص النفسي، وهناك في الوقت الحالي عدد كبير من روائز الشخصية تذكر في الكتب والمجلات الخاصة بإصدار الروائز ونشرها، إلا أنَّ عدداً قليلاً منها فقط هو الذي يمكن استخدامه بأمان، واختيار أي رائز من روائز الشخصية يستند إلى عدة أمور في مقدمتها اثنان: أولهما شعبية استخدام الرائز وقبوله لدى العاملين في ميدان الخدمة النفسية، وثانيهما نظرية الشخصية والمبادئ التي يستند إليها الرائز المُستخدم (عز، 1990، ص ص 4-5).

يُلاحظ مما سبق أهمية ومكانة الشخصية في علم النفس والقياس النفسي، وكذلك أهمية القياس الإكلينيكي ودوره في مقاييس الشخصية، لأنه يساعدنا في اتخاذ قرارات يمكن وصفها بالموثوقية والموضوعية. بناء على ذلك تم استخدام واعتماد رائز الشخصية MMPI-2 أداةً رئيسية في البحث الحالي باعتباره أداة تشخيصية هامة للاضطرابات النفسية، يغطي عشرة مقاييس عيادية أساسية، يمكن الركون إليها بكثيرٍ من الثقة في أغراض البحث العلمي والغربلة.

ثانياً: العينات الإحصائية

يتضمن هذا المحور العينات الإحصائية من حيث شروطها وطرق اختيارها، ومعايير واستراتيجيات تحديد حجمها، كما يتضمن مصادر الخطأ في العينات، وأخيراً مميزات البحث بالمعاينة وأهميته.

1- المجتمع الإحصائي والعينة sample and Statistical Population:

يُعرَّف المجتمع الإحصائي: بأنَّه مجموعة من الأفراد أو الأشياء أو الحوادث التي يتناولها البحث، أي أنه المجموعة الشاملة التي يجري اختيار العينة منها.

ولذلك عند إجراء أي بحث لابد من تعريف المجتمع، وتحديد ما يشمله من مفردات (أفراد، أشياء ...)، ولابد عند تحديده وتعريفه من تحديد خصائصه من حيث الجنس، والعمر، والمنطقة الجغرافية، والوضع الاقتصادي الاجتماعي.

وعموماً يتصف المجتمع الإحصائي بصفتين رئيستين:

1- المرونة: ويقصد بها إمكانية التحكم بالإطار الزماني والمكاني.

2- التجانس: ويقصد به أن جميع الأفراد متجانسين من حيث الخاصية.

(درويش ورحمة، 2011، ص ص 42- 43)

وعادةً ما يُنظر إلى العينة على أنها جزء أو شريحة من المجتمع، تتضمن خصائص المجتمع الأصلي الذي نرغب في التعرف على خصائصه، ويجب أن تكون العينة ممثلة لجميع مفردات هذا المجتمع تمثيلاً صحيحاً. (القصاص، 2007، ص 78)

ويتم اختيارها من المجتمع الإحصائي نظراً لضخامته وصعوبة الوصول إلى جميع أفراده، لهذا تُستخدم النتائج المستخلصة من العينة للاستدلال على معالم المجتمع، أي تعميم خصائص العينة على المجتمع الذي سحبت منه، لذا يلزم أن تكون ممثلة تماماً في خصائصها لخصائص المجتمع حتى تتم عملية التعميم. مع الأخذ بالحسبان أن عملية سحب العينة تسهم في توفير الوقت والتكاليف التي يمكن أن تُنفق في حال دراسة المجتمع بأكمله، وغالباً ما تقل الفروق بين خصائص المجتمع والعينة كلما كان اختيارها دقيقاً وسليماً.

ويسمى كل مقياس إحصائي محسوب من العينة بالتابع الإحصائي الذي قد يكون متوسطاً، أو انحرافاً معيارياً، أو معامل ارتباط، بينما يسمى كل مقياس إحصائي محسوب من المجتمع بالثابت الإحصائي. ومقدار الفرق بين التابع والثابت يعرف بخطأ المعاينة.

(درويش ورحمة، 2011، ص 44)

2- شروط العينة:

لا يختار الباحث العينة دون التقيد بنظام أو وسيلة علمية، بل هناك شروط ينبغي توافرها في العينة حتى يُستعاض بها عن المجتمع الأصلي الكبير، أهمها ما يلي:

- أن تكون العينة ممثلة Representative للمجتمع الأصلي، فكلما اشتملت العينة على عناصر أكثر من المجتمع، كلما كانت العينة صالحة لتمثيل المجتمع.
- أن تكون لوحدات المجتمع الأصلي فرص متساوية Equal Chances في الاختيار. وكثيراً ما يقع الباحث في خطأ عدم استيفاء هذا الشرط في العينة التي يختارها دون قصد منه، فإذا تعلق البحث بإجراء استبيان على مجموعة خاصة كان من السهل عليه أن يختار الأشخاص القريبين منه أو المُحتكّين به، وهكذا يجعل الاختيار مقتصراً على مجموعة دون غيرها، وبالتالي لا يعطي جميع أفراد المجتمع فرصاً متساوية في الاختيار، وغالباً ما يكتفي الباحث بالشرط الثاني لأن فيه عادةً ضمان لاستيفاء الشرط الأول فإذا ضمنا تساوي فرص الاختيار لجميع الأفراد حصلنا عادةً على عينة ممثلة للمجتمع الأصلي (خيري،1997، ص 197)
- أن تكون الظاهرة المراد عمل معاينة لها سائدة ومنتشرة في المجتمع الأصلي، ولا تكون نادرة الحدوث.
- ضرورة افتراض تجانس مفردات المجتمع الأصلي، وفي حالة تعذر ذلك في بعض المجتمعات غير المتجانسة يلجأ الباحث إلى تقسيمها إلى مجتمعات صغيرة متجانسة.
- ضرورة إجراء حصر مسبق لجميع مفردات المجتمع الأصلي المراد بحثه، مع تقسيم هذا المجتمع إلى وحدات معاينة كل منها داخل قوائم أو ما نسميه إحصائياً بالأطر، فعلى سبيل المثال عند دراسة سكان مجتمع ما فإن وحدة المعاينة إما أن تكون الأسرة كوحدة تحليل أو الفرد أو الجماعة، وقد يكون المجتمع بالنسبة للمجتمعات الكبيرة.
- أن يتناسب اختيار حجم ونوع العينة مع الهدف الأساسي للبحث، أو نوع المشكلة وموضوع الدراسة (القصاص، 2007، ص ص 88- 89).

3- عشوائية المعاينة Random Sampling :

لكي تكون العينة عشوائية ينبغي أن تفي بشرطين أساسيين:

- أن تكون لجميع أفراد المجتمع فرصاً متكافئة للانتقاء في العينة.
- انتقاء أي فرد في العينة لا يؤثر على انتقاء فرد آخر، أي أنه لا يمكننا التنبؤ بالفرد الذي يُنتَقى من معرفتنا لفرد آخر تم انتقاؤه، وتسمى هذه بخاصية الاستقلال Independence، فمثلاً لا يجوز في بحوث استطلاع الرأي في منطقة معينة أنه كلما انتقى الباحث أحد الأزواج في عينته ينتقي زوجته معه، فعلى الرغم من أن هذا يوفر عليه الوقت والجهد، حيث ينتقي فردين من مسكن واحد، إلا أن هذه العينة في هذه الحالة لم تعد تفي بشرط

العشوائية، حيث أن انتقاء الزوجة مرهون بانتقاء الزوج مما لا يحقق خاصية الاستقلال، لذلك فإن العشوائية في انتقاء العينات Randomness لا تعني العرضية أو العفوية، وإنما تتحقق بالعناية التامة بتصميم أساليب المعاينات (علام، 2005، ص 19)

4- **طرق المعاينة(Sampling Methods):**

يمكن حصر كافة الطرق التي تتعامل معها المعاينة في فئتين رئيسيتين هما:

أولاً: طرق المعاينة الاحتمالية: Probability Sampling Methods:

وهي تلك الطرق التي تعتمد على العشوائية Randomization في اختيار مفردات العينة من مجتمع الدراسة، حيث يكون لكل فرد من أفراد المجتمع فرصة متساوية لاختياره فيها، وأيضاً تكون له الفرصة نفسها لاختياره ضمن الفئة الواحدة المتجانسة المكونة للمجتمع الإحصائي.

وتشمل هذه العينات: العينة العشوائية البسيطة والطبقية والمنتظمة والعنقودية، وكذلك العينة العشوائية متعددة المراحل.

ثانياً: طرق المعاينة غير الاحتمالية: Non Probability Sampling methods

في حالات كثيرة تكون طرق المعاينة الاحتمالية غير ممكنة، فتكون طرق المعاينة غير الاحتمالية موضع الاستعمال، ويجب الأخذ بالحسبان أن رغبة الباحث وأحكامه الشخصية تتدخل في هذه الطرق، أي أن الباحث له دور مباشر في عملية اختيار المفردات التي ستشملها عينة البحث. وتكون أحياناً هذه المعاينة مفيدة عندما يشعر الباحث أن بعض المفردات أو العناصر هي أفضل من المفردات الأخرى في تمثيل مجتمع البحث (الجادري، 2007، ص ص 29 – 31).

والجدولان الآتيان يوضحان طرق كل من نوعي المعاينة، وخطوات سحب كل منها، ومزاياها وعيوبها، ومجالات مناسبتها.

الجدول رقم 1: طرق المعاينة الاحتمالية

نوع العينة	خطوات سحبها	مزاياها وعيوبها	مجالات مناسبتها
العشوائية البسيطة	يتم سحبها باستخدام أسلوب الاختيار العشوائي، وذلك باستخدام الاقتراع المباشر أو جداول الأرقام العشوائية، أو باستخدام الحاسب الآلي.	مزاياها: – تضمن أن يكون لجميع عناصر المجتمع نفس الفرصة في الاختيار. – تعتبر أسهل طرق المعاينات الاحتمالية وأكثرها استخداماً. – تمثل الأساس الذي يرتكز عليه اختيار العينات العشوائية الأخرى. عيوبها: – إذا كان المجتمع مكوناً من طبقات غير متجانسة من حيث المتغيرات المراد دراستها، قد يكون تمثيل كل مجموعة من المجتمع غير دقيق، وبالتالي تكون العينة متحيزة.	– في المجتمعات المتجانسة وفي الدراسات التي يكون الغاية منها التعميم.
العشوائية المنتظمة	يتم سحبها باتباع الخطوات الآتية: – تحديد أرقام لمفردات المجتمع كافة. – اختيار عشوائي لمفردة البداية للعينة من بين الأرقام (1 إلى 9) – تحديد مقدار الزيادة المنتظمة وفقاً للقانون: الزيادة المنتظمة = حجم المجتمع / حجم العينة –إضافة مقدار الزيادة المنتظمة إلى مفردة البداية لنحصل على المفردة التالية لشمولها في عينة الدراسة، وتستمر إضافة الزيادة المنتظمة إلى أن يتم الوصول إلى العدد المطلوب.	مزاياها: – قلة تكاليفها وسهولة إجرائها. – قلة الأخطاء الناتجة عن اختيار مفرداتها. عيوبها: – عدم صلاحيتها في الحالات التي توجد فيها علاقة دورية عند ترتيب عناصر المجتمع ، وفي هذه الحالة تكون العينة متحيزة بدرجة كبيرة.	– في المجتمعات الكبيرة التي تحتاج إلى تقسيم إلى فئات متساوية.

العينة العشوائية الطبقية	يتم سحبها باتباع الخطوات الآتية: – تقسيم المجتمع إلى طبقات على أساس متغير أو أكثر حسب نوع الدراسة، وتكون كل طبقة من هذه الطبقات على درجة كبيرة من التجانس قدر الإمكان. – ثم اختيار عينة عشوائية (بسيطة أو منتظمة) من كل طبقة من المجتمع.	مزاياها: – تساعد على تكوين مجموعات تتسم بالتجانس الداخلي بالنسبة للخصائص المرتبطة بالدراسة. – حجم الخطأ المعياري فيها أقل. – تمكن الباحث من الحصول على درجة عالية من الدقة في النتائج. عيوبها: – من الصعب الحصول على عينة عشوائية طبقية، لأنها تتطلب من الباحث معرفة دقيقة بخصائص مجتمع البحث.	– في المجتمعات كبيرة الحجم. – وكذلك المجتمعات غير المتجانسة بالنسبة للخصائص والمتغيرات موقع الدراسة.
العينة العشوائية متعددة المراحل	يتم سحبها باتباع الخطوات الآتية: – تقسيم المجتمع إلى عدد من الأقسام المتشابهة أو المتجانسة. – الاختيار العشوائي لعدد من المفردات بكل قسم بحيث يتم ذلك تتابعياً. – الاختيار العشوائي من القسم الأول كمرحلة أولى. –ثم الاختيار العشوائي من القسم الثاني كمرحلة ثانية. – حتى يتم الوصول إلى الاختيار العشوائي في المرحلة النهائية، ولذا أُطلق على هذا النوع من العينات بأنه متعدد المراحل.	مزاياها: – قدرتها على تمثيل المجتمعات الكبيرة. عيوبها: – كثرة المراحل التي تتضمنها تضعف العلاقة بين معالم المجتمع الأصلي وخصائص العينة. – تتطلب جهداً كبيراً في تحديد كل مرحلة وحجم العينة المطلوبة منها.	– في المجتمعات الكبيرة والتي أفرادها متفرقون في أنحاء متباعدة من المجتمع. – يلائم هذا النوع الدراسات السكانية والدراسات في مجال الجغرافيا الاقتصادية.

العينة العنقودية	يتم سحبها باتباع الخطوات الآتية: – تقسيم المجتمع الإحصائي إلى عدد من العناقيد (الوحدات) وفقاً لمعيار معين غالباً ما يكون جغرافياً، تسمى كل مجموعة عنقوداً، وتكون وحدة الاختيار في هذا النوع من العينات هي العنقود. – ثم تؤخذ عينة من هذه العناقيد (الوحدات) بشكل عشوائي. – وبعد ذلك تقسم وحدات المرحلة الأولى إلى وحدات أصغر تسمى بالوحدات الثانوية ويطلق عليه وحدات المرحلة الثانية. – ومن ثم تؤخذ عينة منها. وتستمر عملية التقسيم إلى وحدات أصغر حتى نصل إلى الوحدات أو العناصر النهائية لغرض الدراسة.	مزاياها: – تحقق الدقة وترفع من درجة تمثيل العينة للمجتمع الأصلي. عيوبها: – أسلوبها يكتنفه التعقيد. – مكلفة من حيث الوقت والجهد والنفقات.	– يتم اعتمادها عندما يكون مجتمع الدراسة كبيراً وأفراده متفرقون في أنحاء متباعدة من المجتمع.

الجدول رقم 2: طرق المعاينة غير الاحتمالية

نوع العينة	خطوات سحبها	مزاياها وعيوبها	مجالات مناسبتها
العينة المتيسرة	يتم اختيارها بواسطة الباحث لأنها متاحة فقط وليس لأي سبب آخر.	مزاياها: – تكوين فكرة أولية عن موضوع البحث. – سهولة الإجراء، وقلة التكاليف. عيوبها: – فيها تحيز عالي، ولذلك يستحيل تمثيل المجتمع بواسطة هذا النوع من العينة، وبالتالي لا يمكن تعميم نتائجها.	– في الأبحاث التي لا تستخدم نتائجها للتعميم. – في الأبحاث التي تتطلب الحصول على معلومات من أفراد في المجتمع يسهل الوصول إليهم.

العينة المقصودة	يتم سحبها من قبل الباحث من أفراد يعرف مسبقاً بأنهم الأقدر على تقديم المعلومات عن الظاهرة قيد الدراسة، أو تتوافر لديهم معايير محددة تفيد الدراسة.	مزاياها: - تضمن وجود وحدات معاينة مناسبة للدراسة. - سهولة الإجراء، وقلة التكاليف. عيوبها: - العينة غير ممثلة للمجتمع وبالتالي لا يمكن تعميم نتائجها.	- للحصول على معلومات من شريحة محددة، إما لأنها قادرة على إعطاء المعلومات المطلوبة، وإما لأن بعض المعايير التي وضعها الباحث تتوافر فيهم.
العينة الحصصية	يتم سحبها باتباع الخطوات الآتية: - تقسيم المجتمع إلى حصص تشبه الطبقات على أساس الخصائص موضع الدراسة. - تحديد نسبة المجتمع في كل طبقة أو حصة. - اختيار عينة من كل حصة ويتم الاختيار بطريقة غير عشوائية.	مزاياها: -قليلة الجهد والتكاليف. -ذات فائدة عملية في المراحل الأولى من البحث، وبذلك يمكن استخدامها في الدراسات الاستطلاعية لزيادة خبرة الباحث بمشكلته البحثية. عيوبها: -نظراً لأن الباحث يصنف مفردات عينته حسب خبرته، فقد يؤدي ذلك إلى التحيز، وبالتالي لا يمكن تعميم نتائجها.	- استطلاعات الرأي العام. - عندما يهدف الباحث إلى الحصول على عينة مماثلة في خصائصها لخصائص المجتمع.

(الجادري، 2007، ص ص 29-36، الصياد وربيع، 1983، ص ص 105- 117، القصاص، 2007، ص ص79- 87).

5- تقدير حجم العينة:

هناك عوامل عديدة لتحديد حجم العينة أهمها:

- أن التباين بين أفراد المجتمع وليس حجم المجتمع هو العامل الحاسم في تقدير حجم العينة، فكلما كبر حجم التباين بين أفراد المجتمع، كلما استلزم ذلك أن يكون حجم العينة كبيراً بغض النظر عن حجم ذلك المجتمع .
- عند تقدير حجم العينة لابد من الأخذ في الحسبان التكلفة المادية، والزمن اللازم لإنجاز الدراسة.
- مدى التفصيل المطلوب في نتائج العينة، فكلما زادت درجة التفصيل المطلوبة زاد حجم العينة المسحوبة.

- مدى الخطأ الذي يسمح به في نتائج العينة، فكلما قل مدى الخطأ الذي يمكن السماح به زاد حجم العينة.
- درجة الثقة المطلوب توفرها في تحقيق السمات السابقة . فكلما زادت درجة الثقة المطلوبة، زاد حجم العينة اللازم.
- حجم المجتمع: يجب أن يتناسب حجم العينة مع حجم المجتمع الأصلي.
- عندما يكون هناك توقع أن للمعالجة التجريبية أثراً قليلاً على المتغير التابع، أو عندما يتوقع أن تكون العلاقة بين المتغيرات قيد الدراسة فيها ضعيفة فإنه من المرغوب فيه أن تكون العينة كبيرة الحجم.
- عندما يكون هناك توقع لإهدار وفقدان كبير في عدد المفحوصين، أو هناك عدد من الأفراد لن يُعيد الاستبيانات، لابد من اختيار عينة كبيرة الحجم، حتى لا تتأثر النتائج بهذا الفقدان أو الهدر (المنيزل، 2006، ص ص 25 – 26).

مع الإشارة إلى أنّ استخدام عينات ذات حجوم صغيرة في بعض البحوث ربما كان أفضل من استخدام عينات ذات أحجام كبيرة، مثل الدراسات التي تعتمد على التحليل النفسي وأدوات القياس الإسقاطي. كما أن عينة كبيرة الحجم في التصميم التجريبي تعد ضعيفة لأنه سيكون هناك عدد كبير من المتغيرات غير المضبوطة (درويش ورحمة، 2011، ص 55).

6-معايير تحديد حجم العينة:

هنالك ثلاثة معايير هامة عادةً ما نحتاج إلى تحديدها لتقدير حجم العينة المناسب وهي كالآتي:

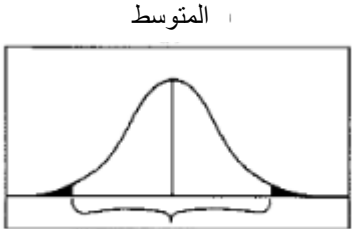
- **مستوى الدقة: The Level of Precision:**

ويسمى أحياناً مستوى الخطأ في اختيار العينات، وهو المدى الذي يحدد القيمة الحقيقية لحجم العينة، وكثيراً ما يُعبَّر عنه بنسب مئوية (على سبيل المثال، $\pm 5\%$)، من مثل الطريقة التي يُعبَّر بها عن نتائج استطلاعات الرأي للحملات السياسية في وسائل الإعلام. وبالتالي إذا وجد الباحث أن 60% من المزارعين مثلاً في العينة قد اعتمدوا الممارسة الموصى بها لزراعة معينة بمعدل دقة $\pm 5\%$ ، فإنه يمكن أن نخلص إلى أن ما بين 55% و 65% من المزارعين في مجتمع البحث قد اعتمدوا هذه الممارسات الزراعية.

- **مستوى الثقة: The Confidence Level :**

يستند مستوى الثقة بشكل أساسي إلى مقاييس النزعة المركزية، والفكرة الأساسية في هذه النظرية هي أنه عندما يتم أخذ عينات من مجتمع البحث مراراً وتكراراً، فإن متوسط قيمة السمة التي يتم الحصول عليها عن طريق تلك العينات يساوي القيمة الحقيقية لهذه السمة عند مجتمع البحث، إضافةً إلى ذلك فإن القيمة التي تم الحصول عليها عبر هذه العينات يتم توزيعها توزيعاً

طبيعياً حول القيمة الحقيقية، وبالتالي بعض العينات يكون لها قيم أعلى من القيمة الحقيقية والبعض الآخر من هذه العينات يكون لها قيمة أدنى من هذه القيمة في التوزيع الطبيعي، وبعبارة أخرى أنه إذا تم تحديد مستوى ثقة 95%، فإن 95 من أصل 100مفردة في العينة لديها قيم المجتمع الحقيقي نفسها في مستوى الثقة المحدد مسبقاً، إلا أنه هناك دائماً خطر أن لا تمثل العينة التي نحصل عليها القيمة الحقيقية للمجتمع الأصلي، يتم تمثيل مثل هذه العينات مع القيم المتطرفة في المناطق المظللة في الشكل الآتي، ولابد من الإشارة إلى أن هذا الخطر يقل في مستوى الثقة 99%، ويزداد في مستويات الثقة 90% فما دون.



الشكل رقم 1 : توزُّع قيم أفراد العينة ضمن انحرافين معياريين، عند مستوى الثقة 95%

- **الدرجة التباينية في السمات التي يجري قياسها: Degree of Variability:**

ويمثل هذا المعيار درجة توزُّع السمات في مجتمع البحث، وكلما كان هذا المجتمع غير متجانس بدرجة أكبر، أي كلما كان التباين أكبر في السمة موضع البحث، كلما تتطلب ذلك زيادة حجم العينة المطلوبة للحصول على مستوى الثقة المطلوب، وكلما قل التباين في السمة بين أفراد المجتمع كلما صغر حجم العينة المطلوبة، ويمكن القول بأن نسبة 50% تشير إلى وجود مستوى أعلى من التباين من نسبة 20 أو 80% ، وذلك لأن 20 و 80% تشير إلى أن أغلبية الأفراد في المجتمع لا يمتلكون أو يمتلكون هذه السمة المقاسة، ولذلك النسبة 50% تشير إلى الحد الأقصى للتباين في مجتمع البحث (Glenn, 2009, PP. 1 – 2)

7- استراتيجيات تحديد حجم العينة:

هناك العديد من الاستراتيجيات لتحديد حجم العينة وهي كالآتي:

- استراتيجية التعداد للمجتمعات الصغيرة: Using a Census Strategy for Small Populations:

يتم فيها استخدام كامل مجتمع البحث كعينة، وعلى الرغم من أن مشكلة التكاليف المادية تجعل هذا الأمر مستحيلاً بالنسبة للمجتمعات كبيرة العدد، إلا أن هذه الطريقة أي التعداد، مفيدة جداً بالنسبة للمجتمعات البحثية الصغيرة (على سبيل المثال 200 نسمة أو أقل)، لإن هذه الاستراتيجية تلغي الخطأ في اختيار العينات، وتقلص التكاليف المادية، إضافةً إلى ذلك توفر بيانات عن جميع أفراد المجتمع، حيث عملياً كامل المجتمع سيكون قد ضُمِّن في العينة باستخدام هذه الاستراتيجية لتحقيق المستوى المطلوب من الدقة.

-استخدام حجم العينة نفسها لدراساتٍ مماثلة: Using a Sample Size of Similar Studies:

في هذه الاستراتيجية يتم استخدام حجم العينة نفسه المستخدم في دراسة مماثلة لتلك الدراسة التي يخطط لها الباحث، ولكن استخدام هذه الاستراتيجية دون مراجعة الإجراءات المستخدمة في تلك الدراسات قد تُعرِّض الباحث لخطر تكرار الأخطاء نفسها التي تم الوقوع بها في تحديد حجم العينة لتلك الدراسات، لكن ومع ذلك، فإن مراجعة الأدبيات بدقة وإحكام يمكن أن يوفر للباحث دليلاً حول حجم العينة النموذجي التي يجب أن يستخدمها.

- استخدام جداول النشر: Using Published Tables:

يتم الاعتماد على الجداول المنشورة، والتي توفر حجم العينة المناسب لمجموعة من المعايير، والملحقان رقم (1) و(2) يوضحان أحجام العينات المناسبة للوصول إلى مستويات الدقة والثقة المطلوبة، لكن لابد من أخذ أمرين مهمين بعين الاعتبار: الأمر الأول: أن هذه الأحجام من العينات تعكس عدد الاستجابات التي تم الحصول عليها وليس بالضرورة كامل المسوحات المطبقة والمخطط لها، وغالباً ما يتم رفع هذا العدد لتعويض عدم الإجابة، والأمر الثاني: أن أحجام العينات في الملحق (2) تفترض أن السمات التي يتم قياسها موزعة توزيعاً طبيعياً أو ما يقارب التوزيع الطبيعي، وفي حال كان هذا الافتراض غير مُحقَّق يجب أن يتم إجراء مسح كامل للمجتمع موضع البحث.

- استخدام الصيغ الرياضية لحساب حجم العينة: Using Formulas to Calculate a Sample Size:

على الرغم من أن الجداول يمكن أن توفِّر دليلاً مفيداً لتحديد حجم العينة، إلا أن الباحث قد يحتاج لحساب حجم العينة اللازمة لمستويات معينة من الدقة أو الثقة المحددة بما يتناسب مع متطلبات ونوع دراسته، تبعاً لتباين السمات في المجتمع، لذا يتم اللجوء إلى تطبيق واحدة من الصيغ المستندة إلى المعادلة الآتية:

$$n = \frac{N}{1+N(e)^2}$$

حيث:

n = حجم العينة

N = عدد أفراد مجتمع البحث

e = مستوى الدقة المطلوب

(Glenn, 2009, PP. 2 – 3)

على الرغم من وجود هذه الاستراتيجيات الأربعة إلا أنه يمكن القول بوجود اتجاهين يُعمَل بهما في تحديد حجم العينة، هما:

الاتجاه الأول : يعتمد على خبرات الباحث وخبرات الباحثين الآخرين في هذا المجال، ووفقاً لهذا الأسلوب فإننا نجد أن الإحصائيين ينصحون بأن يكون الحد الأدنى للأفراد كما يلي (50-100) في الدراسات الارتباطية Correlational studies، ويفضل أن يكون حجم العينة في الدراسات التجريبية experimental studies ذات المعالجة الواحدة 30 فرداً، أما في الدراسات التجريبية ذات المعالجتين أو أكثر فيفضل ألا يقل عدد الأفراد في الخلية الواحدة عن 5 أفراد، أما في الدراسات المسحية Survey studies: فالعدد المناسب 20% من أفراد المجتمع الكلي إذا كان عدد المجتمع (500 إلى 1000) وتنخفض النسبة إلى 5% في المجتمعات الكبيرة جداً.

والاتجاه الثاني: يعتمد على القواعد الاحتمالية الرياضية لتحديد حجم العينة، كاحتمال الوقوع في الخطأ من النوع الأول أو الخطأ من النوع الثاني (درويش ورحمة،2011، ص 55).

يقترح عدد من الباحثين أن يكون عدد أفراد العينة كما يوضحه الجدول الآتي:

الجدول رقم 3: عدد أفراد العينة من وجهه نظر الباحثين

أسلوب البحث	عدد أفراد عينة الدراسة
الدراسات الارتباطية	**(30) فرداً على الأقل**
الدراسات التجريبية	**(15) فرداً في كل مجموعة من المجموعات التجريبية والضابطة، وكلما زاد عدد أفراد العينة كلما كانت أكثر صدقاً في نتائج الدراسة**
الدراسات الوصفية	**(20%) من أفراد مجتمع صغير نسبياً (بضع مئات)** **(10%) لمجتمع كبير (بضعة آلاف)** **(10%) لمجتمع كبير جداً (عشرات الآلاف)**

إلا أنَّ زيادة حجم العينة عن هذه الحدود يتطلب الأخذ بالحسبان ما يلي:

- وجود متغيرات غير مضبوطة بأي من طرق ضبط المتغيرات الدخيلة.
- توقع معاملات ارتباط أقل، لأنه يترتب على زيادة حجم العينة زيادة في متغيرات البحث. (ملحم، 2001، ص155)

8- حجم العينة وقوة الاختبار:

تتوقف قوة الاختبار إلى حدٍ كبير على حجم العينة المستخدمة في الدراسة، ويرى Stevens أنه عندما يكون حجم العينة كبيراً (كأن يكون – مثلاً- عبارة عن مائة فرد أو أكثر)، لاتكون قوة الاختبار مشكلة في حد ذاتها، أما إذا كان حجم العينة في إحدى الدراسات صغيراً (على سبيل المثال n = 20)، فقد يرجع احتمال وجود نتيجة غير ذات دلالة إلى عدم وجود قوة كافية. ويقترح Stevens أنه عند وجود عينات صغيرة الحجم، قد يكون من الضروري ضبط

مستوى الدلالة لتعويض انخفاض الحجم، على سبيل المثال، أن يكون الحد الموضوع هو (0.10) أو (0.15) بدلاً من مستوى (0.05) التقليدي (بالانت، 2006، ص223).

9- مصادر الخطأ في العينات:

في أغلب الأبحاث يتم الحصول على نتائج من العينات قد لا تماثل تماماً النتائج التي يتم الحصول عليها من الحصر الشامل، وذلك لأن العينات عرضة لنوعين من الأخطاء هي:

- أخطاء التحيُّز.
- الأخطاء الاحتمالية.

أخطاء التحيُّز: ويُقصد بها تلك الأخطاء الناتجة عن انحياز الباحث إلى فئة دون أخرى وإعطائها وزناً أكبر مما تستحق، مثل اختيار الطلاب المتفوقين في الصف ليختبر ذكائهم الأمر الذي يؤدي إلى رفع متوسطات الذكاء، أو أن يختار الباحث عينته ممن هم في متناول يده، كأن يختار أفراد العينة من مدرسة أو مدرستين قريبتين من مكان سكنه، أو ممن يرتبطون بصلة أو علاقة ما بالقائمين على هذه المدارس، وفي هذه الحالة ينطوي تعميم النتائج على مخاطر كثيرة، لأنه لم تؤخذ صفة العشوائية في الاختيار ولم تتوفر فرص متكافئة أمام جميع أفراد المجتمع للظهور في العينة. وقد يتم اختيار بعض الأفراد ليسوا من مجتمع الدراسة. (درويش ورحمة،2011، ص 56)

ويرجع حدوث أخطاء التحيُّز لعدد من العوامل من بينها:

- سوء التقدير وعدم توفر الدقة من جانب الباحث، وذلك عند قيامه بعمليات الحصر حيث قد تفوته الدقة الكافية في حساب المتغيرات، وكذلك عدم توفُّق الباحث في صياغة الفروض الصحيحة.
- صياغة أسئلة غامضة وغير واضحة للمفحوصين.
- عدم استجابة بعض أفراد العينة لأسئلة المقياس.
- الاختيار المقصود غير العشوائي لأفراد العينة.
- سوء اختيار العينة، وقد يحدث ذلك نتيجة لسحب العينة من إطار غير كامل، كاستخدام إطار قديم أو إطار غير شامل لجميع مفردات المجتمع، حيث يؤدي ذلك إلى تحيز العينة للمفردات الموجودة في الإطار فقط.
- استبدال أفراد العينة بأفراد آخرين لعدم تمكن الباحث من الوصول لهؤلاء الأفراد. (القصاص، 2007، ص ص 99- 100)

الأخطاء الاحتمالية: هذه الأخطاء لا يد للباحث فيها، وهي ناتجة عن دراسة جزء من المجتمع، وهي تنشأ في الاتجاهين الموجب والسالب وبالحجم نفسه، ولذلك فهي تلغي بعضها بعضاً وخصوصاً مع زيادة حجم العينة وذلك وفق نظرية النزعة المركزية central limit

Theorem ، إذ أنه مهما كان شكل توزيع المجتمع الأصلي، فإن توزيع المعاينة للمتوسط الحسابي يؤول إلى التوزيع الطبيعي تدريجياً مع زيادة حجم العينة، أي أن الفرق بين قيمة متوسط العينة ومتوسط المجتمع يصغر مع زيادة حجم العينة، بمعنى أن مقدار الخطأ المعياري يتقلص مع زيادة حجم العينة، لأن سحب عينة كبيرة يعني الحصول على معلومات أكثر، مما يقود إلى دقة أكبر في تلك المعلومات (درويش ورحمة،2011، ص 56).

يُلاحظ ممّا سبق أهمية البحث بالمعاينة وفوائده، من حيث اختصار الوقت والجهد اللازمين لإتمام البحث وبالتالي اقتصاد التكاليف، كما يمكن بواسطته الحصول على الردود الكاملة الدقيقة إذا ما تم استخدام جزء ممثل من المجتمع الكلي، وتتبع غير المستجيبين، وكذلك الحصول على بيانات أكثر من أفراد العينة، وتلخيصها، وتحليلها على وجه السرعة، بينما يكون ذلك صعباً في حالة الحصر الشامل.

ثالثاً: الخصائص السيكومترية للاختبار والعوامل المؤثرة فيها

يقفُ هذا المحور عند مفهوم الثبات وأنواع معاملاته والعوامل المؤثرة فيها، والعلاقة بين طول الاختبار وثباته، كما يتضمن مفهوم الصدق وأهميته وأنواع معاملاته والعوامل المؤثرة فيها، وكذلك العلاقة بين الصدق والثبات، وأخيراً إضاءة على مشكلة ثبات اختبارات الشخصية.

1- مفهوم الثبات وأهميته:

يُقصد بمفهوم ثبات درجات الاختبار مدى خلوّها من الأخطاء غير المنتظمة التي تشوب القياس، أي مدى قياس الاختبار للمقدار الحقيقي للسمة التي يهدف لقياسها، فدرجات الاختبار تكون ثابتة Reliable إذا كان الاختبار يقيس سمة معينة قياساً متسقاً في الظروف المتباينة التي قد تؤدي إلى أخطاء القياس، فالثبات بهذا المعنى يعني الاتساق أو الدقة في القياس.

وأخطاء القياس التي تؤثر في درجات الاختبارات قد تكون أخطاء عشوائية يصعب التنبؤ بها من موقف إلى آخر ولذلك تعمل على خفض ثبات الدرجات، وهذه الأخطاء ترجع إلى عوامل بعضها يتعلق بالاختبار ذاته مثل: عدم وضوح مفرداته وغموض تعليماته، وعدم تحديد محكات تصحيح مفرداته، وبعضها يتعلق بالظروف البيئية مثل: الإضاءة والتهوية والضوضاء ومدى ملاءمة غرفة الاختبار، وبعضها الآخر يتعلق بخصائص الأفراد المُختبرين مثل: قلّة دافعيتهم، وشعورهم بالتعب أو الملل أو القلق، وحالتهم النفسية والصحية بعامة وقت إجراء الاختبار، وقد تكون أخطاء القياس منتظمة حيث تؤدي إلى تضخيم أو تقليل درجات الاختبار ولكن بطريقة متسقة وبالتالي لا تؤثر في ثبات الدرجات. (علام، 2002، ص131)

فالأخطاء المنتظمة Systematic Error تؤثر في جميع المفحوصين بالدرجة نفسها، أما الأخطاء العشوائية Random Error فيختلف تأثيرها من مفحوص إلى آخر من حيث المقدار أو الاتجاه سواء أكان سلبياً أم إيجابياً.

وبهذا يمكن القول إن مفهوم الثبات بالمعنى العام يُستخدم للدلالة على مدى اعتماد الفروق في درجات الاختبار على الفروق الحقيقية في السمة أو الخاصية المقيسة من جهة، ومدى خلوها من أخطاء القياس العشوائية أو الناتجة عن الصدفة من جهة أخرى. (ميخائيل، 2008، ص ص181 – 182)

ويدرج عبد الرحمن ثلاثة مفاهيم أساسية لثبات درجات الاختبار على النحو الآتي:

- الحصول على النتائج نفسها تقريباً عند إعادة التطبيق.
- أن يكون التباين الحقيقي أكبر ما يمكن بالنسبة للتباين العام، وتباين الخطأ أقل ما يمكن بالنسبة للتباين العام.

- وجود العلاقة القانونية بين وحدات الاختبار: فالعلاقة القانونية بين وحدات الاختبار تدل على التناسق في البناء الداخلي للاختبار، وهذا يعني أن معامل ثبات الاختبار سوف يتوقف على العلاقة أو الارتباط بين كل وحدة ووحدة أخرى من وحدات الاختبار، وهذا ما يسمى بالارتباطات البينية. (عبد الرحمن، 1998، ص ص 165- 166)

ومن التعريفات المهمة للثبات التعريفات الآتية:

- يشير مفهوم الثبات إلى: الدرجة الحقيقية التي تُعبّر عن أداء الفرد في اختبار ما، ومعنى ثبات الدرجة أن المفحوص يحصل عليها في كل مرة يُختبر فيها سواء بالاختبار نفسه أو بصورة مكافئة له تقيس الخاصية نفسها، وسواء اختبر في الظروف نفسها أو في ظروف مختلفة لا تتدخل فيها عوامل عشوائية. (فرج، 2007، ص295)
- ويشير مفهوم الثبات عند جيلفورد Guilford, 1954 إلى: نسبة التباين الحقيقي في الدرجة المستخلصة من اختبار ما، إلى التباين الكلي للدرجة على الاختبار. (المرجع السابق، ص 296)
- ويشير مفهوم الثبات عند علام إلى: اتساق درجات الاختبارات والمقاييس لمجموعة معينة من الأفراد، أي الاتساق عبر الزمن، أو اتساق صيغ مختلفة من نفس الاختبار، أو اتساق مفردات الاختبار ذاته، أو الاتساق عبر أفراد مختبرين ومصححين مختلفين. (علام، 2006، ص89)
- كما يمكن تعريف الثبات بلغة الإحصاء على أنه: نسبة التباين الحقيقي إلى التباين الكلي، مع ملاحظة أن التباين الذي ينطوي عليه مفهوم الثبات يشمل التباين الناتج عن فروق حقيقية في أداء المفحوصين والتباين الناتج عن الخطأ المنتظم، ومن الواضح أن الثبات يزداد كلما نقص تباين الخطأ – أي التباين الناتج عن الخطأ العشوائي– وينخفض كلما ازداد هذا الأخير. (ميخائيل، 2008، ص183)

2- درجة الثبات والدرجة الحقيقية:

إنَّ الدرجة التي يحصل عليها الفرد في مقياس معين تسمى الدرجة الملاحظة obtained Score، غير أن هذه الدرجة تكون في كثير من الأحيان مشوبة بأخطاء القياس التي إذا تم التمكُّن من تحديد مقدارها فإنه يتم الحصول على درجة الخطأ Error Score، وإذا تم طرح درجة الخطأ من الدرجة الملاحظة فإنه يتم الحصول على درجة الفرد خالية من الأخطاء العشوائية، وهذه الدرجة تسمى الدرجة الحقيقية True Score أي أنَّ:

الدرجة الحقيقية = الدرجة الملاحظة – درجة الأخطاء العشوائية

وأنَّ الدرجة الملاحظة = الدرجة الحقيقية + درجة الأخطاء العشوائية.

وبهذا فإن الدرجة الملاحظة للفرد في الاختبار لا تساوي درجته الحقيقية إلا إذا كانت درجة الخطأ صفراً، وهي تزيد عن درجته الحقيقية إذا كانت درجة الخطأ موجبة، وتقل عنها إذا كانت درجة الخطأ سالبة.

فإذا لجأ الفرد إلى التخمين في إجابته عن مفردات الاختبار وتوصل إلى الإجابة الصحيحة، فإن درجته الملاحظة تزيد عن درجته الحقيقية، أما إذا كان الفرد متعباً أو أن مفردات الاختبار وتعليماته غامضة فإن درجته الملاحظة سوف تقل عن درجته الحقيقية.

غير أن هذه الأخطاء الموجبة والسالبة يمحو بعضها البعض الآخر على المدى البعيد استناداً إلى قوانين الصدفة ، وبذلك يصبح متوسط درجات الخطأ صفراً، ولذلك فإن متوسط الدرجات الملاحظة لمجموعة الأفراد الذين يطبق عليهم الاختبار يعد أفضل تقدير لمتوسط درجاتهم الحقيقية في الاختبار.

كما أن تباين الدرجات الملاحظة يساوي مجموع تباين كل من الدرجات الحقيقية وتباين الخطأ، ومن المعروف أن مقدار التباين = مربع الانحراف المعياري، فإذا كانت الدرجات الملاحظة خالية من الخطأ فإن مقدار تباين الخطأ يساوي صفر، ويصبح مقدار تباين الدرجات الملاحظة مساوياً مقدار تباين الدرجات الحقيقية، أي تكون الدرجات متسقة اتساقاً تاماً، وتكون قيمة معامل الثبات واحداً صحيحاً.

أما إذا كانت درجات الاختبار غير متسقة على الإطلاق، فمعنى ذلك أن الدرجات لا تمثل سوى أخطاء القياس، أي أن الفروق الملاحظة بين الأفراد في هذه الحالة لا يكون لها أي علاقة بالفروق الحقيقية بينهم، وتكون قيمة معامل الثبات صفراً.

لذلك فإن الانحراف المعياري لدرجات الخطأ يسمى الخطأ المعياري للقياس Standard Error Of Measurement. فإذا كان مقدار هذا الخطأ المعياري صفراً تكون درجات الاختبار متسقة اتساقاً تاماً، ولكن هنالك عادةً في القياس التربوي والنفسي قدراً معيناً من الخطأ المعياري، مما يؤدي إلى خفض ثبات درجات الاختبارات بقيم متفاوتة بحسب مقدار هذا الخطأ. (علام، 2002، ص ص131 – 132)

3- مصادر أخطاء القياس التي تؤثر في الثبات:

تتميز درجات الاختبارات بقدر مرتفع من الثبات إذا كانت تعكس الجوانب الحقيقية للسمة أو الخاصة أو القدرة المراد قياسها، ولا تعكس عوامل الخطأ والعشوائية، ويزداد ثبات الدرجات إذا قلّت هذه العوامل. وتوجد مصادر متعددة تسهم في أخطاء القياس التي تؤدي إلى انخفاض ثبات درجات الاختبارات والمقاييس التربوية والنفسية، بعضها يتعلق بأداة القياس ذاتها، وبعضها يتعلق بإجراءات وظروف تطبيقها، وبعضها الآخر يتعلق بالأفراد المختبرين، وهي على النحو الآتي:

- **أخطاء تتعلق بأداة القياس:** فأداة القياس هي عينة من البنود، وقد لا تكون هذه العينة ممثلة لجميع البنود، وهذا ما يسمى خطأ العينة أو المعاينة Sampling Error. (ميخائيل، 2008، ص 182)، وهناك خصائص عامة ينبغي توافرها في أدوات القياس لكي تكون درجاتها متسقة بدرجة كبيرة ، فعدم ملاءمة محتوى أداة القياس للأفراد المختبرين، ومستوى صعوبة المفردات، أو غموضها وعدم دقتها، وكذلك غموض تعليمات الإجابة عليها، وأيضاً طول الاختبارات والزمن اللازم للإجابة عنها، واشتمالها على عينة غير ممثلة لمكونات السمة المراد قياسها، وغير ذلك تعد من مصادر الخطأ المتعلقة بأداة القياس، وتؤدي إلى عدم اتساق الدرجات، كما تؤثر أيضاً في صدق الاختبار، حيث إن المعلومات والقرارات التي تستند إليها تكون غير موثوق بها، أي غير صادقة.
- **أخطاء تتعلق بإجراء الاختبار وظروف تطبيقه:** يتطلب تطبيق الاختبارات أن تكون تعليمات الإجابة واضحة وكافية ومناسبة للمختبرين، كما أن ظروف تطبيق الاختبار ينبغي أن تكون جيدة ومقننة، فالاختبار الذي يطبق على مجموعة كبيرة من الأفراد في غرفة ضيقة وغير جيدة التهوية وقليلة الإضاءة وتحيط بها الضوضاء، تتأثر درجاته تأثُّراً كبيراً بهذه الظروف البيئية السيئة، مما يؤدي إلى انخفاض ثبات هذه الدرجات وعدم صدق المعلومات المستمدة منها. وكذلك تعد عملية وإجراءات تصحيح الاختبار أو تقدير درجاته وتدوين نتائجه من مصادر الأخطاء الشائعة، فمن السهل الوقوع في الأخطاء الحسابية عند جمع الدرجات أو إغفال تصحيح بعض الأسئلة، أو عدم الدقة في تصحيح الأسئلة التي تتطلب استجابات مفتوحة وغير ذلك.
- **أخطاء تتعلق بالأفراد المختبرين:** ومن أمثلتها: انخفاض الدافعية لدى المختبَرين وقت تطبيق الاختبار لأسباب عديدة، ربما ترجع إلى عدم معرفتهم بالغرض من الاختبار أو اتجاههم السلبي نحوه، أو قلة انتباههم، أو شعورهم بالإجهاد أو التعب أو الملل، أو انخفاض قدرتهم القرائية، أو حالتهم الصحية والنفسية، وأيضاً مستواهم الاجتماعي والاقتصادي، أو تفاوت تدربهم على نوع معين من الاختبارات، كما أن عدم استقرار بعض السمات الإنسانية وتذبذبها عبر الزمن يجعل من الضروري الحصول على معلومات تتعلق بهذه السمات في الوقت الذي سوف تستخدم فيه هذه المعلومات في اتخاذ قرارات بشأن الأفراد المختبرين.

لذلك ينبغي معرفة مصادر الخطأ الممكنة التي يمكن أن تؤثر في استجابات الأفراد للاختبارات والمقاييس التربوية والنفسية، مما يسهم في تحسين طرق القياس واستخدام نتائج القياس استخداماً مستبصراً، ومعرفة حدود الدقة في تفسير هذه النتائج (علام، 2006، ص ص 89- 91).

4- الانحراف المعياري والخطأ المعياري للقياس:

الانحراف المعياري Standard Deviation من أهم مقاييس التشتت، يقوم في جوهره على حساب انحرافات الدرجات عن متوسطها كما تدل عليه تسميته، وهو الجذر التربيعي لمجموع مربعات انحرافات القيم عن متوسطها مقسوماً على عددها.

يحسب وفق الصيغة الآتية:

$$S= \sqrt{\frac{\sum(\bar{x}-x)^2}{n}}$$

حيث $\bar{x}$ يشير إلى المتوسط.

و x تشير إلى الدرجة.

(درويش ورحمة، 2011، ص ص 140– 142)

أي أنَّ الانحراف المعياري هو الجذر التربيعي لمتوسط الانحرافات عن المتوسط وهو بالتالي الجذر التربيعي للتباين، ويتطلب حسابه أولاً حساب التباين للقيم ثم استخراج الجذر التربيعي للتباين فيكون الناتج هو الانحراف المعياري، فلو كان التباين لمجموعة من العلامات يساوي 64 فإن الانحراف المعياري يساوي $\sqrt{64}$ = 8 (البهي السيد، 2005، ص ص 105 – 106).

والسبب في اللجوء إلى فكرة الانحراف المعياري وجود صعوبة في الإشارات السالبة في الانحرافات عن المتوسط، وكانت الطريقة للتخلص من هذه الإشارات بتربيع القيم، وهذا هو الأساس الذي تقوم عليه فكرة الانحراف المعياري، وبذلك يعرّف الانحراف المعياري كما سبق القول بأنه الجذر التربيعي لمتوسط مربعات الانحرافات عن المتوسط، والسبب في الحصول على الجذر التربيعي لمتوسط هذه الانحرافات هو أنه تم تربيع هذه الانحرافات في أثناء العملية الحسابية، ولذلك يجب أن تعود إلى أصلها. (عيسوي، 2000، ص 41)

بعض خواص الانحراف المعياري:

- الانحراف المعياري حساس لبعد أو قرب العلامات من المتوسط، ولذلك كلما صغرت قيمته دلّ ذلك على أن طبيعة البيانات متقاربة ومتراكمة حول الوسط، وبالتالي التشتت قليل والعكس صحيح.
- لا يتأثر الانحراف المعياري بعمليتي الجمع والطرح، فإضافة مقدار ثابت للقيم الأصلية أو طرح مقدار ثابت منها لن يغير في قيمة الانحراف المعياري.
- يتأثر الانحراف المعياري بعمليتي الضرب والقسمة. (المنيزل، عايش، 2006، ص70)
- قابل للعمليات الجبرية، ولذلك فهو كثير الاستخدام في القوانين والنظريات الاحصائية.
- يُقاس الانحراف المعياري بالوحدات نفسها التي يُقاس بها المتغير، بينما يُقاس التباين بوحدات مربعة.

- يتأثر بالقيم المتطرفة، وذلك لأنه يعتمد في حسابه على المتوسط الحسابي الذي بدوره يتأثر بها.
- لا يمكن حسابه من الجداول التكرارية المفتوحة. (فهمي، 2005، ص 190)

الخطأ المعياري للقياس Standard Error of Measurement (SEM):

نظراً لأنَّ القياس النفسي والتربوي يتأثر بمصادر أخطاء متعددة، فإنه لا يوجد اختبار يقيس سمة معينة لدى الفرد قياساً متسقاً اتساقاً تاماً دون أخطاء، لذلك ينبغي تقدير قيمة هذه الأخطاء وأخذها بعين الاعتبار عند تفسير درجات الاختبار، وإذا أمكن تحديد درجة الأخطاء العشوائية التي أثرت في الدرجة الملاحظة لكل فرد من الأفراد المختبرين، فإنه يمكن إيجاد الانحراف المعياري لدرجات الخطأ، والذي يطلق عليه "الخطأ المعياري للقياس (SEM) Standard Error of Measurement" وهو مؤشر لمدى اختلاف الدرجات الملاحظة للأفراد في الاختبار عن درجاتهم الحقيقية، ويفيد هذا المؤشر في تقدير قيمة الخطأ المعياري للقياس لكل فرد في الاختبار دون الحاجة إلى تكرار تطبيق الاختبار عليه عدة مرات لتعرف مدى اتساق درجاته.

فالخطأ المعياري للقياس هو الانحراف المعياري لدرجات الخطأ، ويزود بقيمة مطلقة وليست نسبية لمدى انطباق الدرجات الملاحظة والدرجات الحقيقية، وتقديره لا يتطلب معرفة الدرجات الحقيقية.(علام، 2006، ص101)

ويمكن إيجاد قيمة الخطأ المعياري، للقياس إذا أمكن الحصول على قيمة الانحراف المعياري للدرجات الملاحظة، وكذلك قيمة معامل ثبات درجات الاختبار، باستخدام الصيغة الرياضية الآتية: $ع_{خ} = ع_{م}\sqrt{1 - ر}$

حيث $ع_{خ}$ هو قيمة الخطأ المعياري للقياس

و $ع_{م}$ هو قيمة الانحراف المعياري

وبذلك يكون الخطأ المعياري للقياس = حاصل ضرب الانحراف المعياري للدرجات الملاحظة في الجذر التربيعي لمعامل الثبات مطروحاً من الواحد الصحيح.

فعلى فرض أن قيمة معامل ثبات درجات الاختبار تساوي (0.80)، وقيمة الانحراف المعياري للدرجات الملاحظة تساوي (6.3)، فإن الخطأ المعياري للقياس =2.82 $6.3\sqrt{1 - 0.80}$

ويلاحظ أن الخطأ المعياري للقياس يساوي صفراً إذا كانت قيمة معامل الثبات مساوية للواحد الصحيح، ولكنه يصبح مساوياً للانحراف المعياري للدرجات الملاحظة إذا كانت قيمة معامل الثبات مساوية صفراً، وهذا يعني أنه إذا كانت درجات الاختبار غير متسقة على الإطلاق، فإن

تشتت الدرجات الملاحظة يعزى إلى محض الصدفة. (علام، 2002، ص ص 174- 176) ويلاحظ أيضاً أنه كلما كان الخطأ المعياري قليلاً كلما كان هناك تقارب أو تجانس أكثر بين القيم، وكلما زاد الخطأ المعياري قلّت دقة القياس ودلّ ذلك على تشتت القيم. (المنيزل، عايش، 2006، ص71)

وبعبارة أخرى فإن الخطأ المعياري للقياس يصف المدى الذي من المحتمل أن تختلف فيه درجات الفرد من اختبار لآخر، عندما يتم التطبيق في ظروف مختلفة، وهو الجذر التربيعي لتباين درجات الخطأ (Lee, 2004, PP. 410- 411).

وللخطأ المعياري للقياس عدة استعمالات منها:

- حساب ثبات أو تحديد مدى دقة كل درجة على حدة.
- حساب معامل الثبات، فالعلاقة عكسية بينه وبين الخطأ المعياري للقياس، فإذا ارتفع معامل الثبات تنخفض قيمة الخطأ المعياري للقياس.
- معرفة إذا كان أصحاب الدرجات القريبة من الدرجات الدنيا (درجة النجاح) يستحقون النجاح أم لا، فإذا حصل الطالب على الدرجة 48%، وكان الخطأ المعياري للقياس 3، فإن هذا الطالب يستحق النجاح، إذ أن درجته في أحسن الأحوال يمكن أن تصل إلى (48+3=51) (أبو لبدة، 2008، ص241).
- والخطأ المعياري للقياس يفيد في تقدير مدى تشتت الدرجات حول الدرجة الحقيقية الافتراضية في اختبار ما، كما يفيد الخطأ المعياري في تقدير قيمة الخطأ العشوائي الذي يؤدي إلى تذبذب الدرجة الملاحظة للفرد حول درجته الحقيقية. (علام، 2002، ص 177)

ومن المفيد في النهاية الإشارة إلى أن استخدام الخطأ المعياري للقياس والتعبير عن أداء الفرد بصورة زمرة من الدرجات تقع ضمنها درجته الحقيقية يقف في وجه التشدّد المفرط في اتخاذ درجة رقمية واحدة لتعبر عن أداء الفرد (ميخائيل، 2006، ص210)

5- أنواع معاملات ثبات درجات الاختبار:

يوضح الجدول الآتي أنواع معاملات ثبات درجات الاختبار، وطريقة إجرائه، وبعض مزاياه وعيوبه، ومجالات مناسبته:

جدول رقم 4: أنواع معاملات ثبات درجات الاختبار

نوع معامل الثبات	طريقة إجرائه	بعض مزاياه وعيوبه	مجالات مناسبته
معامل ثبات الاستقرار أو الثبات بالإعادة – Test retest Method	يقوم هذا المعامل على تطبيق الاختبار نفسه على الأفراد نفسهم ولكن في نقطتين زمنيتين مختلفتين، وحساب الارتباط بين درجات التطبيقين.	– مشكلة الفترة الزمنية الفاصلة بين مرتي التطبيق، فإذا كانت هذه الفترة قصيرة سوف تتأثر معاملات الثبات بعوامل التعلم والتذكر من التطبيق الأول، أما إذا كانت هذه الفترة طويلة فإن المعاملات ستتأثر بعامل النضج. – في التطبيق الثاني للاختبار قد يتغيب أي مفحوص اشتُمل في المرة الأولى، وبالتالي يتم استبعاده .	– يصلح للاختبارات الموقوتة وغير الموقوتة. –لا يصلح للاختبارات التي تهدف إلى قياس التذكر أو التي ترتبط ارتباطاً مباشراً بهذه العملية العقلية، وذلك لتأثر عملية التذكر تأثراً مباشراً بالفاصل الزمني بين التطبيقين.
طريقة التجزئة النصفية Split – half Reliability Method	يقوم هذا المعامل على إمكانية تقسيم المقياس إلى نصفين بحيث تكون بنود كل نصف مجمعة لتشكل مقياساً جديداً ، ثم تطبيق الاختبارين بالوقت نفسه، وحساب الارتباط بين درجات النصفين. وتتم التجزئة باستخدام أساليب عديدة منها: – الأسلوب الفردي والزوجي. – أسلوب جزئي الاختبار. – تحديد معامل السهولة لكل سؤال ثم توزيع الأسئلة على نصفي الاختبار بحسب هذه المعاملات.	– إن هذا النوع من معاملات الثبات يتلافى مشكلة تأثير التعلم والتذكر المرتبط بالفترة الزمنية الفاصلة بين التطبيقين. –كما يعد اقتصادياً أكثر من غيره من طرائق حساب الثبات حيث يقلل الوقت والتكاليف المادية والجهد. –إلا أنه يعاب عليه أن المصدر الرئيس للخطأ يتعلق بالأسئلة ذاتها، وهذا يعني أن عدم الثبات ينتج من الفروق في عينة الأسئلة التي يشتمل عليها كل من نصفي الاختبار، أي أن الارتباط بين النصفين سوف يختلف نسبياً اعتماداً على كيفية تقسيم البنود.	يستخدم في الاختبارات التي يكون تصميمها قائماً على توفير وحدات متجانسة بقدر الإمكان، وألا يكون الأداء عليها تراكمي بمعنى أن كل خطوة تترتب على سابقتها، وألا يكون الأداء مطلوباً خلال مدى زمني يتطلب السرعة في العمل والتوقف عنه تحت محك انقضاء الوقت.

طريقة الصور المتكافئة Parallel Forms Method	– يستلزم هذا المعامل إعداد صورتين متكافئتين من الاختبار وتطبيقهما على الأفراد أنفسهم، بحيث يتساوى عدد البنود وطريقة الإجابة والتصحيح، والزمن، إضافةً إلى تساوي معاملات السهولة والصعوبة والتمييز والتباين والمتوسط الحسابي والانحراف المعياري. –يتم تطبيق إحدى الصورتين على مجموعة من الأفراد ثم تطبق الصورة الثانية بعد فاصل زمني قصير على الأفراد أنفسهم. –تصحح الصورتان ثم يحسب معامل الارتباط بين درجات كل منهما.	– يتم بواسطته تقدير ثبات الاختبار بطوله الكامل دون حاجة إلى تنصيف أو تصحيح لأثر الطول. – كما يتميز هذا النوع بعدم التأثر بالتدريب أو الخبرة من صورة لأخرى، ويختفي أثر الألفة لاختلاف بنود الصورتين. – لكن يعاب عليه صعوبة وضع مقياس ثانٍ معادل للمقياس الأول، سواء من حيث تمثيل البنود للسمة المقيسة، ومستوى صعوبة البنود والصياغة والتعليمات والوقت المعطى للمقياس وغيرها، مما يتطلب مضاعفة الوقت والجهد.	– يصلح لحساب معامل ثبات اختبارات الذاكرة. – كذلك يصلح لحساب معاملات ثبات الاختبارات التحصيلية.
طريقة معامل ألفا Coefficient Alpha Method	يجري حساب هذا النوع من معاملات الثبات وفق المعادلة الآتية: $a = \frac{K}{K-1}\left[1 - \frac{\sum S_i^2}{S_t^2}\right]$ حيث K: عدد أسئلة الاختبار. و $\sum S_i^2$: مجموع تباين كل سؤال من أسئلة الاختبار، أي $S_1^2 + S_2^2 + \cdots + S_i^2$ و s_t^2: تباين درجة الاختبار الكلية.	– من أكثر أشكال معاملات الاتساق الداخلي شيوعاً واستخداماً. –إن قيمة معامل ألفا تساوي متوسط القيم التقديرية لمعامل ثبات كل من نصفي الاختبار لجميع طرق التجزئة النصفية الممكنة. – ويمكن تفسير معامل ألفا على أنه قيمة ارتباط البند مع جميع البنود الأخرى الممكنة التي تقيس الجانب نفسه.	يشيع استخدام هذا النوع في تقدير ثبات مقاييس الاتجاهات واستطلاع الرأي وفي مقاييس الشخصية، وكذلك في حالة الاختبارات التحصيلية.

ثبات المحكمين Inter- judge Reliability	يقوم هذا المعامل على حساب الارتباط بين الدرجات التي يعطيها محكمان أو أكثر، ويمكن أيضاً تقدير الثبات عن طريق تحديد النسبة المئوية للاتفاق بين المحكمين إذا كان الهدف مقصوراً على تصنيف الأفراد. ويمكن حسابه باستخدام الصيغة الآتية: معامل ثبات المحكمين= (عدد مرات الاتفاق بين المحكمين/ عدد مرات الاتفاق+ عدد مرات الاختلاف) × 100	لا يكفي معامل ثبات المحكمين وحده لتحديد التباين الحقيقي للقدرة التي يقيسها الاختبار، بل يتعين حساب الثبات أيضاً بأحد الأساليب المناسبة للاختبار بالإضافة إلى حساب ثبات المحكمين.	– في الاختبارات التي تعتمد عملية إعطاء الدرجات فيها على الأحكام التي يصدرها المحكمون. – في اختبارات الانتقاء والتصنيف. – وكذلك في الاختبارات المقالية التي تُصحح من أكثر من مصحح واحد.

(Drost, 2011, P.110, Garson, 2009, PP.3-8, Lee, 2004, P.396, البهي السيد، 2005، ص 381، النبهان، 2004، ص ص 248– 249، فرج، 2007، ص ص 314– 340، علام، 2006، ص ص98 – 96)

6- العلاقة بين طول الاختبار وثباته:

بشكل عام تعطي الاختبارات الأكثر طولاً درجات أكثر ثباتاً، ذلك لأن الأخطاء العشوائية الموجبة والسالبة في هذه الاختبارات ستحظى بفرصة أن يلغي بعضها الآخر، وبذلك تقترب الدرجة الملاحظة من الدرجة الحقيقية، وبالطريقة نفسها تعطي الاختبارات القصيرة أو الفرعية درجات أقل ثباتاً، وهذا بطبيعة الحال ينطبق على كافة المقاييس، ذلك لأن العدد الأكبر من البنود عبارة عن عينة أكبر من السلوك، وكلما كانت العينة أكبر كان من المتوقع أن يمثل فيها بشكل مستقر العدد الأكبر من مكونات السلوك أو القدرة المقيسة، والقابل للظهور في مرتي التطبيق أو نصفي أو جزأي الاختبار. (النبهان، 2004، ص ص 255– 256)

ووفقاً لمعادلة سبيرمان – براون يتحدد طول الاختبار على الوجه الآتي:

$$\text{ن} = \frac{\text{ر ن ن}(1-\text{ر ت ت})}{\text{ر ت ت}(1-\text{ر ن ن})}$$

حيث ن = مضاعفات الطول المطلوبة للاختبار للوصول به إلى معامل ثبات معين

ر ت ت = معامل الثبات الحالي للاختبار قبل رفعه.

ر ن ن = معامل الثبات المطلوب للاختبار

فإذا كان معامل ثبات الاختبار الحالي 0.60 مثلاً، والمطلوب رفعه ليصل إلى 0.80، فإنه يجب التعويض في المعادلة الآتية لتحديد مضاعفات الطول المطلوبة للاختبار، كالآتي:

ن = $\frac{(0,6-1)0,8}{(0,8-1)0,6}$ = 2.7 أي أن المطلوب مضاعفة عدد بنود الاختبار إلى مرتين وسبعة من عشرة مرة للوصول إلى معامل ثبات قدره (0.80). (فرج، 2007، ص ص347- 348)

لكن لابد من مراعاة شرطين عند زيادة عدد بنود الاختبار، أولهما أن يكون للأسئلة المضافة الخواص الإحصائية نفسها التي تتصف بها أسئلة الاختبار، أي أن تكون متجانسة معها، والشرط الثاني ألا يؤدي تطويل الاختبار إلى تعب أو ملل المفحوص ومن ثم التأثير في أدائه. (أبو لبدة، 1985، ص274)

7- عوامل أخرى تؤثر في معامل ثبات الاختبار:

يتأثر معامل ثبات أداة القياس بعدد من العوامل إضافةً إلى طول الاختبار، منها ما يتعلق بأداة القياس نفسها، ومنها ما يتعلق بخصائص الأفراد الذين يطبق عليهم الاختبار، ومنها ما يتعلق بظروف التطبيق والتصحيح، وفيما يلي بعض هذه العوامل:

- **الموضوعية في التصحيح:** إن الفقرات التي تصحح بموضوعية عالية تؤدي إلى تحسين ورفع تقدير معامل الثبات. لهذا فإن ثبات الاختبار الذي يتألف من فقرات موضوعية أعلى عموماً من ثبات الاختبار الذي يتألف من فقرات إنشائية، شريطة تساوي العوامل المؤثرة الأخرى (النبهان، 2004، ص 258)، والسبب في ذلك أنه لكي يعطي المقياس نتائج ثابتة يجب ألا تختلف الدرجات من مصحح لآخر أو من وقت لآخر، ومن المعلوم أن الاختبارات الموضوعية عموماً بخلاف الاختبارات المقالية لا تفسح المجال لتدخل العوامل الذاتية للمصحح (ميخائيل، 2008، ص211)
- **زمن الاختبار:** يتأثر ثبات الاختبارات الموقوتة بالزمن المحدد لها، حيث يزداد الثبات تبعاً لزيادة الزمن حتى يصل إلى الحد المناسب للاختبار، فيصل الثبات إلى نهايته العظمى، ثم يقل الثبات بعد ذلك كلما زاد الزمن عن ذلك الحد.
- **التخمين:** ينقص الثبات تبعاً لزيادة التخمين، ذلك لأن الإجابة التي تعتمد على التخمين في المرة الأولى لإجراء الاختبار لا تعتمد على هذا التخمين نفسه في المرة الثانية على المجموعة نفسها، وبذلك تضعف الصلة بين نتائج المرة الأولى ونتائج المرة الثانية، وتنخفض تبعاً لذلك قيمة معامل الثبات، وهكذا يؤثر الغش والتخمين تأثيراً ضاراً على ثبات الاختبار.

• **صياغة الأسئلة:** الأسئلة الغامضة، الخادعة، العاطفية، تقلل الثبات، والأسئلة واضحة المبنى، الموضوعية، تزيد الثبات، ولذا يجب أن يدقق الباحث في اختيار ألفاظ أسئلته ونوعها، حتى يصل بذلك إلى الثبات الحقيقي للاختبار. (البهي السيد، 2005، ص ص394- 396)

• **تباين الدرجات (انتشار الدرجات):** يزداد ثبات الاختبارات تبعاً لزيادة التباين في الدرجات، ويرجع ذلك إلى أن معامل الثبات هو نسبة التباين الحقيقي إلى التباين الكلي، كما أن زيادة التباين يدل على قياس مدى متسع من السمة موضع القياس، ومعنى ذلك أن الأسئلة منخفضة التباين (السهلة أو الصعبة) تؤدي إلى خفض معامل الثبات، بينما الأسئلة مرتفعة التباين (متوسطة السهولة) تؤدي إلى زيادة معامل الثبات. (مراد وسليمان، 2002، ص369)، والسبب في ذلك أن الفقرات السهلة جداً أو الصعبة جداً لا تبرز الفروق الفردية، وهي بالتالي لا تظهر التباين بين درجات المفحوصين. الأمر الذي يؤدي الى جعل معامل الثبات منخفضاً. (النبهان، 2004، ص 258)

كما أنَّ زيادة صعوبة الاختبار وصعوبة مفرداته يجعل الأفراد يلجؤون إلى التخمين العشوائي في إجاباتهم مما يزيد من الأخطاء العشوائية ويقلل من ثبات درجات الاختبار. (علام، 2006، ص104)

• **خصائص العينة:** يختلف معامل الثبات بين الجماعات التي يتفاوت متوسط مستويات القدرة فيها، وهذه الفروق لا يمكن التنبؤ بها أو تقديرها بأي معادلة إحصائية، ولكن يمكن اكتشافها وتحديدها بتجريب الاختبار على جماعات تختلف في العمر أو في مستوى القدرة، وعموماً معاملات الثبات قد تكون منخفضة نسبياً في الجماعات الأصغر عمراً أو الأقل قدرةً، لأن هذه الدرجات تتأثر في هذه الجماعات بالتخمين، كما أن كل معامل الثبات يجب أن يصحبه وصف دقيق لنوع المجموعة التي حُسِب من درجات أفرادها، ولا يكون معامل الثبات الذي يتضمنه دليل الاختبار صالحاً إلا للجماعات المشابهة لعينة التقنين، ولذلك يُفضَّل أن يتم تقسيم عينة التقنين هذه إلى مجموعات فرعية أكثر تجانساً بالنسبة لعوامل الجنس والعمر والصف الدراسي والمهنة وغيرها، ثم حساب معاملات ثبات منفصلة لكل من هذه المجموعات، وفي هذه الحالة يمكن الاستفادة من معاملات الثبات التي حسبت من عينات التقنين حين نستخدم الاختبار في الأغراض التطبيقية والعملية (أبو حطب، 2003، ص 127).

8- مفهوم الصدق وأهميته:

يشير مصطلح الصدق إلى كون الاختبار يقيس فعلاً ما أُعدَّ لقياسه، حيث يُقصد بصدق الاختبار أن تكون هناك دلائل مقبولة حول ما يقيسه وكيفية قياسه، وعلى الرغم من أن كل اختبار يحمل مُسمى يشير إلى ما يقيسه إلا أن هذا المُسمى غالباً ما يكون عاماً وغير محدد، كما أنه لا يُعد – في حد ذاته- دليلاً على ما يقيسه الاختبار، حيث يفيد كأداة تعريفية فقط

للاختبار، لذلك من الضروري التحقق من صدق الاختبار بأساليب منهجية محددة قبل الشروع في الاعتماد على النتائج المستخلصة منه، ولا يُعد صدق الاختبار خاصية مطلقة- بمعنى أن الاختبار إما أن يكون صادقاً أو غير صادق- بل يُعد خاصية نسبية، فبدلاً من القول: إن الاختبار صادق أو غير صادق، يجب القول بأنه صادقٌ بقدرٍ ما، وهذا يعني أن الصدق تعبير عن درجة، وأنه ليس ثمة اختبار صادق بصورة كلية أو مطلقة، كما يختلف صدق الاختبار نفسه من مجتمع إلى آخر نتيجة للفروق الحضارية التي تسبغ على السلوك دلالات مختلفة، كما قد يختلف صدق الاختبار أيضاً بين فترة زمنية وأخرى، فنتيجة للتقدم في التعليم والتطور الاجتماعي والثورة التكنولوجية الاتصالية، أصبحت البنود التي كانت صادقة على مقياس الذكاء في عشرينيات وثلاثينيات القرن الماضي لا تصلح الآن ولا يمكن استمرار وجودها في مقياس للذكاء في القرن الواحد والعشرين. (فرج، 2007، ص ص 238 – 239)

ويمكن توضيح مفهوم الصدق بالنقاط الآتية:

- يشير الصدق إلى تفسير نتائج الاختبار وليس الاختبار نفسه.
- يُستنتج الصدق من الأدلة المتوافرة وليست المناسبة.
- الصدق خاص باستخدام درجات الاختبار، مثل استخدام الدرجات في اختيار الأفراد أو توزيعهم أو تقويم الأداء.
- يُعبَّر عن الصدق بدرجة وصفية مثل مرتفع أو متوسط أو منخفض، وقد تُستخدم الأرقام لتوضيح الدرجة الوصفية. (مراد وسليمان، 2002، ص 350)

ومن التعريفات المهمة للصدق:

- تعريف تايلر (Tyler) للصدق بأنه: مدى قياس الاختبار لما وضع لقياسه، وإن أهم سؤال يتصل بالصدق هو: ماذا يقيس الاختبار؟

وإذا لم تكن هنالك إجابة كافية بدرجة معقولة لهذا السؤال، فإن الاختبار سوف يكون عديم الفائدة. (تايلر، 1989، ص ص 50- 51)

- تعريف ليندكوست (Lindquit,1942) للصدق بأنه: درجة الصحة التي تُقاس بها السمة المراد قياسها، أو أنه الدرجة التي يتم الاقتراب منها بنجاح تام لقياس السمة أو الخاصية موضع القياس.
- ويذكر كرونباخ (Cronbach, 1960) أنه بقدر اكتمال تفسير درجة الاختبار للسمة المعينة والثقة في هذا التفسير، بقدر ما يتحقق صدق الاختبار، وهو يربط بذلك بين الدرجة على الاختبار وقدرتها التفسيرية (الأنصاري، 2000، ص ص 91- 92).

9- أنواع معاملات صدق الاختبار:

أ- **صدق المحتوىContent Validity:**

يرتكز صدق المحتوى أو الصدق التمثيلي على الفحص المنظم لمحتوى الاختبار، لتحديد ما إذا كان عينة ممثلة لمحتوى الموضوع ومجال السلوك الذي يُراد قياسه، ويتطلب هذا النوع من أنواع الصدق تحديداً ووصفاً دقيقاً للمحتوى أو المجال المقيس، وتحليله إلى عناصره ومكوناته حتى يتم التأكُّد من أن البنود الموضوعة تغطي الجوانب الأساسية لهذا المحتوى وبالنسب الصحيحة، دون أن يطغى أي من هذه الجوانب على غيره أو يُعطى وزناً أكبر من الوزن الذي يستحقه.

ليس ثمة صيغة كمية يمكن استخدامها للتحقق من صدق المحتوى أو تعبير عددي يمكن استعماله للدلالة عليه، فهذا النوع من أنواع الصدق يعتمد على التحليل المنطقي والمقارنة أو المطابقة بين محتوى الاختبار ومحتوى الموضوع المقيس، ولذلك يُشار إليه في الكثير من الأحيان بالصدق المنطقي، وصدق المحتوى بحد ذاته لا يكفي لضمان الصدق وخاصة في المقاييس النفسية، حيث يصعب القول إن هذه الاختبارات صادقة من مجرد فحص محتواها دون الاعتماد على محكات عملية وتجريبية تُستخدم في التثبت من صحة الفروض التي اعتمدت أساساً في بنائها، وتقول آنا أنستازي في هذا الصدد: مع أنّ صدق المحتوى يجب أن يدخل في المراحل الأولية لبناء أي اختبار، إلا أنّ تأسيس الصدق النهائي لدرجات الاختبار يتطلب التحقق التجريبي بإجراءات أخرى (ميخائيل، 2008، ص ص144- 146).

وهكذا فإن صدق المحتوى يُعد من أنواع الصدق الكيفي أو النوعي، حيث يحدد مجال المفهوم أو السمة موضع القياس بوضوح، وبالتحليل الاحصائي يمكن الحكم ما إذا كانت المقاييس تمثل بشكل كامل المجال أو السمة موضع القياس (Drost, 2011, P. 118).

ب- **الصدق الظاهري أو السطحي Face Validity:**

يُستخدم هذا النوع من الصدق للتأكد أن المقياس يبدو بنظر المفحوصين قادراً على قياس السمة أو الخاصية موضع القياس، وعلى الرغم أن هذا النوع من الصدق ليس علمياً بالمعنى الدقيق للكلمة، إلا أنه عنصر أساسي لإثارة الدافعية لدى المفحوصين، فإذا كان المفحوصون لا يعتقدون أن هذا المقياس يقيس بدقة القدرة أو أنه غير مرتبط بالسمة موضع القياس، فمن الممكن أن لا يعيروه اهتماماً وبالتالي يفشلوا عملية القياس (Phelan & Wren, 2005, P.3)

ومما لاشك فيه أن الاختبار لابد أن يبدو صادقاً كي يُقبل المفحوصون عليه برغبة وتستثار دافعيتهم لأدائه بالحد الأقصى، وتؤكد آنا أنستازي أنه لابد من توفير صفة الصدق الظاهري للاختبار حتى يكون أكثر فاعلية في المواقف العملية، ولضمان تعاون المفحوصين في الموقف الاختباري، ويمكن تحسين الصدق الظاهري بنظر أنستازي بمجرد إعادة صياغة الأسئلة بحيث

تبدو أكثر ارتباطاً بموقف القياس الفعلي، ومهما يكن من أمر فإن الصدق الظاهري للاختبار لابد أن يكون في خدمة الصدق الحقيقي لهذا الاختبار وليس بديلاً عنه، ولابد تغليب الصدق الحقيقي حين يتعذر الجمع بينهما. (ميخائيل، 2008، ص ص146-147)

ت- **صدق المحك Criterion Validity:**

يرتكز هذا النوع من الصدق على دراسة الارتباط بين درجات المقياس ودرجات مقياس آخر يُعتمد كمحك له، ويُقصد بالمحك: مقياس مباشر ومستقل لما يهدف الاختبار إلى قياسه، والتنبؤ بـه أو تشخيصـه، أو هـو ميـزان لتحديـد مـدى صـلاحية الاختبـار، أي هـو اختبـار للاختبـار (أبو حطب، 2003، ص 148)

ويمكن تعريف الصدق المحكي على أنه الدرجة التي يترابط عندها الأداء على مقياس مع الأداء على مقياس آخر عُدَّ محكاً للمقياس الأول وأساساً في الحكم على صلاحيته، ويتطلب هذا النوع من الصدق إجراءات تجريبية إحصائية (ميخائيل، 2008، ص ص 147- 148).

وللصدق المحكي أو الصدق التجريبي طرائق حساب عديدة منها:

- **الصدق التلازمي Concurrent Validity:**

يشير الصدق التلازمي إلى قدرة المقياس على التنبؤ بحدوث السلوك بالوقت الحالي، ويكون الصـــدق تلازميـــاً عنـــدما يطبـــق المحـــك فـــي الوقـــت نفســـه مـــع المقيـــاس (Drost, 2011, PP. 118- 119).

حيث يدل هذا النوع من الصدق على حجم العلاقة بين درجات الأفراد على الاختبار ودرجاتهم على محك آخر بحيث لا يكون هناك فاصل زمني، أو يكون هذا الفاصل قصير بين أداء الأفراد على الاختبار وأدائهم على المحك، ويطلق البعض على الصدق التلازمي اسم الصدق التجريبي لأن حسابه يعتمد على تطبيق اختبارات وقياس أداء فعلي عملي، ثم حساب معامـــل الارتبـــاط بـــين الـــدرجات علـــى الاختبـــار والـــدرجات علـــى المحـــك. (مراد وسليمان، 2002، ص353)

- **الصدق التنبؤي Predictive Validity:**

عندما يتحقق المحك أو المعيار في المستقبل فإن الصدق يكون تنبؤياً، فالصدق التنبؤي يشير إلى قدرة المقياس على قياس السلوك والتنبؤ به في المستقبل (Drost, 2011, P. 119). يهتم هذا النوع من الصدق باستخدام درجات الاختبار في التنبؤ بالأداء المستقبلي على مقاييس أخرى هي المحكات، فقد تستخدم درجات اختبار الاستعداد للنجاح في المدرسة بالتنبؤ بدرجات الطلبة في مواد دراسية معينة، ومن أمثلتها اختبار الاستعداد للنجاح في الجامعة أو اختبارات القبول في الجامعة.

الفرق بين الصدق التلازمي والصدق التنبؤي:

- تستخدم الطريقتان في حساب الصدق محكاً خارجياً، لدراسة حجم الارتباط بين درجات الاختبار والمحك، إلا أن الفرق ينحصر في الفارق الزمني، ففي حالة الصدق التلازمي يتم الحصول على درجات الاختبار والمحك في الوقت نفسه تقريباً، أما في حالة الصدق التنبؤي فيتم الحصول على درجات الاختبار أولاً وبعد فاصل زمني لا يقل عن ستة شهور يتم الحصول على درجات المحك.

- الصدق التنبؤي يحاول قياس صلاحية درجات الاختبار في التنبؤ بدرجات الأفراد في المستقبل، ومثال ذلك درجات التحصيل في الثانوية العامة كمؤشر للنجاح في الجامعة، أما الصدق التلازمي فيحاول قياس صلاحية درجات الاختبار في استخدامها بدلاً من درجات المحك الذي قد يكون من الصعب الحصول عليها بدقة وفي وقت قصير، ومثال على ذلك استخدام مقياس لتقدير خصائص الشخصية للفرد بدلاً من إجراء المقابلات الشخصية التي تستغرق وقتاً وجهداً كبيراً. (مراد وسليمان، 2002، ص ص353- 354)

ث-**<u>صدق التكوين الفرضي Construct Validity:</u>**

يُستخدم هذا النوع من الصدق للتأكد من أن المقياس يقيس في الواقع ما يهدف لقياسه (أي ما هو مبني على أساسه) ولا يقيس متغيرات أخرى، ويُبنى عن طريق لجنة من الخبراء ممن هم على دراية بالطريقة التي يُبنى على أساسها هذا النوع من الصدق، حيث يقوم هؤلاء الخبراء بفحص بنود المقياس لتقرير ما يقيس كل بند من بنوده تماماً، كما يمكن للطلاب المشاركة في هذه العملية للحصول على ملاحظاتهم (Phelan and Wren, 2005, P.4)

وتضع لوفنجر (Loevinger, 1979) صدق التكوين الفرضي في إطارين هما:

- إلى أي حد يقيس الاختبار خاصية أو سمة لها وجود فعلي؟
- إلى أي حد يكون هناك تناظر بين التفسير المقترح للسمة أو الخاصية وما يقيسه الاختبار فعلاً؟

ويتصل الإطار الأول بصدق الاختبار، في حين يعرف الإطار الثاني بصدق التفسير، ولاشك أن التساؤلين يكمل أحدهما الآخر، إذ يقدّم الإطار الأول أدلة حول مقدار ونوع الارتباطات بين فقرات الاختبار، وهذا يضفي معنى سيكومتري على الاختبار، في حين يقدم الإطار الثاني معنى عملياً أو مهنياً - سيكولوجياً كان أم تربوياً- للاختبار (النبهان، 2004، ص ص 294 – 295).

وهكذا يشير هذا النوع من الصدق إلى الدرجة التي يقيس معها الاختبار الصفة أو الخاصية التي يفترض أن يقيسها، وينطلق من أن الدرجات على الاختبار يجب أن تتنوع أو تتباين كما تتنبأ النظرية الخاصة بالسمة المقيسة أو المفهوم أو التكوين المفترض، والصدق البنيوي هو

مفهوم شامل يتضمن الأنواع الأخرى من الصدق، ويُعد بالتالي شرطاً مهماً للاختبارات النفسية والتربوية بأنواعها، ويتطلب هذا النوع من الصدق اللجوء إلى الاستنتاجات المنطقية بالإضافة إلى الوسائل التجريبية والاحصائية، والتقنيات والإجراءات الخاصة بتأسيس صدق المحتوى والصدق المحكي يمكن أن تندرج ضمن تقنيات وإجراءات الصدق البنيوي. (ميخائيل،2008، ص ص162- 163)

وفيما يلي بعض الأساليب المستخدمة في تأسيس صدق التكوين الفرضي:

- **الصدق العاملي Factorial Validity:**

الصدق العاملي من أهم الوسائل التي يتم اللجوء إليها في تحديد صدق التكوين الفرضي، وهو في جوهره أسلوب لتحديد بنية السمات النفسية، ويقوم على تحليل العلاقات بين البيانات الإمبريقية، وتتمثل هذه العلاقات بصورة معاملات ارتباط (أبو حطب، 2003، ص 163).

فالهدف الرئيسي من الصدق العاملي هو تبسيط عملية وصف السلوك، عن طريق خفض عدد المتغيرات التي تقيسها الاختبارات إلى عدد ضئيل من العوامل أو السمات المشتركة، وعلى هذا الأساس يمكن وصف كل اختبار بالاعتماد على العوامل الرئيسية المحددة لدرجاته، إضافة إلى تشبعه بكل من هذه العوامل وارتباطه بها. (ميخائيل،2008، ص162- 163)

وبعبارة أخرى يهتم الصدق العاملي بتعرف مدى تشبع بنود الاختبار بعوامل معينة سواء أكانت عاملاً عاماً أم عوامل طائفية، ويستخدم الصدق العاملي طريقة التحليل العاملي لمصفوفة ارتباطات بين درجات بنود الاختبار، في محاولة لإنقاص عدد العوامل أو مكونات السمة التي تتجمع حولها بنود الاختبار. (مراد وسليمان، 2002، ص355)

- **الصدق التجريبي Experimental Validity:**

يتم التحقق من الصدق تجريبياً عن طريق حساب معاملات ارتباط درجات الاختبار بدرجات أدوات أخرى (محكات) تقيس السمة نفسها أو تكون السمة مستخدمة فيها، وتسمى هذه الطريقة بالصدق التجريبي، وهي طريقة أيضاً لحساب الصدق التلازمي والصدق التنبؤي، وبمعنى آخر فإن الصدق التجريبي هو طريقة تستخدم لحساب صدق المحك (تلازمي أو تنبؤي) إضافةً إلى صدق التكوين الفرضي. (مراد وسليمان، 2002، ص356)

- **طريقة المقارنات الطرفية أو الفرق المتقابلة Comparison of Extreme groups:**

تعتمد هذه الطريقة على دراسة الفروق بين مجموعات مختلفة من الأفراد، وتهدف كثير من هذه الدراسات للمقارنة بين مجموعتين مختلفتين في ضوء الدرجات التي يحصل عليها الأفراد في اختبار أو مقياس معين، حيث من المتوقع أن المجموعتين سوف تختلف استجاباتهما أو الدرجات التي يحصل عليها أفراد كل منهما في الاختبار أو المقياس، ويمكن التحقق من ذلك بطريقة مباشرة، حيث للتحقق مثلاً من صدق التكوين الفرضي لاختبار يقيس العصابية

Neuroticism، يمكن اختيار مجموعتين من الأفراد استناداً إلى تشخيص الأطباء النفسيين بحيث تكون هذه السمة مرتفعة لدى إحدى المجموعتين ومنخفضة لدى المجموعة الأخرى، ويطبق الاختبار على كل من المجموعتين، فإذا تبين أن هناك اختلافاً واضحاً بين درجات كل من المجموعتين، فإنه يمكن اعتبار ذلك أحد الأدلة على صدق التكوين الفرضي للاختبار، وتُستبقى عندئذٍ المفردات التي ميّزت بدرجة أكبر بين المجموعتين (علام، 2002، ص222)

10- العوامل المؤثرة في صدق الاختبار:

ثمة عوامل عديدة يمكن أن تؤثر سلباً في صدق الاختبار وتضعفه. وقد صنف ن. جرونلند (Gronlund, 1971) هذه العوامل ضمن أربع فئات رئيسية على النحو الآتي:

- **عوامل تتعلق بالاختبار نفسه ومن أهمها:**

– التعليمات غير الواضحة: فالتعليمات التي لا تشير بوضوح إلى ما يجب أن يفعله المفحوص، وكيف يجيب، وكيف يسجل إجاباته، تضعف الصدق.

– المفردات والتراكيب الصعبة: فهي قد تجعل الاختبار يقيس الفهم القرائي، أو مظاهر من الذكاء وليس مظاهر السلوك التي أُعد لقياسها.

– البنود الصعبة جداً أو السهلة جداً: إذ أن البنود الصعبة جداً وكذلك السهلة جداً لا تميز بين الأقوياء والضعاف تحصيلاً، وبالتالي فهي تضعف الصدق.

– البنود الموحية بالإجابة: فمثل هذه البنود قد تقيس القدرة على اكتشاف الإجابات الصحيحة اعتماداً على الإيحاءات أو التلميحات التي تنطوي عليها البنود ذاتها، ولا تكشف بالتالي الفروق في السمة أو الخاصية موضع القياس.

– عدد البنود: فالاختبار هو عينة من بنود أو أسئلة كثيرة يمكن أن تطرح، وإذا كان قصيراً وتضمن عدداً ضئيلاً من البنود يضعف تمثيله وبالتالي يضعف صدقه.

– الغموض: حيث يؤدي إلى سوء التفسير، وقد يضلل المفحوصين الأقوياء أكثر من الضعاف ويجعل الاختبار يميز بالاتجاه المعاكس.

– ترتيب البنود بصورة غير ملائمة: يرتب الاختبار عادةً وفق مبدأ التدرج في الصعوبة، فإذا وضعت البنود الصعبة في البداية يقضي الطلاب الكثير من الوقت للإجابة عنها، وقد لا يتسع الوقت للإجابة عن البنود السهلة، ثم إن هذا الترتيب قد يضعف دافعيتهم مما ينعكس سلباً على أدائهم.

– تخصيص مكان واحد للإجابة الصحيحة: فإذا أخذ الجواب الصحيح الرقم (3) أو الرمز (ج) مثلاً في أكثر من مرة يسهل اكتشافه، مما يؤدي إلى إضعاف الصدق بطبيعة الحال.

- **العوامل المتصلة بشروط الإجراء والتصحيح ومنها:**

– الشروط البيئية المحيطة: فالحرارة والبرودة الزائدة والضوضاء وغيرها، هي من العوامل التي يمكن أن تؤثر سلباً في الأداء الاختباري.

– الوقت المخصص للإجابة: فإذا لم يكن كافياً أو كان طويلاً أكثر من اللزوم يتأثر الأداء الاختباري بالضغط على المفحوصين في الحالة الأولى وإرباكهم، وإتاحة الفرصة للغش والفوضى في الحالة الثانية.

– أخطاء التصحيح: وقد تنتج هذه الأخطاء عن عملية التقدير ذاتها كما في الأسئلة المقالية التي تعاني من عيوب التقدير الذاتي وضعف مستوى الثبات، وقد تنتج عن عملية جمع العلامات أو نقلها وما شابه ذلك.

– الغش: فإتاحة الفرصة للغش أو تقديم المساعدة للمفحوصين يحولان دون الكشف عن الفروق الحقيقية بينهم، ويضعف الصدق.

- **العوامل المتصلة باستجابات المفحوصين:**

قد يتعرض المفحوصون لأزمات انفعالية حادة تؤثر في أدائهم الاختباري، حيث يعاني بعض المفحوصين مما يسمى قلق الاختبار، كما أن بعضهم قد تضعف دافعيتهم ولا يبذلون جهودهم على النحو المطلوب. (ميخائيل، 2008، ص ص 175– 178).

11- العلاقة بين الصدق والثبات:

يُعدّ الصدق والثبات أهم سمتين للاختبار الجيد، ومن الثابت أن العلاقة بين صدق الاختبار وثباته هي علاقة أكيدة، خاصةً أن كلا المفهومين يبحثان في مدى كفاءة الاختبار ومناسبته للمسلمات الرئيسية لنظرية القياس.

وفي حين يبحث مفهوم الثبات في مدى استقرار درجات الاختبار عندما تتغير الظروف الخارجية، أي يختص بالاختبار ودرجاته، يتجاوز مفهوم الصدق الاختبار ودرجاته إلى محك خارجي من أجل تعيين معامل صدق الاختبار، سواء بصورة بسيطة مباشرة أي بحساب معامل الارتباط بين الاختبار والمحك أو المقارنة الطرفية، أو بصورة أكثر تعقيداً عندما يستخدم منهج التحليل العاملي للوقوف على صدق الاختبار في ضوء تشبّعه بالعوامل التي يقيسها.

والاختبار الثابت هو اختبار أيضاً عالي الصدق من الناحية النظرية، ولكن قد يكون غير ذلك تماماً من الناحية العملية التطبيقية، أما الاختبار الصادق فلا بد أن يكون اختباراً ثابتاً من الناحيتين النظرية والتطبيقية. (عبد الرحمن، 1998، ص198)

وهكذا فإن الثبات خاصية تتوفر في كل اختبار صادق، بينما الصدق لا يتوافر بالضرورة لكل اختبار ثابت، حيث يمكن القول بأن كل اختبار صادق ثابت بالضرورة ولكن ليس كل اختبار ثابت صادق بالضرورة.

فإذا كان الاختبار صادقاً، أي يقيس ما أُعد لقياسه بالفعل، فإن الدرجة عليه ستكون معبرة عن الأداء الحقيقي أو القدرة أو الوظيفة الفعلية كائنة ما تكون هذه الوظيفة أو القدرة النفسية، وما دامت الدرجة على المقياس الصادق تعبر عن هذه الوظيفة بدقة، يمكن توقع أن تكون هذه الدرجة ثابتة في الوقت نفسه أي متسقة في تعبيرها عن مقدار الوظيفة، ومستقرة عبر الزمن في تعبيرها وتقديرها لهذه الوظيفة. (فرج، 2007، ص295)

ولابد من الإشارة في هذا السياق إلى أن الفرق بين الصدق والثبات هو أن الثبات مفهوم إحصائي، بخلاف الصدق الذي يعتمد على التحليل المنطقي كما يتطلب العمل الإحصائي أيضاً. (ميخائيل، 2008، ص184)

غير أنّ كون الصدق مفهوم أشمل من الثبات، لا يعني أنه يمكن الاكتفاء بتقديرات الصدق الخاصة باختبار معين باعتبارها تتضمن بالضرورة تقدير ثباته، فحتى المرحلة الراهنة من التقدم العلمي لا تتوافر مقاييس نفسية ذات صدق تام، كما أن صدق المقياس يخضع لاعتبارات عملية متعددة تختلف من حالة إلى أخرى ومن مجتمع إلى آخر، وبالتالي يتعين حساب وتقدير كل من صدق وثبات المقاييس باستقلال كل منهما عن الآخر، قبل استخلاص نتائج تجريبية منها تتعلق بمجتمع معين أو عينة معينة (فرج، 2007، ص295).

12- مشكلة ثبات اختبارات الشخصية:

أظهرت الدراسات أن بعض معاملات ثبات اختبارات الشخصية منخفض، وبعضها مرتفع، و غالبيتها متوسط، مع ملاحظة أن ثبات الاتساق الداخلي أعلى غالباً من ثبات الاستقرار عبر الزمن.

ومن المُلاحظ أن معاملات ثبات الاستخبار الواحد تختلف باختلاف نوع الثبات المستخدم، حيث لاحظ "مولار" Muller " أن معاملات ثبات التنصيف مرتفعة بوجه عام، أما معاملات ثبات الاستقرار أي بتطبيق الاختبار وإعادته فهي منخفضة بوجه عام. ويفسر ذلك أنه في حين أن استجابات المفحوص لمختلف بنود الاستخبارات متسقة بدرجة مقبولة في التطبيق الواحد، فإنها تتذبذب بدرجة كبيرة من تطبيق إلى آخر حيث يفصلهما فترة من الزمن، وهنا لابد من الإشارة إلى أن ثبات الاستقرار أهم من ثبات التنصيف، ففي حين يشير الأول إلى استقرار حقيقي و ثبات للسمة المقيسة، فقد يشير ثبات التنصيف إلى مجرد اتساق سطحي أو شكلي.

ويمكن إجمال أهم أسباب انخفاض ثبات اختبارات الشخصية على النحو الآتي:

1- مشكلة التزييف في اختبارات الشخصية، والتي قد تكون إلى الأسوأ أو إلى الأحسن أي تشمل كلا النوعين .

2- دافعية المفحوص نحو الاختبار.

3- ظروف التطبيق الفيزيقية.

4- الحالة النفسية الراهنة للمفحوص.

5- نوعية وعمومية السمة المقاسة، حيث افترض عدد من علماء النفس انخفاض ثبات سمات الشخصية نتيجة لما لها من نوعية موقفية، أي أن السلوك الذي يصدر عن الفرد قد يكون سلوكاً نوعياً خاصاً وليس عاماً، أي أن محدداته تكمن في الموقف الخاص الذي يصدر فيه (الأنصاري، 2000، ص ص135- 137).

وقد اقترح جيلفورد Guilford (1954) اتباع خطوة عامة لتفسير قوة العلاقة لمعاملات الثبات وفق الجدول الآتي:

الجدول رقم 5: تفسير قوة العلاقة لمعاملات الثبات

معامل الارتباط	تفسير العلاقة
أقل من 20%	العلاقة ضعيفة الأهمية – الارتباط غير جدير بالاهتمام
من 20% إلى 40%	الارتباط منخفض – واقع العلاقة صغير
من 40% إلى 70%	الارتباط معتدل – العلاقة أساسية
من 70% إلى 90%	الارتباط مرتفع – العلاقة قوية
من 90% إلى 1	الارتباط عالٍ جداً – العلاقة معتمد عليها جداً

(Crask, 1999, P.3)

غير أن هذه المشكلات التي تواجه قياس الشخصية لا تعني الابتعاد عن قياسها، فقد قدّم علماء النفس العديد من الحلول لمعالجة هذه المشكلات وخاصة مشكلة التزييف، حيث يوصي ألبورت (Allport, 1937) باستثارة دوافع الأمانة والصراحة لدى المفحوص للتجربة، كما ينصح كرونباخ (Cronbach, 1960) بتكوين علاقة تعاون مع المفحوص واستخدام الباحث مهاراته في تكوين علاقة ودية، وتذكر أنستازي (Anastasi, 1988) أن تعليمات الاختبار وتكوين علاقة ودّية يمكن أن تدفع الشخص إلى أن يستجيب بصراحة، وذلك إذا أمكن إقناعه أن مصلحته الشخصية في أن يفعل ذلك، وتضيف أن وضع بنود خفيّة نسبياً أو محايدة اجتماعياً يمكن أن يقلل من التزييف، ولو أن جاكسون Jaxon يذكر أن ما يُسمى بالبنود الخفيّة ربما تكون ذات صدق منخفض بالنسبة للبعد الذي نهتم بقياسه، ويذكر الأستاذ الدكتور أحمد عزت راجح أنه للتحوط من خداع المفحوص في الإجابة يمكن تكرار السؤال ذاته في أجزاء مختلفة من الرائز مع اختلاف بسيط في صيغته، ومقارنة أجوبة المفحوص عن هذه الأسئلة المتشابهة والمتباينة في آنٍ واحد، أما مولار (Muller, 1944) فيرى أنه بدلاً من صياغة العبارات في صورة أسئلة مباشرة تبدأ كالآتي: هل أنت ..... ؟ وهل تفعل ..... ؟، يجب تقديم عبارات وصفية غير شخصية (عبد الخالق، 2005، ص ص 225 – 227).

ومن الجدير بالذكر أن هناك العديد من روائز الشخصية مثل رائز مينيسوتا المتعدد الأوجه، والمُعتمد كأداة رئيسية في البحث الحالي، وكذلك رائزي "أيزنك" و "كاتل" للشخصية ، تتضمن مقاييس كشف الكذب (Lie Scales)، حيث تحتوي هذه المقاييس مجموعة من الأسئلة المُدرجة في الرائز الأساسي، يُستخرج لها درجة منفصلة، إذا زادت هذه الدرجة عن حدٍ معين دلّت أن المفحوص قد زيّف في إجاباته.

رابعاً: ما وراء التحليل الإحصائي وقياس الشخصية

يتناول هذا المحور: التعريف بمنهج ما وراء التحليل الإحصائي، وذكر مراحله، وبعض الاختبارات الإحصائية المُستخدمة في هذا المنهج والموظفة في البحث الحالي، كما يتضمن إضاءة على القانون الإحصائي M والقوانين الإحصائية التي يستند إليها، ومن ثم التطرُّق لبعض مشكلات ومحدودات هذا المنهج، ويتضمن أيضاً الحديث عن درجات القطع من حيث تعريفها، وطرائق تحديدها، وأهم الأسس التقويمية لهذه الطرائق.

1- تعريف منهج ما وراء التحليل الإحصائي Meta-Analysis:

ما وراء التحليل الإحصائي: منهجية بحثية تهدف إلى الوصول إلى رؤيا تكاملية كمية لنتائج الدراسات التجريبية، حول موضوع بحثي معين، ولهذا الغرض يتم حساب مؤشرات حجم الأثر بالطرق الإحصائية الخاصة بهذا المنهج لكل جزء من الدراسة، كما يتم فحص وترميز متغيرات الدراسة، لغرض فحص علاقتها مع حجم الأثر.

وفي العقود الثلاثة الأخيرة أصبح هذا المنهج أداة بحثية منهجية ضرورية ومفيدة جداً لتجميع وتكامل النتائج حول موضوع معين، حيث أن النمو الهائل للأبحاث في علم النفس جعل من الصعب جداً تجميع هذه النتائج في أي مجال دون مساعدة الأساليب الإحصائية، وهنا لابد من الإشارة إلى أن نتائج كل جزئية من الدراسة في هذا المنهج يجب أن تكون معتمدة في القياس الكمي على المقياس نفسه، وكذلك التحليل الإحصائي في هذا المنهج يتطلب الأخذ بالاعتبار كل تقدير لحجم الأثر كدلالة ومؤشر لدقة هذا التحليل.

وفي الوقت الحالي يجري تطبيق هذا المنهج في العديد من مجالات علم النفس، وخصوصاً في تقييم فاعلية البرامج العلاجية، وبرامج الوقاية والصحة النفسية مثل برامج الصحة العقلية، والبرامج التعليمية، وبرامج الخدمات الاجتماعية، والموارد البشرية. كما يتم استخدام منهج ما وراء التحليل الإحصائي في مجالات نفسية أخرى مثل دراسة الفروق بين الجنسين في مرحلة الطفولة والمراهقة، كما يستخدم في أبحاث نفسية مع البالغين لقياس العديد من الجوانب مثل المواهب والميول والاتجاهات، ويطبق أيضاً في اختبارات الصحة النفسية المستخدمة لغرض التوظيف، حيث من الشائع جداً في الوقت الحالي أن نجد دراسات قائمة على هذا المنهج حول موضوعات مختلفة في أية مجلة علمية في علم النفس، لذا يجب أن يكون لدى الأطباء النفسيين والباحثين قاعدة معرفية كافية حول هذا المنهج، كما يجب أن تتوفر لديهم القدرة على تنفيذ الأبحاث بالاستناد إليه وتحليل وتفسير نتائج هذه الأبحاث.

2- مراحل منهج ما وراء التحليل الإحصائي:

ينطوي منهج ما وراء التحليل الإحصائي كمنهج بحث علمي على تنفيذ مراحل الدراسات التجريبية ذاتها، مع وجود بعض الخصوصية في مرحلة أو أكثر من مراحله الست الأساسية، والمتمثلة بالمراحل الآتية:

أ- تحديد أسئلة البحث: Defining the research question:

كما هو الحال في أية دراسة تجريبية، فإن الخطوة الأولى في منهج ما وراء التحليل الإحصائي هي تحديد أسئلة البحث بشكل واضح ودقيق وموضوعي قدر الإمكان، كما تتضمن هذه المرحلة أيضاً تحديد التعريفات الإجرائية والاصطلاحية للمفاهيم ذات الصلة بأسئلة البحث.

ب- تحديد أدبيات البحث: Literature search:

تتضمن هذه الخطوة تحديد الأسس النظرية لكل جزء من الدراسة البحثية، وهذا يعني أن خصائص الدراسة يجب أن تشبع في هذا الجانب ليتم تضمينها وإدراجها في منهج ما وراء التحليل الإحصائي، واختيار هذه الأسس يعتمد بشكل أساسي على تحديد الغرض من استخدام هذا المنهج، إضافةً إلى ذلك من الضروري في هذه المرحلة تحديد الدراسات ذات الصلة بالموضوع، وتحديد الاستراتيجيات المختلفة المستخدمة في الدراسة، كذلك قاعدة البيانات يجب أن تكون محددة في هذه المرحلة.

ت- تحديد وترميز المتغيرات: Coding variables:

يتم في هذه المرحلة تسجيل الخصائص والمتغيرات الأساسية في الدراسة، بغرض أن تكون واضحة ومُعتَمدة في الشرح اللاحق لتفسير النتائج.

ث- حساب مؤشرات حجم الأثر بين المتغيرات: Calculating an effect – size index:

يتم هنا حساب مؤشرات حجم الأثر والتباين بين المتغيرات، لغرض الوصول إلى نتائج كمية لكل جزء من الدراسة بالمقياس ونوع الإحصاء المناسب لها، حسب نوع هذه الدراسة ومتغيراتها المستقلة والتابعة.

ج- التحليل والتفسيرات الإحصائية:Statistical analysis and interpretation:

السمة الرئيسية لمنهج ما وراء التحليل الإحصائية هو أن يتم استخدام الأساليب والقوانين الإحصائية الخاصة بهذا المنهج، للوصول إلى تكامل نتائج الدراسة البحثية ككل، استناداً إلى قاعدة البيانات الخاصة بالدراسة، وذلك لتحقيق ثلاثة أهداف رئيسية هي:

- حساب حجم الأثر بين المتغيرات: وحجم الأثر عبارة عن مجموعة من الإحصائيات التي تحدد: القوة النسبية للفروق بين قيم الوسط الحسابي، وكذلك تحدد قيمة وقوة الفرق بين عدة معاملات ارتباط – التي تمثل الصدق أو الثبات – والتي توضح العلاقات بين متغيرات الدراسة.

- تقييم التباين لحجم الأثر في جميع نقاط الدراسة: لمعرفة إذا ما كان هذا التباين أو الفرق دال إحصائياً أم لا.
- البحث عن متغيرات البحث التي يمكن أن تفسر التباين.

ح- النتائج وإعداد البحث للنشر: Results and Publication:

أخيراً لابد من نشر نتيجة استخدام منهج ما وراء التحليل الإحصائي في بنية أية دراسة تعتمد هذا المنهج، حيث يُذكر في هذه المرحلة المقدمة، والمنهجية، والنتائج ومناقشتها، والاستنتاجات، كما يرد فيها مراجعة الأدبيات حول هذا الموضوع جنباً إلى جنب مع التعريفات الإجرائية والاصطلاحية، وذكر المتغيرات والأسئلة البحثية، والأهداف والفرضيات، وكذلك ينبغي أن يُدرج في هذه المرحلة استراتيجية البحث، والقوانين الإحصائية المستخدمة، والتحليل الإحصائي الذي تم إجراؤه في مجال التكامل في منهج ما وراء التحليل الإحصائي، وأخيراً مناقشة الآثار المترتبة للبحث في المستقبل، ومقارنة نتائج البحث مع سابقتها (Meca and Martinez, 2010, PP. 151-153).

3- بعض الاختبارات الإحصائية المستخدمة في منهج ما وراء التحليل الإحصائي والموظفة في البحث الحالي:

أ- الاختبار الإحصائي (V) لدراسة دلالة الفرق بين عدة معاملات ارتباط بيرسون:

هذا الاختبار الإحصائي وضعه هيز Hays ، والذي يتبع توزيع كاي تربيع بدرجات حرية تساوي (عدد معاملات الارتباط-1) ، والخاص بمعرفة دلالة الفروق بين عدة معاملات ارتباط بيرسون ، وفي البحث الحالي تم توظيف هذا الاختبار لدراسة الفروق بين معاملات ارتباط الاختبار بالإعادة لكل من نوعي بنود الاختبار الطويل والقصير ولكل حجم من أحجام العينات المدروسة، وكذلك بين معاملات الارتباط المحكي مع محك قائمة الأعراض لكل من نوعي بنود الاختبار الطويل والقصير ولأحجام العينات المختلفة المدروسة، بعد تحويل معاملات الارتباط إلى قيم Z بطريقة فيشر : ويحسب هذا الاختبار باستخدام المعادلة الآتية:

$$V = \sum_1^K (ni - 3)(Z_i - U)^2$$

وتحسب قيمة U من المعادلة الآتية:

$$U = \frac{[\Sigma_1^k (n_{i-3}) Z_i]}{\Sigma_1^k (n_{i-3})}$$

حيث n_i : عدد أفراد العينة في النموذج (i)

و Z_i : قيمة فيشر المقابلة لمعامل الارتباط

و i : رقم النموذج

(Hays, 1977, P. 664)

ب- الاختبار الإحصائي (M) لدراسة دلالة الفرق بين عدة معاملات ألفا كرونباخ لعينات مستقلة وحالاته المختلفة:

تم استخدام الاختبار الإحصائي M المقترح من Hakstian and Whalen والذي يتبع توزيع كاي تربيع بدرجات حرية تساوي (عدد معاملات الثبات-1)

يستند هذا القانون في حسابه إلى عدد من القوانين الإحصائية والتي سيتم توضيحها آنفاً، يحسب هذا الاختبار بصورته العامة التي تمثل الحالة الأولى من القانون والتي تختبر عدد بنود مختلف للاختبارات وأحجام عينات مختلفة، بالشكل الآتي:

$$M = \sum_{K=1}^{K} A_K - \frac{\left[\sum_{K=1}^{K} A_K (1 - r_{ak})^{-1/3}\right]^2}{\sum_{k=1}^{k} A_K (1 - r_{ak})^{-2/3}}$$

حيث A_K والتي تعبر عن قانون التباين للحالة الأولى تحسب بالشكل الآتي:

$$A_K = \frac{(J_K - 1)(9n_k - 11)^2}{18J_K(n_k - 1)}$$

حيث J : عدد فقرات الاختبار

K : رقم النموذج

r_{ak} : معامل الارتباط للنموذج k

n_k : عدد أفراد النموذج k

وبهذه الحالة تكون دلالة القانون M متمثلة بـ:

مجموع قيم قانون التباين الخاص بالحالة الأولى لكل نموذج معتمد من قبل الباحث، مطروحاً منه قيمة X*

---

* X يمثل ناتج قسمة { مربع مجموع قيم قانون التباين الخاص بهذه الحالة (عندما تتغير قيم K – رقم النموذج – بدءاً من رقم واحد وحسب عدد النماذج التي اعتمدها الباحث) مضروباً ب: واحد على الجذر التكعيبي لمتمم معامل الارتباط } مقسوماً على:

{ مجموع قيم قانون التباين الخاص بهذه الحالة مضروباً ب واحد على الجذر التكعيبي لمربع متمم معامل الارتباط } للتوضيح:

$$(1 - r_{ak})^{-1/3} = \frac{1}{\sqrt[3]{1 - r_{ak}}}$$

$$(1 - r_{ak})^{-2/3} = \frac{1}{\sqrt[3]{(1 - r_{ak})^2}}$$

وتكون دلالة قانون التباين الخاص بهذه الحالة متمثلة ب:

ناتج قسمة { درجات الحرية لعدد فقرات الاختبار مضروب ب: ناتج مربع (9 × عدد أفراد العينة في هذا النموذج مطروحاً منه 11)} مقسوماً على {(18 × عدد فقرات الاختبار) مضروباً ب (درجات الحرية لعدد أفراد العينة في هذا النموذج)}

- **وتختص الحالة الثانية من القانون باختبار الحالات التي يكون فيها عدد البنود متساوٍ بينما تكون أحجام العينات مختلفة وتحسب بالشكل الآتي:**

$$M = \frac{J-1}{18J}\left\{\sum_{K=1}^{K} B_K - \frac{\left[\sum_{k=1}^{k} B_k(1-r_{ak})^{-1/3}\right]^2}{\sum_{k=1}^{k} B_k(1-r_{ak})^{-2/3}}\right\}$$

حيث B_K والتي تعبر عن قانون التباين الخاص بالحالة الثانية تحسب بالشكل الآتي:

$$B_k = (9n_k - 11)^2/(n_k - 1)$$

وبهذه الحالة تكون دلالة القانون M متمثلة ب:

ناتج قسمة (درجات الحرية لعدد فقرات الاختبار في هذا النموذج) على (18 × عدد فقرات الاختبار في هذا النموذج) مضروباً ب:

مجموع قيم قانون التباين الخاص بالحالة الثانية لكل نموذج مُعتمد من قبل الباحث، مطروحاً منه قيمة X*

وتكون دلالة قانون التباين الخاص بهذه الحالة متمثلة ب:

ناتج قسمة { مربع (9 × عدد أفراد العينة في هذا النموذج مطروحاً منه 11)} على {درجات الحرية لعدد أفراد العينة في هذا النموذج }

- **في حين تختص الحالة الثالثة من القانون باختبار الحالات التي يكون فيها عدد البنود مختلف بينما تكون أحجام العينات متساوية وتحسب بالشكل الآتي:**

$$M = \frac{(9n-11)^2}{n-1}\left\{\sum_{K=1}^{K} C_K - \frac{\left[\sum_{k=1}^{k} c_k(1-r_{ak})^{-1/3}\right]^2}{\sum_{k=1}^{k} c_k(1-r_{ak})^{-2/3}}\right\}$$

حيث C_K والتي تعبر عن قانون التباين الخاص بالحالة الثالثة تحسب بالشكل الآتي:

$$C_k = (J_k - 1)/18\,J_K$$

وبهذه الحالة تكون دلالة القانون M متمثلة ب:

ناتج قسمة { مربع (9 × عدد أفراد العينة في هذا النموذج مطروحاً منه 11)} على {درجات الحرية لعدد أفراد العينة في هذا النموذج } مضروباً ب:

مجموع قيم قانون التباين الخاص بالحالة الثالثة لكل نموذج معتمد من قبل الباحث، مطروحاً منه قيمة X*

وتكون دلالة قانون التباين الخاص بهذه الحالة متمثلة ب:

ناتج قسمة (درجات الحرية لعدد فقرات الاختبار في النموذج) على ناتج ضرب (18 × عدد فقرات الاختبار لهذا النموذج)

- **بينما تختص الحالة الرابعة والأخيرة من القانون باختبار الحالات التي يكون فيها عدد البنود متساوٍ وأحجام العينات متساوية أيضاً وتحسب بالشكل الآتي:**

$$M = \frac{(J-1)(9n-11)^2}{18J(n-1)}\left\{K - \frac{\left[\sum_{K=1}^{K} (1-r_{ak})^{-1/3}\right]^2}{\sum_{K=1}^{K} (1-r_{ak})^{-2/3}}\right\}$$

نلاحظ عدم وجود قانون تباين خاص بالحالة الرابعة لهذا الاختبار الإحصائي، يعود ذلك لأن هذه الحالة تختص باختبار الحالات التي يكون فيها عدد البنود متساوٍ وأحجام العينات متساوية أيضاً.

وبهذه الحالة تكون دلالة القانون M متمثلة ب:

ناتج قسمة { درجات الحرية لعدد فقرات الاختبار مضروباً ب مربع (9 × عدد أفراد العينة في هذا النموذج مطروحاً منه 11)} على {(18 × عدد فقرات الاختبار) مضروباً ب (درجات الحرية لعدد أفراد العينة في هذا النموذج)} مضروباً ب:

رقم النموذج المعتمد من قبل الباحث، مطروحاً منه ناتج قسمة { مربع مجموع القيم (عندما تتغير قيم K – رقم النموذج – بدءاً من رقم واحد وحسب عدد النماذج التي اعتمدها الباحث) ضرب واحد على الجذر التكعيبي لمتمم معامل الارتباط } على

{ مجموع قيم قانون التباين (عندما تتغير قيم K – رقم النموذج – بدءاً من رقم واحد حسب عدد النماذج التي اعتمدها الباحث) مضروباً ب واحد على الجذر التكعيبي لمربع متمم معامل الارتباط } (Hakstian & Whalen, 1976, PP. 222- 224)

ت- حجم التأثير/ قوة الارتباط (Effect size):

من الطرق التي يتم بها تقييم أهمية النتائج، ويعرف أيضاً بقوة الارتباط (Strength of association)، وحجم التأثير عبارة عن مجموعة من الإحصائيات التي تحدد القوة النسبية للفروق بين قيم الوسط الحسابي، بعبارة أخرى، يصف حجم التأثير مقدار التباين الكلي في المتغير التابع، الذي يمكن التنبؤ به من معرفة مستويات المتغير المستقل، ويعد

اختبار إيتا تربيع (eta squared) من أكثر الإحصاءات المستخدمة لحساب حجم التأثير شيوعاً، حيث يمثل هذا الإحصاء نسبة تباين المتغير التابع التي يفسرها ويوضحها المتغير المستقل، ويمكن أن تتراوح قيم إيتا تربيع من (0 إلى 1). (بالانت، 2006، ص225)

ونستطيع الاستعانة بالإرشادات الآتية لتفسير قوة قيم إيتا تربيع:

- .01 = تأثير ضئيل
- .06 = تأثير معتدل
- .14 = تأثير كبير

وتحسب قيمة حجم الأثر من خلال المعادلة الآتية:

قيمة إحصاء إيتا تربيع = $\frac{\text{مجموع المربعات البينية}}{\text{المجموع الكلي للتربيع}}$ (المرجع السابق، ص 246)

ث- إحصاء Z فيشر:

تم استخدام الاختبار الإحصائي Z لفحص الفروق بين معاملات الارتباط بطريقة الإعادة ومعاملات الارتباط مع قائمة الأعراض لكل من نوعي بنود الاختبار الطويل والقصير، والذي يحسب بعد تحويل معاملات الارتباط إلى قيم Z بطريقة فيشر وذلك بالمعادلة الآتية:

$$Z = \frac{z_1 - z_2}{\sqrt{\left(\frac{1}{n_{1-3}}\right)+\left(\frac{1}{n_{2-3}}\right)}}$$

حيث z_1 و z_2 قيمة فيشر للمعاملين

و n_1 و n_2 عدد أفراد كل من العينتين (علام، 2005، ص221)

وتحول معاملات الارتباط إلى قيم Z بطريقة فيشر وفق المعادلة الآتية:

$$Z_i = \frac{1}{2}\ell n\left[\frac{(1+r_i)}{(1-r_i)}\right]$$

(علام، المرجع السابق، ص156)

4- القوانين الإحصائية التي يستند إليها القانون الإحصائي M:

أ- مقياس كاي مربع بدرجات حرية (N -1):

يُعتبر هذا المقياس من أهم مقاييس الدلالة الإحصائية، وقد أُطلق عليه هذا الاسم نسبة إلى واضعه كارل بيرسون، والغرض من هذا المقياس هو الكشف عن مدى اختلاف التكرار التجريبي أو الواقعي عن التكرار المتوقع أو المحتمل، وبالتالي التأكُّد مما إذا كانت القيم أو التكرارات التجريبية تختلف اختلافاً دالاً عن القيم أو التكرارات المتوقعة التي يفترض حدوثها على المستوى الاحتمالي النظري، وتزداد قيمة كاي مربع مع ازدياد الفروق بين التكرار التجريبي

والتكرار المتوقع أو المحتمل، وتنخفض هذه القيمة مع تناقص هذا الفرق، إلى أن تنعدم وتصل إلى الصفر، مما يشير في هذه الحالة إلى مطابقة التكرار التجريبي للتكرار المتوقع وهو التكرار الاعتدالي.

ترتكز الطريقة العامة لحساب كاي مربع على الصيغة الآتية:

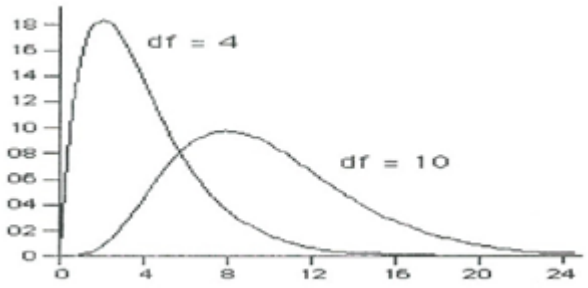
$$\text{كا}^2 = \text{مج} \frac{(\text{ت و} - \text{ت م})^2}{\text{ت م}}$$

حيث يشير الرمز ت و إلى التكرار الواقعي (أو التجريبي أو الملاحظ)

والرمز ت م إلى التكرار المتوقع (الاحتمالي)

ولحساب كا^2 لابد من حساب درجات الحرية وتحسب وفق الصيغة (n – 1) (ميخائيل، 1997، ص ص 135- 137)

حيث يعتمد توزيع كاي مربع اعتماداً كاملاً على درجات الحرية فكلما زادت درجات الحرية كلما قل التواء التوزيع واقترب من التماثل كما يبينه الشكل الآتي:



الشكل 2 توزيع كاي تربيع

(حريري، 2009، ص 2)

ب- درجات الحرية:

تُمثّل درجات الحرية عدد الدرجات أو التكرارات التي يمكن أن تتغير حول قيمة ثابتة أو مقياس معين للمجتمع الأصلي، فإذا كان لدينا مجموعة من الدرجات عددها عشرون درجة، ولها متوسط معروف 10 مثلاً، فإنه يترتب على ذلك أن تكون أي تسع عشرة درجة من هذه الدرجات العشرين حرة في تغيير قيمتها، بينما تكون الدرجة العشرين مقيدة بقيمة معينة تُضاف إلى القيم التسع عشرة حتى يصبح المتوسط 10، لذلك تكون درجات الحرية التي تتشتت حول متوسط ذلك التوزيع مساوية (ن _ 1). (أبو النيل، محمود السيد، 1987، ص 241)

ت- **إحصاء فيشر الخاص بالاختبار:**

والمتمثل بالقانون :

$$F = \frac{1 - r_a}{1 - p_a}$$

حيث تمثل r_a قيمة معامل الارتباط للعينة المدروسة

وتمثل p_a قيمة معامل الارتباط للمجتمع المدروس

(Hakstian & Whalen, 1976, P. 220)

ث- قانون حساب معامل كرونباخ ألفا:

ويجري حساب معامل ألفا وفق المعادلة الآتية:

$$a = \frac{K}{K-1}\left[1 - \frac{\sum S_i^2}{S_t^2}\right]$$

حيث K: عدد أسئلة الاختبار.

و $\sum S_i^2$: مجموع تباين كل سؤال من أسئلة الاختبار، أي $S_1^2 + S_2^2 + \cdots + S_i^2$

و S_t^2: تباين درجة الاختبار الكلية. (Lee, 2004, P. 396)

ج- **قانون التباين العام للاختبار (var) لأحجام عينات مختلفة، وأطوال اختبار مختلفة:**

ومنه تم اشتقاق مجموعة قوانين التباين الخاصة بكل حالة من حالات الاختبار، ويحسب بصورته العامة وفق المعادلة الآتية:

$$var = \frac{18J_K(n_k - 1)(1 - r_{ak})^{2/3}}{(J_K - 1)(9n_k - 11)^2}$$

(Hakstian & Whalen, 1976, P. 223)

5- بعض مشكلات ومحدودات منهج ما وراء التحليل الإحصائي:

يُعدّ منهج ما وراء التحليل الإحصائي مدخلاً علمياً جديداً، يسهم في تعزيز الأدلة المستندة إلى الممارسة التجريبية في علم النفس، وبالتالي يهدف إلى تشجيع المهنيين لإسناد ممارساتهم إلى أقصى حدٍّ ممكن من الأدلة العلمية التي يتم الحصول عليها من الأبحاث، ومع ذلك فإن هذا المنهج لديه بعض المشكلات التي لابد من الوقوف عندها ومحاولة الحد منها، والتي تتلخص في أن صدق وصحة النتائج المُستمدة من هذا المنهج تعتمد على نوعية الدراسات التجريبية المتكاملة، وفي حال كان أي جزء من الدراسة القائمة على هذا المنهج يستند على تقديراتٍ منحازة، فإن النتائج المستخرجة من هذا المنهج ستكون منحازة بمجملها، ولذلك فإن تقييم

الجودة المنهجية لكل جزء من الدراسة على حدة يُعد من المستلزمات الرئيسية في هذا المنهج، علاوةً على ذلك فإن منهج ما وراء التحليل الإحصائي يمكن أن يعاني من القصور في الإجراء والممارسة، لذا ينبغي تنفيذ خطواته بشكل دقيق، كما ينبغي إجراء دراسة مفصلة للمتغيرات المتضمنة في الدراسات المعتمدة عليه.

(Meca and Martinez, 2010, PP. 161-162)

6- درجات القطع Cut - off scores:

أ- مسمياتها وتعريفها:

تتعدّد مسميات درجة قطع الاختبارات بتعدد آراء وخبرات علماء القياس والتقويم، ومن تلك المسميات ما ذكرها هالبن Halpin وهي كما يلي:

- درجة القطع Cut of score.
- الحد الأدنى من الكفاية Minimum Competency.
- درجة المحك Criterion Score.
- درجة الاجتياز Passing Score.

وبالمقابل فقد تتعدّد تعريفات علماء القياس والتقويم النفسي والتربوي لهذا المفهوم بتعدُّد المنظور المُستخدم في ذلك.

فقد عرفها هامبلتون (Hambleton) بأنها:

نقطة على متصل درجات الاختبار، تستخدم لتصنيف الطلاب إلى فئتين تعكس مستويات الأداء المختلفة بالنسبة لهدف معين، أو بالنسبة لمجموع الأهداف المراد قياسها في الاختبار (التميمي، 1999، ص 25).

في حين عرفها بيرك (Berk) بأنها:

تلك الدرجة التي تزيد من نسبة التصنيفات الصحيحة، وفي المقابل تقلل من نسبة التصنيفات الخاطئة.

وعرفها بوبام (Popham) بأنها:

مقياس لمدى ملاءمة أداء المفحوص لهدف محدد.

ولعل البحث الحالي يأخذ بالتعريف الذي يشير إلى أن:

درجة القطع هي نقطة على متصل درجات الاختبار، تستخدم لتصنيف الأفراد إلى فئات تعكس مستويات الأداء لمختلف الجوانب المُراد قياسها. (مجيد،2007، ص ص 186- 187).

ب- طرق تحديد درجة القطع:

تَبنَّى خبراء القياس والتقويم التربوي العديد من الطرق لتحديد درجة القطع، اختلفت باختلاف خبرات وتخصصات هؤلاء الخبراء، فقد ذكر بيرك (Berk, 1986) أن هذه الطرق بلغت ثمانٍ وثلاثين طريقة مختلفة من حيث أسلوب بنائها، وتطبيقها وتفسير نتائجها، ويمكن تصنيفها إلى ثلاثة تصنيفات على النحو الآتي:

أولاً: الطرق التحكيمية Judgmental – methods:

تستند هذه الطرق على آراء المتخصصين في مجال القياس أو في محتوى الاختبار المراد قياسه، لتحديد درجة القطع المناسبة، حيث يُسترشد بآراء هؤلاء في تحديد مستوى الطالب ذو الحد الأدنى من الكفاية، أي تحديد عدد الأسئلة التي يجب أن يجيب عنها الطالب بصورة صحيحة، واضعين باعتبارهم المتغيرات التي تؤثر في تحديد هذه الكفاية.

مميزات وعيوب الطرق التحكيمية:

لكل طريقة من الطرق التحكيمية مميزات وعيوب خاصة بها، إلا أن هناك مميزات وعيوب تكاد تكون مشتركة بينها، أما مميزات هذه الطرق فهي كما ذكرها (Berk, 1986) على النحو الآتي:

1- سهلة الاستخدام مقارنةً بالطرق الأخرى.
2- سهلة الفهم والتفسير.
3- سهلة الحساب.
4- لا تعتمد على أساليب إحصائية معقدة.
5- لا تحتاج إلى وقت وجهد كبيرين.

وأما عن عيوب هذه الطرق فهي كما أشار إليها كل من بيرك، وهامبلتون، وإيغنور (Hambelton &Eignor 1980, Berk 1986) على النحو الآتي:

1- تعتمد كلياً على آراء المحكمين، وبالتالي كفاية هؤلاء المحكمين لها دور بارز في الوصول إلى المعيار المناسب لتصنيف الطلاب، وتحديد بنود الاختبار التي يمكن لأقل الطلاب كفاءةً الإجابة عنها بشكل صحيح.
2- صعوبة تحديد الطلاب الأقل كفاءةً.
3- لا تعتمد على بيانات تجريبية فعلية لكي يسترشد بها المحكمون في تقدير أحكامهم.
4- الحصول على درجات قطع مختلفة عند تطبيقها على اختبار موحد وعينة دراسية واحدة.
5- ذاتية المحكمين في اختيار الحد الأدنى من الكفاية، وذلك بسبب اختلاف تصوراتهم حول عملية الإتقان.
6- تفاوت عملية الاتساق الداخلي في عملية التحكيم.
7- صعوبة التنبؤ بأداء الطلاب في الاختبار.

وعلى الرغم من وجود بعض العيوب في استخدام الطرق التحكيمية، إلا أنه توجد بعض المعايير التي تحد من تأثيرها، وقد ذكر بيرك (Berk, 1986) بعضاً منها على النحو الآتي:

أ- استخدام طرق لتوحيد بيانات الأداء الفعلي.

ب-يجب أن تكون عملية تحكيم مفردات الاختبار متزامنة مع تطبيق الاختبار.

ت-الاهتمام بتدريب المحكمين.

ث-تنظيم عملية تحكيم بنود الاختبار.

ج- إيجاد الحلول المناسبة لتسوية الاختلافات في أحكام المحكمين.

ح- تصميم استمارات تنظم عمليات التحكيم.

خ- تزويد المحكمين ببيانات مسبقة حول أداء الاختبار.

والجدول التالي يوضح بعض هذه الطرق، كما يوضح مميزات وعيوب كلٍ منها.

الجدول رقم 6: أنواع الطرق التحكيمية لتحديد درجة القطع

طريقة تحديد درجة القطع	طريقة إجرائها	مميزاتها	عيوبها
طريقة أيبل Ebel's Method	تعتمد على تعليقات المحكمين على مفردات الاختبار، باعتبار أن مواصفات المفردات لها بعدين هما: بعد الصلة، الذي يتضمن أربعة مستويات (أساسية، مهمة، مقبولة، غير مهمة)، وبعد الصعوبة، الذي يتضمن ثلاثة مستويات(سهلة، متوسطة، صعبة)، ويستخدم المحكم أعداداً مختلفة من تصنيفات الصعوبة والصلة، ثم يطلب من كل محكم ما يلي: - تحديد الخلية المناسبة لكل مفردة بناءً على توافق محتواها مع تصنيفات البعدين. - تحديد النسبة المئوية لعدد الطلاب ممن لديهم الحد الأدنى من الكفاية، ثم يُضرب عدد مفردات كل خلية في النسبة المئوية، ويُقسم مجموع نواتج جميع الخلايا على عدد المفردات، وبالتالي يتم الحصول على درجة القطع (Jeager, 1989)	- سهلة الإجراء والفهم والتطبيق. - لا تحتاج لعمليات إحصائية معقدة.	- لم يلتزم إيبل عند وضع هذه الطريقة بمواصفات محددة لكلا البعدين، فربما يعتمد كل محكم على مواصفات يحددها لنفسه مما يؤثر في دقة التصنيف. - تعتمد على تحكيم المفردة ككل، ولا تأخذ بالاعتبار بدائل الإجابات التي تشتملها مفردات الاختيار من متعدد. - تحتاج إلى تحكيم خاص يتمثل بنسب مئوية للمفردات وذلك لكل خلية عبر بعدي الصلة والصعوبة المستمرين، فإذا لم تكن هنالك معايير خارجية يُبنى عليها الحكم فإنه سيغدو حكماً اعتباطياً. (Berk 1986)

طريقة أنجوف Angoff'sMethod	يُطلب من كل محكم فحص كل فرة من فقرات الاختبار، ثم يُطلب منهم تصوُّر مجموعة الأفراد الذين حققوا الحد الأدنى للكفاية التي يقيسها الاختبار، ثم تقدير نسبة عدد الأفراد الذين يُحتمل أن يجيبوا إجابة صحيحة عن كل مفردة ، ويمثل متوسط هذه النسب الحد الأدنى لمستوى الاجتياز في الاختبار. (Angoff, 1971)	– هلة الفهم والتنفيذ. – تعتمد على تصور عدد من الأفراد الذين حققوا الحد الأدنى للكفاية المطلوبة بدلاً من تصور فرد واحد.	– صعب على كثير من المحكمين معرفة الطلاب من ذوي الحد الأدنى للكفاية. – ويصعب أيضاً تقدير القيم الاحتمالية لبعض المفردات المصاغة من عبارات منفية أو المتضمنة عمليات حسابية. (Berk 1986)
طريقة نيدلسكاي Nedlesky's Method	اقترح نيدلسكي (Nedlesky, 1954) هذه الطريقة لتحديد درجة قطع الاختبارات التي تشتمل على أسئلة الاختيار من متعدد، حيث يُطلب من كل محكم فحص كل مفردة من مفردات الاختبار، ثم تحديد من بين الخيارات الخاصة لكل مفردة تلك التي يمكن أن يتجنب الطلاب ذوي الحد الأدنى للكفاية اختيارها، بعد ذلك يضع المحكم درجة للسؤال بناءً على ما تبقى من الخيارات، وذلك بأخذ مقلوب الخيارات المتبقية. فمثلاً: إذا كانت المفردة تشتمل على خمسة خيارات ورأى المحكم أن الطالب ذو الحد الأدنى للكفاية يمكن أن يتجنب اختيار ثلاثة منها، عندها تكون درجة السؤال كالتالي: $0.5 = \frac{1}{2} = \frac{1}{3-5}$ بعد ذلك يقوم المحكم بجمع درجات الأسئلة، ويكون المجموع الكلي هو درجة القطع بالنسبة لذلك المحكم، وتجمع في النهاية الدرجات التي يتحصل عليها المحكمون، ثم يحسب المتوسط وبالتالي نحصل على درجة القطع للاختبار ككل.	– سهلة الفهم والتنفيذ.	–لا يوجد تبرير منطقي يؤيد افتراض ندلسكاي أن الطالب الذي لا يعرف الإجابة الصحيحة لسؤال معين لا يلجأ إلى التخمين العشوائي لأن مشتتات الأسئلة تصمم عادة لجذب غير المتمكنين. – كما أن هذه الطريقة تسمح فقط بقيم احتمالية منفصلة تعتمد على عدد خيارات الإجابات، لذلك يضطر المحكمون لتعيين القيمة (0.5) لمعظم المفردات. – هذا بالإضافة إلى أن هذه الطريقة تحتاج لقدر كبير من التدريب لكي يستخدمها المحكمون استخداماً مناسباً. (Berk 1986)

طريقة أنجوف - نادسكاي المختلطة	في هذه الطريقة يتم حساب المعيار لتحديد درجة القطع على أساس حساب متوسط المعايير الناتجة عن استخدام طريقتي أنجوف ونيدلسكاي.	لها نفس مميزات طريقتي أنجوف ونيدلسكاي.	لها نفس عيوب طريقتي أنجوف ونيدلسكاي بالإضافة إلى احتمالية زيادة مصادر الخطأ المختلفة، وذلك عندما نحسب متوسط المعيارين. (Berk 1986)
طريقة تقدير أهمية الصعوبة difficulty importance estimate	يقوم المحكمون بفحص كل هدف لتقدير الحد الأدنى للنجاح، ويجب أن تتفق عينة المحكمين على الحد الأدنى لكل هدف من الأهداف، وذلك لأن للحد الأدنى وزن خاص طبقاً لأهمية صلته بالهدف، ثم تضرب النسبة المئوية لدرجة النجاح في نسبة أهمية الدرجة لكل هدف، والناتج يُجمع من خلال جميع الأهداف، ويُضرب في عدد بنود الاختبار، وذلك للوصول لمعيار تحديد درجة القطع، ويجب أن يضع الحكام في الاعتبار الخطأ المعياري عند اتخاذهم القرارات الخاصة بدرجة القطع.	– سهلة الفهم. – سهلة الاستخدام وتعمل على ربط درجة القطع بقائمة خصائص النطاق السلوكي.	– استخدمت هذه الطريقة تقديرات المحكمين للأهداف لتقدير درجة القطع لبنود الاختبار، لكن ليس بالضرورة افتراض أن تقديرات الأسئلة تساوي تقديرات الأهداف. – كما أنه من المحتمل أن يكون نوع الخطأ المعياري المستخدم غير ملائم لتفسير درجة القطع. (Berk 1986)
طريقة مواصفات المفردة	في هذه الطريقة يقوم المحكمون بقراءة مواصفات كل مفردة من المفردات التي تحتوي على المهارة المراد قياسها، ووصف محتوى البند، وخواص المثير، والقوانين التي تحكم إحداث خيارات الاستجابة، وعينة المفردات، ومن ثم فإن على المحكمين إدراك وتصور الإطار الكلي للمفردات، والتنبؤ بنسبة البنود التي يمكن أن يجيب عليها المُمتحن ذو الحد الأدنى من الكفاية إجابة صحيحة.	– سهلة الفهم والتطبيق. – تقوم بربط درجة القطع بمواصفات النطاق السلوكي وكافة البنود التي تقيسه. –تستغرق زمناً أقل مما لو تم تقدير كل مفردة بشكل منفصل. –تكون درجة القطع فيها مرتبطة بالنطاق السلوكي وليس بصيغة الاختبار المحدد.	تُستخدم في هذه الطريقة تقديرات المحكمين لمواصفات البنود لتقدير درجة القطع لمفردات الاختبار، إلا أنه ليس من الضروري أن تكون تقديرات درجة القطع مماثلة لتقديرات مواصفات البنود. (Berk 1986)

(التميمي، 1999، ص ص 26- 32، ذيب، 2010، ص ص 350- 359، مجيد، 2007، ص ص 187- 194 ،PP. 3- 13 ,Socha& Braclay, 2005)

ثانياً: الطرق التي تعتمد جزئياً على التحكيم وتسترشد ببيانات تجريبية:

أدرك الباحثون قصوراً ملحوظاً في الطرق التحكيمية، وذلك بسبب اعتمادها على تحكيم مفردات الاختبار دون النظر إلى الأداء الفعلي للمختبرين، مما يؤدي إلى الحصول على معايير لا تتصف بدرجة كافية من الصدق. (علام، 1995، ص ص 241-242)

لذا اقترح علماء القياس العديد من الطرق المختلفة لتحديد مستويات الأداء، والتي تتضمن طرقاً تجريبية، محاولين بذلك تجاوز عيوب الطرق التحكيمية، والحصول على مستويات أداء دقيقة يمكن الاعتماد عليها، والجدول الآتي يوضح بعض هذه الطرق، ومميزات وعيوب كلٍ منها:

الجدول رقم 7:

أنواع الطرق التي تعتمد جزئياً على التحكيم وتسترشد ببيانات تجريبية لتحديد درجة القطع

طريقة تحديد درجة القطع	طريقة إجرائها	مميزاتها	عيوبها
الطريقة التحكيمية المعززة بالمعلومات	تعتمد هذه الطريقة على بيانات مسبقة تتعلق بالأداء الفعلي للأفراد. وتتضمن هذه البيانات إحصاءات خاصة بالمفردات مثل معاملات الصعوبة والتمييز، ومتوسط درجات الاختبار بعد تطبيقه على العينة المناسبة من الأفراد، كما تتضمن بعض تعليقات وأحكام الخبراء حول مفردات الاختبار. وتُقدَّم جميع هذه المعلومات للمحكمين لفحصها وتبادل المقترحات حولها، ومناقشتها وإعادة النظر فيها، وذلك لكي يتم التوصل إلى اتفاق حول تحديد درجة القطع المناسبة.	- سهلة التطبيق والفهم والحساب. - تجمع وتوحد بين بيانات الأداء، والبيانات التحكيمية المُستمدة من مجموعات المحكمين المختصة.	- تتطلب جمع الكثير من البيانات التجريبية، والمعلومات المتعلقة بالاختبار ومحتواه من مصادر متعددة، وقد يستغرق ذلك وقتاً وجهداً كبيرين. - تحليلات المحكمين للبيانات في هذه الطريقة غير منظمة، لذلك تظهر الحاجة لوجود خطوط تفصيلية تسهل استخدام أنواع البيانات المختلفة والمناقشة الجماعية. (Berk,1986)

طريقة أنجوف التجريبية المعدلة	وهي عبارة عن تعديل لطريقة أنجوف التحكيمية، تتطلب تزويد المحكمين ببعض البيانات المستمدة من التجريب الميداني لمفردات الاختبار، ويُطلب من كل محكم تقدير احتمال إجابة الأفراد ذوي الحد الأدنى من الكفاية المطلوبة لمفردات الاختبار ككل وليس من كل مفردة على حدة كما في الطريقة السابقة، وبعد الحصول على هذه التقديرات يتم تزويد المحكمين ببيانات تجريبية، تشتمل على متوسط صعوبة كل مفردة، ثم يُطلب منهم تعديل تقديراتهم في ضوء تلك البيانات التجريبية، ويكون الحد الأدنى لمستوى الاجتياز في الاختبار متوسط مجموع تقديراتهم المعدلة.	-سهلة التطبيق والفهم والحساب. - وتعمل على ربط درجة القطع بالكفايات المقاسة. - وتدمج معامل صعوبة البند الحقيقي خلال عمليات التحكيم.	- عملية استخدام تقديرات المحكمين للكفايات يجب أن تكون على درجة عالية من التوافق والانسجام، وذلك بين المفردات والكفاية ولا يمكن افتراض أن تكون التقديرات الواقعية متطابقة (Berk,1986)
طريقة توفق بين الطرق المطلقة والطرق النسبية	تعمل هذه الطريقة على التوفيق بين أحكام المحكمين المتعلقة بأداء الأفراد في الامتحان، والأداء الفعلي المُستمد من البيانات التجريبية. ويُطلب من المحكمين تحديد قيمتين هما: - الحد الأدنى للنسبة المئوية لعدد المفردات التي يجب أن يجيب عنها الفرد إجابة صحيحة، ويرمز لها بالرمز (k). - النسبة المئوية لعدد الأفراد المتوقع اجتيازهم للاختبار، ويرمز لها بالرمز (v). ثم رسم منحنى للمتوسطات والانحرافات المعيارية لقيم (v) ,(k) التي حددها المحكمون، وكذلك النسبة المئوية لعدد الأفراد المتوقع نجاحهم ويرمز لها بالرمز (y) كدالة تناقصية لدرجات الاختبار ويرمز لها بالرمز (x) ثم يرسم منحني لتحديد النقط(i) والتي تمثل الحد الأدنى للنجاح.	- توفق هذه الطريقة بين الأحكام المتعلقة بمفردات الاختبار، وبين الأداء الفعلي للمختبرين أي بين المستويات المطلقة والمستويات النسبية.	- تتطلب من المحكمين أن يوفروا تقديرين، وهذا ما يجعل تنفيذ هذه الطريقة أكثر صعوبة من غيرها، سواء في عملية التحكيم أو من الناحية الإحصائية. - بالإضافة إلى أن تفسير المعيار من قبل المتخصصين قد يكون صعباً جداً.

(التميمي، 1999، ص ص 32- 36، ذيب، 2010، ص ص 360- 364، مجيد، 2007، ص ص 195- 200 ،PP. 3- 13 ,Socha& Braclay, 2005)

ثالثاً: طرق تعتمد على البيانات التجريبية وتسترشد بالتحكيم Empirical Judgmental Method: يتضمن هذا النوع العديد من طرق تحديد درجة القطع، يبين الجدول التالي بعضها:

الجدول رقم 8: أنواع الطرق التي تعتمد على البيانات التجريبية وتسترشد بالتحكيم

طريقة تحديد درجة القطع	طريقة إجرائها	مميزاتها	عيوبها
طريقة المجموعات المحكية	تتطلب هذه الطريقة انتقاء مجموعتين من الطلاب: الأولى متقنة (تلقت التعليم)، والثانية غير متقنة (لم تتلق التعليم)، ويكون مدى توزيع درجات الطلاب من الدرجة (صفر) إلى الدرجة (ن) حيث (ن) تمثل عدد الأسئلة التي تقيس كل هدف. وتتطلب هذه الطريقة تقسيم الطلاب بناءً على درجة القطع (س) إلى قسمين (متقنين ، غير متقنين)، ثم تقسيم طلاب المجموعة المتقنة بناءً على درجة القطع إلى فئتين: الأولى متقنين بالفعل والثانية متقنين غير حقيقيين، وبالمثل يتم تصنيف طلاب المجموعة الثانية غير المتقنة إلى فئتين: الأولى غير متقنين فعلاً والثانية غير متقنين غير حقيقيين. - وتُحدد درجة القطع (س) من خلال فحص خلية الاحتمالات لكل درجة قطع، وتكون درجة القطع المثلى: هي تلك التي تزيد نسبة التصنيفات الصحيحة، وتقلل نسبة التصنيفات الخاطئة. ويتم تحديدها بناءً على نقطة تقاطع التوزيعين للطلاب المتقنين وغير المتقنين.	- تُعتبر هذه الطريقة نموذج للصدق التكويني. -كما تُعتبر ذات حساسية خاصة للخصائص الإحصائية للمفردة، وكذلك للصدق البنائي.	من عيوبها أنها لا تحدد المعيار الحقيقي، بل تقوم بتحديد موقع الأداء على متصل الدرجات. وفي بعض الأحيان يتعذر التمييز بين المجموعتين المتقنة وغير المتقنة.

طريقة المجموعات المتضادة	تُعد هذه الطريقة مماثلة إلى حدٍ ما لطريقة المجموعات المحكية، ولكنها تختلف بكيفية تحديد درجة القطع في الاختبار، ففي هذه الطريقة يقوم المحكمون المتخصصون بتحديد مجموعتين من الطلاب، بحيث تكون الأولى متقنة والأخرى غير متقنة، بناء على ما يلي: – المعلومات المتوفرة حول مستويات مهارات الطلاب. – درجات الطلاب وأدائهم داخل الفصل، حيث يتم تصنيف الطلاب إلى فئتين: الأولى الحاصلون على تقدير جيد جداً فما فوق، والثانية الطلاب الحاصلون على تقدير مقبول فما دون. ثم يطبق الاختبار على كلتا المجموعتين، ويرسم منحنيان يمثلان التوزيعين التكراريين لدرجات الطلاب في الاختبار، وتمثل نقطة تقاطع المنحنيين درجة قطع الاختبار، وهذه الطريقة تفترض تداخل التوزيعين التكراريين وعدم تطابقهما.	– سهلة الفهم والتطبيق والإحصاء.	لها نفس عيوب طريقة المجموعات المحكية، بالإضافة إلى أنه من الصعب تجاوز الخطأ الناجم من المحك.
طريقة المجموعة الحدية	تتطلب هذه الطريقة من المحكمين انتقاء عينة من الطلاب الذين يرون بأن كفايتهم المراد قياسها تقع عند الحد الفاصل بين التمكُّن وعدم التمكُّن، ومن المحتمل أن تكون نسبة هؤلاء الطلاب قليلة بالنسبة للمجموعة الكلية، ثم يطبق الاختبار على هذه العينة، وبعد ذلك يتم إيجاد وسيط درجات هذه المجموعة ليكون بمثابة مستوى الأداء المطلوب.	– سهلة الفهم والتطبيق والإحصاء. لكنها تحتاج إلى عدد كاف من الطلاب لكي يتم الوصول إلى تقديرات مستقرة نسبياً.	– يصعب على المحكمين أو المدرسين انتقاء المجموعة الحدية، إذ أن تعريف أو تحديد الطالب الذي يقع عند الحد الفاصل بين الاتقان أصعب بكثير من تحديد الفرد المتقن أو غير المتقن. – كما قد يصعب تحديد الحجم المناسب للعينة.

طريقة المحك معياري المرجع	يتم اختيار المعيار في هذه الطريقة بناءً على العلاقة بين درجات القطع في الاختبار محكي المرجع، وبين درجات الأداء في الاختبار معياري المرجع، ويتم اختيار درجة القطع أولاً في الاختبار معياري المرجع بطريقة اختيارية (على سبيل المثال: مناظرة الدرجة الخام بالدرجة المئينية 50 أو 75) ثم يتم اختيار درجة القطع في الاختبار محكي المرجع على أساس أنها تزيد ما يلي: -نسبة ثبات القرارات الناتجة عن تطبيق الاختبار معياري المرجع ويرمز لها بالرمز (P). - نسبة ثبات القرارات المصححة من عامل الصدفة ويرمز لها بالرمز (K). - تعيين دالة الفقد (Loss Function) ويرمز لها بالرمز (L).	- سهلة الفهم والتطبيق والإحصاء، ويمكن توفير بيانات محك الاختبار معياري المرجع بسهولة ويسر.	- إن استخدام درجات القطع الاختيارية في الاختبار معياري المرجع كمحك يعد خطأ فنياً وغير منطقي.
طريقة النتائج التربوية	تعتمد هذه الطريقة أساساً على اختيار المعيار من خلال العلاقة بين درجات الاختبار ومحك التعلم المستقبلي (نتائج انفعالية أو معرفية)، حيث يتوقع أن يزداد المنحني طردياً ليربط بين المتغيرين، وتُحدد درجة القطع اعتماداً على أثرها في زيادة أداء المحك.	-سهلة التطبيق والفهم والحساب.	- إن معيار (التعلم المستقبلي) يحتاج إلى تحديد واضح. - العلاقة بين المتغيرات المقترحة لا توفر دالة درجية بالضرورة((Step Function لتحديد درجة القطع.

(التميمي، 1999، ص ص 37- 40، ذيب، 2010، ص ص 365 - 369، مجيد، 2007، ص ص 200- 13,204 -PP. 3، Socha& Braclay, 2005)

ج-الأسس التقويمية لطرق تحديد درجة القطع:

ذكر بيرك (Berk,1986) بأنه يوجد معيارين لتقويم طرق تحديد درجة القطع، وهما:

أولاً: <u>الكفاية الفنية Technical- Adequacy : وتتضمن الآتي:</u>

- **أن تُحدث الطريقة تصنيفاً مناسباً للمعلومات Appropriate Classification of information:**

بمعنى أن تُحدث الطريقة تصنيفات ثنائية مثل: الاتقان/ عدم الاتقان، الفشل/النجاح، ويعتبر التصنيف أساسياً لمعظم القرارات المُتخذة.

- **أن تكون حساسة للأداء Sensitive to Examine the Performance:**

يجب أن تكون الطريقة حساسة لمستويات صعوبة الأسئلة المختلفة، ولأي خصائص أخرى مثل معامل التمييز.

- **أن تكون الطريقة حساسة لكل من عمليتي التعليم و التدريب Sensitive to each education and Training:**

يجب أن تراعي الطريقة عملية التعلم والتدريب الفعلي الذي تلقاه الطالب، بحيث تقاس المهارات والكفايات التي تلقاها الطالب بالفعل.

- **أن تتضمن طرقاً إحصائية مناسبة Include appropriate Statistically methods:**

يجب استخدام الطرق الإحصائية المناسبة ، كما يجب أن تفسر بشكل صحيح.

- **أن تحدد الطريقة المعيار الصحيح Identify the true standard:**

كما أنها يجب أن تراعي أخطاء القياس.

- **أن تثبت الطريقة صدق القرار Decision Validity Evidence:**

بمعنى أنه يجب الحصول من الطريقة على تقديرات احتمالات قرارات التصنيفات الصحيحة والخاطئة.

ثانياً: <u>الجانب التطبيقي (العملي) Practicability: ويتضمن أن تكون درجة القطع:</u>

- **سهلة التطبيق Easy to implement:**

يجب أن تكون الخطوات المستخدمة في تطبيق الطريقة واضحة وسهلة الفهم للمشاركين فيها، وأن تتم في قدر معقول من الوقت.

- **سهلة الحساب Easy to Compute:**

يجب أن تكون الطرق الإحصائية المستخدمة للحصول على درجة القطع سهلة الاستخدام، سواء باستخدام الآلة الحاسبة أو البرامج الإحصائية المتوفرة في الحاسوب.

- **سهلة للتفسير Easy to interpretation:**

يجب أن تكون الطريقة المستخدمة في تحديد درجة القطع قابلة للفهم والتفسير ، ويجب أن تكون تفسيرات الطريقة واضحة وسهلة الإدراك والفهم، للعامة والمهتمين على حدٍ سواء.

كما أشار بيرك (Berk,1986) إلى عدة عوامل تساعد على اختيار الطريقة المناسبة وهي:

1- مدى أهمية القرارات المبنية على تحديد درجة القطع.

2- الكمية المتاحة من الوقت لتحديد درجة القطع.

3- مدى توفر المصادر المادية، والطاقة البشرية لإنجاز مهمة تحديد درجة القطع.

4- مدى مناسبة طريقة تحديد درجة القطع لنوعية الاختبار المراد قياسه

(ذيب، 2010، ص ص 373- 375، مجيد، 2007، ص ص 207- 208).

يُلاحظ مما سبق تَعدُّد وتَنوُّع طرائق حساب درجة قطع الاختبارات، كما يُلاحَظ أهمية الأخذ بالحسبان خصائصها وكفاياتها الفنية والعملية التطبيقية، إضافةً إلى مدى ملاءَمتها لنوعية الاختبار المُستخدَم، وقد اعتمدت الباحثة في الدراسة الحالة الطريقة المبنية على المتوسط و الانحراف المعياري، في حساب درجات قطع المقاييس العيادية لرائز مينيسوتا MMPI-2، بسبب اعتماد هذه الطريقة في الدليل الأساسي للاختبار، حيث قامت الباحثة بحساب درجتي القطع العليا والدنيا لكل مقياس عيادي مُستخدم في البحث الحالي باستخدام المعادلتين الآتيتين:

- درجة القطع العليا= متوسط الدرجات + 2 انحراف معياري.
- درجة القطع الدنيا= متوسط الدرجات – 2 انحراف معياري.

الفصل الرابع

خطّة البحث وإجراءاته

أولاً: مقدمة.

ثانياً: منهج البحث.

ثالثاً: مجتمع البحث وعينته.

رابعاً: أدوات البحث.

خامساً: إجراءات التطبيق.

سادساً: المعالجات الإحصائية.

الفصل الرابع

خطة البحث وإجراءاته

أولاً: مقدمة:

يتناول الفصل الحالي إجراءات تنفيذ البحث ميدانياً، متضمناً منهج البحث، ومجتمعه وعينته، وخصائص العينة وحجمها وطريقة سحبها، كذلك أدوات البحث والتعريف بها، والمقاييس العيادية التي تم إجراء الدراسة عليها، حيث اكتفت الباحثة بإجراء الدراسة على أربعة فقط من المقاييس العيادية العشرة التي يتضمنها رائز مينيسوتا MMPI-2، وذلك لأن هذه المقاييس تُظهر فروقات أكبر في عدد البنود بين النموذج الأصلي الذي أعدّه هاثاواي وماكنللي، وبين النموذج البحثي الذي أعدّه هاريس ولنجوس، وبالتالي تُمكِّن من دراسة أثر الفرق في عدد البنود في الخصائص السيكومترية ودرجات القطع، كما يتناول الفصل الحالي أيضاً إجراءات التطبيق، والمعالجات الإحصائية المستخدمة للحصول على البيانات المطلوبة.

ثانياً: منهج البحث:

بعد الاطلاع على بعض الدراسات والبحوث السابقة، ومراجعة المناهج البحثية، توصلت الباحثة إلى أنَّ المنهج الملائم للبحث الحالي هو المنهج الوصفي التحليلي، وهو المنهج الذي يعتمد على دراسة الظاهرة كما توجد في الواقع ويهتم بوصفها وصفاً دقيقاً، ويُعبّر عنها تعبيراً كيفياً وكمياً، فالتعبير الكيفي يصف الظاهرة ويُبيّن خصائصها، بينما التعبير الكمي يعطي وصفاً رقمياً لمقدار الظاهرة أو حجمها ودرجات ارتباطها مع الظواهر المختلفة الأخرى (عبيدات، 1993، ص 219).

ثالثاً: مجتمع البحث وعينته:

1- <u>مجتمع البحث:</u>

تألّف مجتمع البحث من ثلاثة مجتمعات أصلية، وهي على النحو الآتي:

أ- **المجتمع الإحصائي لطلبة التعليم الجامعي:** ويتمثل بطلبة السنة الرابعة في جامعة دمشق، بكافة الاختصاصات النظرية والعملية، ووفقاً لإحصائيات مديرية شؤون الطلاب في وزارة التعليم العالي للعام الدراسي 2012/ 2013 م، فقد بلغ عدد هؤلاء الطلبة (9081) طالباً وطالبة في الأقسام التي تمَّ السحب منها، وهي: الأقسام الطبيّة، والتطبيقية، والتربوية، والعلوم، والآداب والعلوم الإنسانية، ممثلةً بذلك المجتمع الأصلي لطلبة التعليم الجامعي في السنة الرابعة.

ب- **مجتمع طلبة الماجستير:** ويتمثل بطلبة الماجستير من السنة الأولى والثانية في جامعة دمشق بكافة الاختصاصات النظرية والعملية، ووفقاً لإحصائيات مديرية شؤون الطلاب في وزارة التعليم العالي للعام الدراسي 2012/ 2013 م، فقد بلغ عدد هؤلاء الطلبة (1962) طالباً وطالبة في الاختصاصات التي تمَّ السحب منها، وهي: طب الأسنان، والصيدلة، واللغة الإنكليزية، والتاريخ، والجغرافيا، والهندسة المدنية، والفيزياء، والرياضيات، إضافةً إلى كافة أقسام التربية، ممثلةً بذلك المجتمع الأصلي لطلبة الماجستير.

ت- **مجتمع المضطربين نفسياً:**

- جرى تطبيق أداة البحث في مشفى ابن رشد للأمراض النفسية، ومركز الشهيد زهير حبي / العيادة النفسية /.
- استغرق التطبيق ما يقارب أربعة أشهر تضمنت الفترة الممتدة ما بين 27 كانون الثاني و 30 أيار عام 2013 م.
- تمثّل مجتمع البحث بجميع الحالات التي ترددت إلى العيادة النفسية في المشفى والمركز، و الحالات المقيمة في المشفى – السكن الداخلي – في تلك الفترة.

2- <u>عينة البحث:</u>

تألّفت عينة البحث من ثلاث مجموعات أساسية، وذلك لتمثيل خصائص العينة المراد دراسة أثرها في الخصائص السيكومترية ودرجات القطع، وهي على النحو الآتي:

أ- عينة طلبة التعليم الجامعي:

قامت الباحثة بسحب هذه العينة بطريقة العينة العشوائية، من طلبة السنة الرابعة في جامعة دمشق، من كافة الفروع والاختصاصات النظرية والعملية، حرصاً على تمثيل المجتمع الأصلي ككل، وقد جرى سحب العينة باتباع الخطوات الآتية:

- قامت الباحثة بتقسيم مجتمع الطلبة الجامعيين إلى خمسة أقسام هي:

1- العلوم الطبية 2- الآداب والعلوم الإنسانية 3- التربية

4- العلوم التطبيقية 5- العلوم

- أُعطيَت أسماء الكليات المُتضمنة في الأقسام السابقة أرقاماً، ثم تم سحب عدد من الكليات من كل قسم بشكل عشوائي.
- تراوحت نسب السحب بين (11.04%) و(15.45%)، وبلغت نسبة تمثيل عينة طلبة التعليم الجامعي للمجتمع الأصلي ككل (12.48%).
- بلغ عدد أفراد هذه العينة (1134) طالب وطالبة، تراوحت أعمارهم بين /21- 23/ سنة، بمتوسط قدره /22.3/ سنة، والجدول ذو الرقم 11 يبين توزع أفراد هذه العينة حسب الكليات،

كما يبين العدد الكلي للطلبة في كل قسم، والنسبة المئوية لعدد أفراد هذه العينة من عدد أفراد المجتمع.

ب- عينة طلبة التعليم ما بعد الجامعي /طلبة الماجستير/:

قامت الباحثة بسحب هذه العينة بطريقة العينة العشوائية أيضاً، من طلبة الماجستير في جامعة دمشق، بكافة الفروع والاختصاصات النظرية والعملية، باتباع الخطوات المذكورة أعلاه.

- تراوحت نسب السحب بين (10.34%) و (11.16%) من المجتمع الأصلي، وبلغت نسبة تمثيل عينة طلبة الماجستير للمجتمع الأصلي ككل (10.49%).

- بلغ عدد أفراد هذه العينة (206) طالب وطالبة، تراوحت أعمارهم بين /23- 35/ سنة، بمتوسط قدره /26.5/ سنة، والجدول ذو الرقم 12 يبين توزع أفراد هذه العينة حسب الاختصاصات، كما يبين العدد الكلي لطلبة الماجستير في كل اختصاص، والنسبة المئوية لعدد أفراد هذه العينة من عدد أفراد المجتمع.

ت- عينة المضطربين نفسياً:

- جرى سحب عينة المضطربين نفسياً من مشفى ابن رشد للأمراض النفسية، ومركز الشهيد زهير حبي /العيادة النفسية/، بشكل مقصود.
- شملت هذه العينة الحالات التي ترددت إلى العيادة النفسية في المشفى والمركز، والحالات المقيمة في المشفى – السكن الداخلي – في تلك الفترة، والتي كانت مشخصة بواحدة من الاضطرابات الآتية: (الفصام، الاكتئاب، توهم المرض، الانحراف السيكوباتي)، حيث بلغ عدد أفرادها /150/ حالة، تراوحت أعمارهم بين /21- 48/ سنة، بمتوسط قدره /32.17/ سنة.
- روعي في اختيارها الخصائص الآتية:

1- أن يكون المضطرب نفسياً قد شُخِّصت حالته بواحدٍ من الاضطرابات النفسية الآتية : الفصام، الاكتئاب، توهم المرض، الانحراف السيكوباتي، والذي يتضمن في بعض الحالات: الإدمان، أعراض البارانويا، أعراض الهوس الخفيف، سواء بالرجوع إلى السجلات الخاصة بالمريض أو استناداً إلى تشخيص الطبيب النفسي المشرف على حالته، أو كليهما معاً .

2- ألا يكون المضطرب النفسي قد دخل في طور الشفاء.

3- أن تكون حالة المضطرب نفسياً مستقرة، أي يستطيع التواصل مع الباحثة، إذ أن بعض الحالات لا تستطيع التواصل بسبب الهياج أو النعاس الناتج عن المهدئات ومنعكسات الأدوية، والجدول الآتي يبين توزع أفراد هذه العينة حسب نوع الاضطراب، ونسبة عدد الأفراد في كل اضطراب من العدد الكلي لأفراد عينة المضطربين نفسياً.

الجدول رقم 9: توزع أفراد عينة المضطربين نفسياً

نوع الاضطراب	عدد الأفراد	النسبة المئوية
الفصام	22	%14.7
الاكتئاب	29	%19.3
توهم المرض	30	%20
الانحراف السيكوباتي	69	%46
المجموع	150	%100

ويوضح الجدول الآتي الخصائص الديموغرافية لعينة البحث ككل:

الجدول رقم 10: الخصائص الديموغرافية لعينة البحث

أفراد العينة		متوسط العمر	الانحراف المعياري			
			مقياس الاكتئاب	مقياس البارانويا	مقياس الفصام	مقياس الهوس
الأسوياء	طلبة السنة الرابعة	22.3 سنة	5.61	4.46	9.71	4.97
	طلبة الماجستير	26.5 سنة	3.72	3.36	5.04	3.87
المضطربين نفسياً	عينة الاكتئاب	31.34 سنة	4.76	3.72	9.85	6.00
	عينة الفصام	34.95 سنة	3.14	3.37	2.52	4.41
	عينة توهم المرض	31.33 سنة	4.75	6.33	11.01	7.14
	عينة الانحراف السيكوباتي	31.07 سنة	3.92	2.96	7.93	4.93

الجدول رقم 11: توزع أفراد عينة طلبة السنة الرابعة حسب الكليات

	العلوم الطبية		الآداب والعلوم الإنسانية						التربية			
	الطب البشري	الصيدلة	اللغة العربية	اللغة الإنكليزية	علم الاجتماع	الجغرافيا	الفلسفة	التاريخ	علم النفس	معلم صف	الإرشاد النفسي	التربية الحديثة
العدد الكلي	580	342	1258	1889	454	360	303	396	245	423	219	350
عدد أفراد العينة	73	44	171	226	54	47	42	48	37	55	26	41
النسبة المئوية	%12.58	%12.86	%13.59	%11.96	%11.89	%13.05	%13.86	12.12 %	%15.10	%13	%11.87	%11.71
	العلوم التطبيقية (الهندسات)								العلوم			المجموع اكلي
	الهندسة المدنية	الهندسة المعلوماتية	الهندسة الزراعية	الهندسة الإلكترونية	الهندسة الميكانيكية	هندسة الطاقة الكهربائية	هندسة الغزل والنسيج		الجيولوجيا	علم الحياة	الرياضيات	
العدد الكلي	470	163	337	204	170	194	61		110	163	390	9081
عدد أفراد العينة	54	18	40	23	20	24	7		17	22	45	1134
النسبة المئوية	%11.48	%11.04	%11.86	%11.27	%11.76	%12.37	%11.47		%15.45	%13.49	%11.53	%12.48

الجدول رقم 12: توزع أفراد عينة طلبة الماجستير حسب الاختصاصات

	العلوم الطبية		الآداب والعلوم الإنسانية			العلوم التطبيقية	العلوم		التربية	المجموع الكلي
	طب الأسنان	الصيدلة	اللغة الإنكليزية	الجغرافيا	التاريخ	الهندسة المدنية	الفيزياء	الرياضيات	كافة الأقسام	
العدد الكلي	244	363	122	110	116	338	92	94	483	1962
عدد أفراد العينة	25	39	13	12	12	35	10	10	50	206
النسبة المئوية	%11.16	%10.74	%10.65	%10.90	%10.34	%10.35	%10.87	%10.64	%10.35	%10.49

رابعاً: أدوات البحث:

1- رائز مينيسوتا المتعدد الأوجه للشخصية MMPI-2:

يتكون رائز مينيسوتا النسخة الثانية من 567 بنداً يُجاب عنها من قبل الأفراد الذين تجاوزوا الثامنة عشر من العمر بـ " صح " أو " خطأ "، وتصحح بإعطاء درجة واحدة أو صفر. وتُعد الدرجة – على أي مقياس فرعي – مرتفعة إذا قابلت درجة تائية 65 أو أكثر، ومنخفضة إذا كانت أقل من الدرجة التائية 50. يتضمن هذا الرائز /81/ مقياساً هي:

- (3) مقاييس للصدق Validity Scales.
- (3) مؤشرات الصدق الإضافية Additional Validiyt Indicators.
- (10) مقاييس عيادية Clinical Scales.
- (31) مقياساً فرعياً للمقاييس العيادية Clinical Subscales.
- (15) مقاييس المحتوى Content Scales.
- (13) مقاييس تكميلية Supplementary Scales : (4) تقليدية Traditional و(9) إضافية Additional
- (6) مقاييس (مجموعات) للبنود الحرجة Critical Item Sets.

وفيما يلي وصفاً موجزاً للمقاييس العيادية التي يتضمنها الرائز:

(1) توهم المرض أو المراق Hypochondriasis (Hs): يكشف عن وجود تاريخ من الأعراض المرضية المُميزة لتوهم المرض، ويعبر عن تنوع في الأمراض والشكاوى البدنية يتراوح بين المحددة إلى العامة أو الغامضة.

(2) الاكتئاب Depression (D): يرتبط المقياس بالتفكير الانتحاري، ويعتبر مؤشراً لاستياء الفرد وعدم الرضا عن الحياة. ويتضمن مشاعر تتعلق بالتثبيط وعدم الاهتمام، وفقدان الأمل، والحساسية الزائدة والمشكلات البدنية وأمراضاً أو شكاوى تتضمن التأخر النفسحركي.

(3) الهستيريا Hysteria (Hy): وضع لتحديد الأشخاص الذين يستجيبون للضغوطات بطريقة هستيرية تتضمن اضطرابات حسية أو حركية دون وجود أساس عضوي لها.

(4) الانحراف السيكوباتي Psychopathic Deviate (Pd): يكشف درجة تشابه المفحوص بجماعة السيكوباثيين الذين تتمثل صعوباتهم الرئيسية في نقص الاستجابة الانفعالية العميقة، وعدم القدرة على الإفادة من الخبرة واللامبالاة بالمعايير الاجتماعية.

(5) الذكورة / الأنوثة Masculinity- Femininity (MF): يكشف الاهتمامات الأنثوية أو الذكرية، فارتفاع الدرجات لدى الذكور على هذا المقياس يشير إلى أنهم أكثر أنوثة من هؤلاء الذين هم في نفس الفئة العمرية من الذكور، وبنفس الطريقة لدى الإناث اللواتي يظهرن وجود اهتمامات ذكرية.

(6) البارانويا Paranoia (Pa): وضع لتحديد المرضى الذين يظهرون أعراضاً زورية تتعلق بمرجعية الأفكار، والشك، والشعور بالاضطهاد.

(7) الوهن النفسي أو السيكاثينيا Psychasthenia (Pt): صُمم المقياس ليقيس الزملة العصبية المعروفة بالسيكاثينيا، والتي ترتبط في فئة التصنيف الحالية باضطراب الوسواس القهري. ويغطي محتوى المقياس مدى واسعاً ومتنوعاً من الأعراض التي تتضمن مشكلات في التركيز، والأفكار الوسواسية، والأفعال القهرية، والقلق، ومشاعر الدونية والذنب، والمخاوف الشاذة والهموم.

(8) الفصام Schizophrenia (Sc): وضع لتحديد المرضى الذين يعانون من أعراض الفصام من حيث محتوى عمليات التفكير الغريبة، والإدراكات الخاصة، والاغتراب الاجتماعي، واضطرابات المزاج والسلوك، وصعوبات في التركيز.

(9) الهوس الخفيف Hypomania (Ma): وضع لتحديد المرضى الذين يظهرون أعراضاً للاستثارة الهوسية كالتكلف، والهياج، وطيران الأفكار ، والمزاج المرتفع، والنشاط الزائد سلوكياً ومعرفياً.

الانطواء الاجتماعي Social Introversion (Si): يكشف وجود مشكلات بالعلاقات الاجتماعية كالانسحاب الاجتماعي وقلة المشاركة في الأنشطة الجماعية، والتقدير المنخفض للذات، والحساسية وعدم الارتياح في المواقف الاجتماعية، و الخجل والشعور بالارتياح لدى الابتعاد عن الآخرين (عز، قيد النشر، ص ص 5 – 7).

قامت الدكتورة إيمان عز بتقنين هذا الرائز عام /2012/ (عز، قيد النشر)، على عينة من طلبة السنة الثالثة في جامعة دمشق بلغ عدد أفرادها /4376/ طالباً وطالبة بمتوسط عمر مقداره /21.8/ وانحراف معياري /1.4/، عدد الذكور فيها /2054/ طالباً، وعدد الإناث /2322/ طالبة.

حيث تمت دراسة الخصائص السيكومترية للرائز على النحو الآتي:

أ- **دراسة الصدق** : بالإضافة إلى التحقق من صدق المحتوى وصدق الترجمة، دُرِس صدق الرائز بأسلوبين هما:

- الصدق المحكي: تمّ تطبيق الرائز مع قائمة الأعراض على عينة من طلبة الجامعة قوامها /189/ طالباً، وأظهرت النتائج أن جميع مقاييس رائز مينسوتا كان لها ارتباط ذا دلالة إحصائية مع واحد أو أكثر من مقاييس قائمة الأعراض، باستثناء أربعة مقاييس لم يكن لها أي ارتباط مع أي مقياس من مقاييس القائمة هي: مقياس عدم اتساق الاستجابة المتغيرة VRIN، ومقياس عدم اتساق الاستجابة الصحيحة TRIN، ومقياس الإدمان الكامن APS، ومقياس البدء بالإدمان AAS. مما يعطي مؤشراً هاماًعلى صدق الرائز.

- صدق المجموعات المتناقضة: تمّ تطبيق الرائز كاملاً على مجموعتين ضمت الأولى /360/ مريضاً مراجعاً للمراكز والمشافي النفسية الحكومية والخاصة، ويتلقى العلاج النفسي الدوائي أو

النفسي، بينما تكونت المجموعة الثانية من /360/ طالباً جامعياً غير مراجع لأي طبيب أو معالج نفسي أو متلقي لأي شكل من أشكال العلاج النفسي.

بيّنت النتائج عدم وجود فروق ذات دلالة إحصائية بين مجموعة الأسوياء والمرضى في /10/ مقاييس فقط من أصل /47/ مقياساً. مما يعني قدرة الرائز في معظمه على التمييز بين الأسوياء والمرضى. وهذا بحد ذاته يقدم مؤشراً آخر على صدق الرائز بنسخته العربية السورية.

ب- **ثبات الرائز**: تمّ حساب ثبات الرائز بالأساليب التالية :

- الثبات بالإعادة: بينت النتائج أن معاملات الثبات بالإعادة كانت بمجملها دالة إحصائياً عند مستوى الدلالة /0.01/ وقد تراوحت بين /0.53 و0.82 / وهي معاملات ثبات مرتفعة.

- الثبات بالاتساق الداخلي والتنصيف: بينت النتائج أن معاملات ألفا – كرونباخ للمقاييس المدروسة تراوحت بين /0.44 و 0.89/، ومعاملات الثبات بالتنصيف تراوحت بين / 0.49 و 0.87 / وهي تقترب من مدى معاملات ألفا للرائز الأصلي التي تراوحت بين / 0.24 و 0.90/، وهي مؤشرات هامة على ثبات الرائز بنسخته العربية السورية.

كما تمّ بناء معايير محلية تمثل الإطار المرجعي للحكم على الدرجات، وإعطاء وصف أو تقويم للمفحوص وفقاً لها بالسواء أو الاضطراب، حيث جرى بناء جداول معايير سورية محلية لكل المقاييس المدروسة – أحدها خاص بالذكور والآخر خاص بالإناث – تتمثل بالدرجات التائية التي تحول الدرجات الخام وفقاً لها، كما تم تحديد نقطتي القطع اللتين تفصلان بين السواء والاضطراب بالدرجتين التائيتين/65/ و/50/ (المرجع السابق، ص ص11 – 25).

2- <u>قائمة الأعراض المعدلة :</u>

تُعد قائمة الأعراض/Symptom Checklist/90 التي تعرف اختصاراً ب (SCL-90) من أكثر القوائم استخداماً في العيادات والمشافي النفسية الأمريكية، وذلك لسهولة تطبيقها وقدرتها على إعطاء صورة سريعة عن الاضطراب الذي قد يعانيه المريض بوقتٍ قصير. ويمكن تطبيقها من قبل الممرضة أو الطبيب النفسي، وهناك أكثر من /1000/ دراسة أجريت على هذه القائمة للتحقق من صدقها وثباتها وفوائد استخدامها.

تتألف هذه القائمة من /90/ بنداً، يُجاب عن كل منها بأحد الخيارات التالية: "أبداً، قليلاً، أحياناً، كثيراً، دائماً " ضمن مدى من الدرجات يتراوح ما بين /0-4/، وتصلح للتطبيق الفردي أو الجماعي على الأفراد الذين تجاوزوا عمر/13 / سنة، وتصحح إما يدوياً أو عن طريق الحاسوب، وغالباً ما يستغرق التطبيق مدة تتراوح بين /12-15/ دقيقة.

يجري تصحيح وتفسير الدرجات ضمن تسعة مقاييس أولية للأعراض Primary Symptom Scales، وثلاثة مؤشرات للمعاناة Global Indices Distress، يمكن شرحها كما يلي:

أولاً: المقاييس الأولية للأعراض: Primary Symptom Scales:

(1)–الجسدنة: Somatization: يتكون هذا المقياس من /12/ بنداً، ويعكس المعاناة الناتجة عن الضعف الجسدي. تتركز الشكاوى الجسدية في الأجهزة المعوية، والقلبية، والتنفسية، وأوجاع وآلام العضلات، وقلق دائم على الصحة. ومثال تلك البنود: – آلام في الصدر أو القلب.

(2)– الوسواس القهري: Obsessive – Compulsive (O- C): يتكون هذا المقياس من /10/ بنود، ويعكس التركيز على أفكار واندفاعات ومخاوف لا يستطيع الشخص التخلص منها، وأفعال قهرية في هيئة طقوس حركية مستمرة أو دورية. ومثال تلك البنود: – تكرار أفعال مزعجة لا يمكنك التخلص منها.

(3)–حساسية العلاقات بالآخرين: Interpersonal Sensitivity (I –S): يتكون هذا المقياس من /9/ بنود، ويركز على الشعور بالنقص وعدم الكفاءة مقارنةً بالناس الآخرين، ومن خصائص هؤلاء الأشخاص عدم ارتياح في أثناء اللقاءات الاجتماعية، والشعور بالخجل والاستخفاف بالذات، إضافةً إلى تفسير سلوكيات الآخرين على نحو شخصي. ومثال لتلك البنود: – الشعور بالنقد من الآخرين.

(4)–الاكتئاب: Depression (DEP): يتكون هذا المقياس من /13/ بنداً، تعكس الاكتئاب العيادي الذي يتسم بانخفاض في المزاج والطاقة، وقلة النشاط، وفقدان الأمل، والإحساس بالذنب، والانعزال عن أنشطة الحياة اليومية والتفكير بالانتحار. ومثال لتلك البنود: – الشعور بالخمول وقلة النشاط.

(5)- القلق: Anxiety (ANX): يتكون هذا المقياس من /9/ بنود، وتعكس السمات العامة للقلق كالعصابية والرجفة التي تُعد محددات لنوبة الهلع والشعور بالخوف والرهاب. ومثال تلك البنود: – العصبية أو الرجفة من الداخل.

(6)– العدائية: Hostility (HOS): يتكون هذا المقياس من /6/ بنود، تعكس الأفكار والمشاعر والأفعال التي تُعد من مميزات الحالة السلبية للغضب، كما تعكس الامتعاض والعدوان والنزق والغضب الشديد. ومثال لتلك البنود: – الشعور بسرعة المضايقة أو النزق.

(7)– القلق الرهابي: Phobic Anxiety (PHOB): يتكون هذا المقياس من /7/ بنود، ويُعرّف القلق الرهابي في هذه القائمة بأنه استجابة خوف مستمرة من شخص ما، أو مكان ما، أو شيء ما، أو موقف ما، وهذه الاستجابة غير منطقية وغير ملائمة للمثير، وتؤدي إلى سلوك تجنبي أو هروبي، وتركز البنود هنا على رهاب الساح أو الخلاء. ومثال لتلك البنود: – الشعور بالخوف من الأماكن المفتوحة أو الشوارع.

(8)– الأفكار البارانوئية: Paranoid Ideation (PAR): يتكون هذا المقياس من /6/ بنود، تعكس السلوكات البارانوئية، كعدم القدرة على تنظيم الأفكار، وتُعد العدائية والشكوى والخوف من

الضياع والهلوسة الانعكاسات الأولية لهذا الاضطراب. ومثال لتلك البنود: – وجود أفكار أو معتقدات لديك لا تخطر على بال أحد.

(9)– الذهانية: Psychoticism (PSY): يتكون هذا المقياس من /10/ بنود، تعكس العزلة ونمط الحياة الفصامي والهلاوس التي تُعد أعراضاً أولية للفصام. ومثال لتلك البنود: – الشعور بالوحدة حتى في وجود الآخرين.

ثانياً: البنود الإضافية: Additional Items:

هناك سبعة بنود من هذه القائمة لا تندرج تحت أي من المقاييس الأولية التسعة السابقة الذكر، لكنها مشبعة على المقاييس جميعها، ولها قيمتها العيادية.

ومثال لتلك البنود: – ضعف الشهية للطعام. – فرط الطعام.

ثالثاً: المؤشرات العامة: Global Indices:

هناك ثلاثة مؤشرات للمعاناة صُممت مع قائمة الأعراض، وهي: مؤشر الشدة العامة Global Severity Index (GSI) ، ومؤشر معاناة الأعراض الإيجابية Positive Symptom Distress Index (PSDI)، ومجموع الأعراض الإيجابية Positive Symptom Total (PST)، وعمل كل من هذه المؤشرات هو معرفة مستوى معاناة الفرد النفسية أو عمقها، بوساطة الدرجة التي يحصل عليها المفحوص في واحد من هذه المؤشرات، التي تعد مهمة ومفيدة في تقييم الصورة العيادية، كما تُستخدم خلاصة للقائمة.

قامت الدكتورة إيمان عز بإعداد هذه القائمة للبيئة السورية، كما قامت الدكتورة هيفاء البقاعي بدراسة خصائصها السيكومترية باعتبارها أداة رئيسية في دراستها واعتمادها محكاً لرائز روتر لتكملة الجمل الناقصة (ر. ت. ج. ن) على عينة عشوائية من طلبة المرحلة الثانوية، وعينة من الطلبة الجامعيين، حيث بلغ عدد أفراد العينة الكلي /113/ طالباً وطالبة.

جرى التحقق من ثبات القائمة باستخدام طريقتين هما: الاتساق الداخلي والثبات بالإعادة، تراوحت معاملات الاتساق الداخلي بين /0.56 – 0.80/ بمتوسط قدره /0.67/، وهي تدل على مؤشرات جيدة، في حين تراوحت معاملات الثبات بالإعادة /0.61 – 0.80/بمتوسط قدره /0.74/ وهي تدل أيضاً على مؤشرات ثبات جيدة.

كما جرى التحقق من صدق القائمة باستخدام طريقتين هما:

- صدق المجموعات الطرفية: حيث جرى التحقق من هذا النوع من الصدق بواسطة اختيار أعلى وأدنى15% من الدرجات عند كل مقياس من مقاييس قائمة الأعراض، وذلك عند الدرجات الخاصة بعينة الصدق والثبات السابقة الذكر، بعد ذلك جرى حساب الفرق بين الدرجات بواسطة اختبار "ت" ستيودنت عند عينة الذكور والإناث معاً. وقد دلّت النتائج على

وجود فروق دالة إحصائياً عند كل مقياس من مقاييس القائمة، مما يدل على القدرة التمييزية الجيدة لهذه القائمة.

- صدق المجموعات المتضادة: جرى حساب الفروق بواسطة اختبار "ت" ستيودنت بين العينة السوية وعينة المضطربين نفسياً عند الذكور والإناث معاً، وقد دلّت النتائج على وجود فروق دالة إحصائياً بين الأسوياء والمرضى عند جميع المقاييس الفرعية لقائمة الأعراض (البقاعي، 2007، ص ص 93 – 99).

خامساً: إجراءات التطبيق:

– تم تطبيق رائز مينيسوتا MMPI-2 بشكل جماعي، وفق التعليمات الخاصة بتطبيقه، على عينتي طلبة التعليم الجامعي، وطلبة الماجستير.

– في حين جرى تطبيق الرائز بشكل فردي على عينة المضطربين نفسياً، والسبب في ذلك هو أن بعض أفراد هذه العينة لا يعرف القراءة والكتابة، وبعضهم الآخر لا يستطيع التركيز بسبب الأدوية النفسية التي تؤثر في دقة التركيز، وأيضاً لتلافي الشعور بالملل لديهم والذي قد يؤدي في بعض الأحيان إلى إفشال عملية التطبيق.

– كما تم تطبيق قائمة الشطب بشكلٍ جماعي أيضاً، على عينة الصدق المحكي والبالغ عددها /567/ طالب وطالبة، وذلك لدراسة أثر التفاعل بين طول الاختبار وخصائص العينة في معاملات الصدق المحكي.

وبعد تصحيح البيانات وفقاً لمفتاح التصحيح الخاص بكل رائز، قامت الباحثة بالإجراءات الآتية للإجابة عن تساؤلات البحث:

- **لدراسة أثر الفرق في طول المقياس في الخصائص السيكومترية ودرجات القطع:**

قامت الباحثة بدراسة الفرق في عدد البنود بين كل مقياس من المقاييس العيادية في النموذج الأصلي للرائز الذي أعدّه هاثاواي وماكنللي، ونظيره في النموذج البحثي الذي أعدّه هاريس ولنجوس، وذلك لمعرفة المقاييس العيادية التي تُوضح فروقاً أكبر في عدد البنود بين النموذجين، وبالتالي تُمكِّن من دراسة أثر هذه الفروق في الخصائص السيكومترية ودرجات القطع، أما المقاييس العيادية التي كان فيها عدد البنود متقارباً بين النموذجين، فقد استبعدتها الباحثة لأنها لا تخدم أغراض الدراسة، والجدول التالي يبين المقاييس التي اعتمدتها الباحثة بناءً على ذلك، كما يبين عدد بنود المقاييس العيادية في النموذجين، والنسبة المئوية للفرق في عدد البنود.

الجدول رقم 13: عدد بنود المقاييس العيادية في النموذجين المُستخدمين في الدراسة

المقياس العيادي	عدد البنود عند هاثاواي وماكنللي	عدد البنود عند هاريس ولنجوس	الفرق في عدد البنود	النسبة المئوية للفرق
الاكتئاب	57	49	8	%14.03
البارانويا	40	34	6	%15
الفصام	78	68	10	%12.82
الهوس الخفيف	46	34	12	%26.08

- **ولدراسة أثر الاختلاف في خصائص العينة في الخصائص السيكومترية ودرجات القطع:**

قامت الباحثة بسحب واختيار ثلاث عينات، لتشمل خصائص العينة المراد دراستها، وتتمثل هذه الخصائص بالمتغيرات الآتية:

أ- متغير المرحلة الدراسية: عينة مرحلة التعليم الجامعي /طلبة السنة الرابعة /، وعينة مرحلة ما بعد التعليم الجامعي /طلبة الماجستير/.

ب-متغير الحالة الصحية: عينة الأسوياء، وعينة المضطربين نفسياً / وتشمل حالات الفصام، الاكتئاب، توهم المرض، الانحراف السيكوباتي/.

ت-متغير حجم العينة: حيث جرى سحب واختيار أحجام العينات على أساس حسابها كنسب مئوية من عدد بنود رائز مينيسوتا MMPI-2 ككل أي (567) بنداً، ذلك لأن رائز مينيسوتا وإن كان يتضمن مقاييس عيادية عديدة إلا أنه اختبار كلي ويطبق غالباً ككل، وبنود مقاييسه العيادية موزعة ضمن البنود ال370 الأولى، لذا اعتمدت الباحثة سحب العينات على أساس حسابها كنسب مئوية من عدد البنود الكلي للرائز، وبذلك كانت أحجام العينات المدروس أثرها في الخصائص السيكومترية على النحو الآتي:

- حجم العينة ما يقارب 50% من عدد بنود الرائز، وبالتالي حجمها يساوي (284) طالباً وطالبة.
- حجم العينة ما يقارب 75% من عدد بنود الرائز، وبالتالي حجمها يساوي (425) طالباً وطالبة.
- حجم العينة 100% من عدد بنود الرائز، وبالتالي حجمها يساوي (567) طالباً وطالبة.
- حجم العينة ما يقارب 150% من عدد بنود الرائز، وبالتالي حجمها يساوي (851) طالباً وطالبة.
- حجم العينة 200% من عدد بنود الرائز، وبالتالي حجمها يساوي (1134) طالباً وطالبة.

ذلك بالنسبة لعينة الطلبة الجامعيين والبالغ عدد أفرادها 1134 طالباً وطالبة، أما بالنسبة لعينتي طلبة الماجستير البالغ عدد أفرادها(206) طلاب وطالبات، والمرضى النفسيين البالغ عدد أفرادها (150)، فقد اكتفت الباحثة بدراسة أثر التفاعل بين طول الرائز ومتغيري المرحلة الدراسية والحالة الصحية لهاتين العينيتين في الخصائص السيكومترية ودرجات القطع.

وقد قامت الباحثة بدراسة مدى التوزع الطبيعي لهذه العينات ذات الأحجام المختلفة، ولهذا الغرض جرى حساب بعض من مقاييس التشتت ومقاييس النزعة المركزية كالمتوسط والانحراف المعياري، استناداً إلى منحنى التوزيع الطبيعي، كما تم حساب الالتواء والتفرطح لهذه العينات ومن خواص منحنى التوزيع الطبيعي أن تكون الالتواءات فيه معدومة أو قريبة من الصفر.

وفيما يلي الإحصاءات الوصفية لهذه العينات ذات الأحجام المختلفة:

الجدول رقم 14: الإحصاءات الوصفية للعينات ذات الأحجام المختلفة

نوع المقياس العيادي	حجم العينة	المتوسط	الانحراف المعياري	التباين	الخطأ المعياري	الالتواء	الخطأ المعياري للالتواء	التفرطح	الخطأ المعياري للتفرطح
الاكتئاب	284	28.186	6.030	36.36	0.357	-0.058	0.145	-0.323	0.288
	425	28.371	5.828	33.97	0.282	0.098	0.118	-0.265	0.236
	567	27.321	5.881	34.59	0.247	0.206	0.103	-0.244	0.205
	851	26.199	5.824	33.92	0.199	0.325	0.084	-0.068	0.167
	1134	25.478	5.752	33.09	0.170	0.355	0.073	0.005	0.145
البارانويا	284	14.316	4.682	21.92	0.277	0.293	0.145	-0.344	0.288
	425	14.343	4.901	24.02	0.237	0.305	0.118	-0.166	0.236
	567	13.791	4.826	23.29	0.202	0.359	0.103	-0.082	0.205
	851	13.383	4.724	22.31	0.161	0.436	0.084	0.074	0.167
	1134	13.037	4.608	21.23	0.136	0.417	0.073	0.097	0.145
الفصام	284	28.609	10.728	115.10	0.636	0.038	0.145	-0.173	0.288
	425	28.049	11.41	130.22	0.553	0.133	0.118	-0.291	0.236
	567	26.312	11.32	128.15	0.475	0.248	0.103	-0.268	0.205
	851	25.069	10.70	114.50	0.366	0.421	0.084	-0.109	0.167
	1134	24.160	10.12	102.61	0.300	0.495	0.073	0.121	0.145
الهوس الخفيف	284	21.728	4.669	21.803	0.277	-0.041	0.145	-0.175	0.288
	425	21.788	5.100	26.016	0.247	-0.258	0.118	0.301	0.236
	567	21.384	5.104	26.057	0.214	-0.199	0.103	0.226	0.205
	851	20.960	5.121	26.227	0.175	-0.024	0.084	0.002	0.167
	1134	20.528	4.999	24.996	0.148	-0.017	0.073	0.018	0.145

يُلاحظ من الجدول السابقة أن قيم المتوسط والانحراف المعياري وكذلك الخطأ المعياري كانت متقاربة بين العينات ذات الأحجام المختلفة بالنسبة للمقياس العيادي الواحد، كما يُلاحظ أن جميع الالتواءات كانت قريبة من الصفر، حيث وقعت هذه الالتواءات ضمن الحدود الطبيعية التي تتراوح بين (+1، -1)، ومعامل التفرطح داخل المجال من (+3، -3)، مما سبق يمكن القول أن درجات المقاييس تتوزع بين أفراد العينة توزعاً طبيعياً اعتدالياً.

سادساً: المعالجات الإحصائية:

تمَّ استخدام برنامج الحزم الإحصائية SPSS لتحليل البيانات واستخراج وحساب كلٍ مما يلي:

- معاملات الثبات بالإعادة باستخدام معامل ارتباط بيرسون.
- معامل ألفا كرونباخ لحساب ثبات الاتساق الداخلي لدرجات المقاييس العيادية الأربعة لنوعي بنود الاختبار الطويل والقصير.
- معاملات الصدق المحكي مع محك قائمة الأعراض، لكل من نوعي المقاييس الطويلة والقصيرة.
- اختبار الفروق بين متوسطات درجات المقاييس العيادية الأربعة لكل من مجموعتي الأسوياء والمضطربين نفسياً، ومجموعتي طلبة الماجستير والطلبة الجامعيين لنوعي بنود الاختبار الطويل والقصير باستخدام تحليل التباين الأحادي (ANOVA).
- استخدام الاختبار الإحصائي Dunnett's C لإجراء عمليات المقارنة البعدية بين فروق متوسطات مجموعات العينات السابقة على كل من نوعي بنود الاختبار الطويل والقصير.
- قيمة إحصاء إيتا تربيع لحساب حجم التأثير للفروق بين المتوسطات حيث تحسب هذه القيمة وفقاً للمعادلة الآتية:

قيمة إحصاء إيتا تربيع= مجموع المربعات البينية / المجموع الكلي للتربيع (بالانت، 2006، ص 246)

- حساب الفرق بين متوسطات درجات القطع للمقاييس الطويلة والمقاييس القصيرة باستخدام الاختبار الإحصائي T-student.
- كما تم استخدام الاختبار الإحصائي M المقترح من قبل Hakstian and Whalen والذي يتبع توزيع كاي تربيع بدرجات حرية تساوي (عدد معاملات الثبات−1) وذلك للكشف عن الفروق بين معاملات ثبات الاتساق الداخلي لكل مقياس عيادي من المقاييس العيادية الأربعة لكل من نوعي بنود الاختبار الطويل والقصير ولكل حجم من أحجام العينات المدروسة علماً أن قيمة M تحسب لهذه الحالات وفق المعادلتين الآتيتين:

- **عدد بنود متساوٍ للاختبارات وأحجام عينات مختلفة:**

$$M = \frac{J-1}{18J}\left\{\sum_{K-1}^{K} B_K - \frac{\left[\sum_{k=1}^{k} B_k(1-r_{ak})^{-1/3}\right]^2}{\sum_{k=1}^{k} B_k(1-r_{ak})^{-2/3}}\right\}$$

Where

$$B_k = (9n_k - 11)^2/(n_k - 1)$$

- **عدد بنود مختلف وأحجام عينات متساوية:**

$$M = \frac{(9n-11)^2}{n-1}\left\{\sum_{K-1}^{K} C_K - \frac{\left[\sum_{k=1}^{k} c_k(1-r_{ak})^{-1/3}\right]^2}{\sum_{k=1}^{k} c_k(1-r_{ak})^{-2/3}}\right\}$$

حيث: $C_k = (J_k - 1)/18\,J_K$

حيث أنَّ J : عدد فقرات الاختبار

K: رقم النموذج

r_{ak} : معامل الارتباط للنموذج k

n_k : عدد أفراد النموذج k (Hakstian & Whalen, 1976, P. 224)

- واستخدام الاختبار الإحصائي V الذي وضعه Hays هيز ، والذي يتبع توزيع كاي تربيع بدرجات حرية تساوي (عدد معاملات الارتباط–1) ، والخاص بمعرفة دلالة الفروق بين عدة معاملات ارتباط بيرسون، أي بين معاملات ارتباط الاختبار بالإعادة لكل من نوعي بنود الاختبار الطويل والقصير ولكل حجم من أحجام العينات المدروسة.

وكذلك بين معاملات الارتباط المحكي مع قائمة الأعراض لكل من نوعي بنود الاختبار الطويل والقصير ولأحجام العينات المختلفة المدروسة، بعد تحويل معاملات ارتباط إلى قيم Z بطريقة فيشر :

$$V = \sum_1^K (ni-3)(Z_i - U)^2$$

وتحسب قيمة U من المعادلة الآتية:

$$U = \frac{\left[\sum_1^k (n_{i-3})z_i\right]}{\sum_1^k (n_{i-3})}$$

حيث n_i : عدد أفراد العينة في النموذج (i)

و z_i : قيمة فيشر المقابلة لمعامل الارتباط

و i : رقم النموذج (Hays, 1977, P. 664)

- واستخدام الاختبار الإحصائي Z لفحص الفروق بين معاملات الارتباط بطريقة الإعادة ومعاملات الارتباط مع قائمة الأعراض لكل من نوعي بنود الاختبار الطويل والقصير، والذي يحسب بعد تحويل معاملات الارتباط إلى قيم Z بطريقة فيشر وذلك بالمعادلة الآتية:

$$Z = \frac{z_1 - z_2}{\sqrt{\left(\frac{1}{n_{1-3}}\right)+\left(\frac{1}{n_{2-3}}\right)}}$$

حيث z_1 و z_2 قيمة فيشر للمعاملين

و n_1 و n_2 عدد أفراد كل من العينتين (علام، 2005، ص221)

وتحول معاملات الارتباط إلى قيم Z بطريقة فيشر وفق المعادلة الآتية:

$$Z_i = \frac{1}{2}\ell n\left[\frac{(1+r_i)}{(1-r_i)}\right]$$ (علام، المرجع السابق، ص156).

يُلاحظ من الإجراءات والمعالجات الإحصائية السابقة، أنه تمّ الاعتماد على القوانين الرياضية الخاصة بمنهج ما وراء التحليل الإحصائي في حساب معظم معاملات الدراسة الحالية، للإجابة عن التساؤلات البحثية والوصول إلى النتائج المطلوبة، ذلك لأن هذا المنهج وكما تمّ الإشارة سابقاً، يهدف إلى الوصول إلى رؤيا تكاملية كمية للنتائج حول موضوع معين، حيث يساعد في دراسة أثر التفاعل بين متغيرين أو أكثر من متغيرات الدراسة في بقيّة المتغيرات، وبالتالي يُتيح دراسة أثر التفاعل بين طول الاختبار وخصائص العينة في الخصائص السيكومترية للرائز من جهة، ودرجات قطعه من جهة ثانية، وهو الهدف الأساسي للبحث الحالي.

الفصل الخامس

نتائج البحث وتفسيرها

أولاً: الإجابة عن السؤالين الأول والثاني.

ثانياً: الإجابة عن السؤالين الثالث والرابع.

ثالثاً: الإجابة عن السؤالين الخامس والسادس.

رابعاً: الإجابة عن السؤال السابع.

خامساً: الإجابة عن السؤال الثامن.

سادساً: الإجابة عن السؤال التاسع.

سابعاً: الإجابة عن السؤال العاشر.

ثامناً: الإجابة عن السؤال الحادي عشر.

تاسعاً: مناقشة النتائج.

عاشراً: مقترحات البحث.

الفصل الخامس

نتائج البحث وتفسيرها

يتضمن الفصل الحالي الإجابة عن أسئلة البحث المتعلقة ببيان أثر التفاعل بين طول الاختبار وخصائص العينة، في الخصائص السيكومترية ودرجات القطع لاختبار الشخصية المتعدد الأوجه مينيسوتا MMPI-2، باستخدام بعض القوانين الرياضية الخاصة بمنهج ما وراء التحليل الإحصائي، وقد اكتفت الباحثة بالإجابة عن أسئلة البحث بالنسبة للمقاييس العيادية الأربعة: الاكتئاب، البارانويا، الفصام، والهوس الخفيف، لأن هذه المقاييس تُظهر فروقات أكبر في عدد البنود بين النموذج الأصلي للرائز والنموذج البحثي الذي أعدّه هاريس ولنجوس، وبالتالي تُمكِّن من دراسة أثر الفرق في عدد البنود على الخصائص السيكومترية ودرجات القطع، أما بالنسبة لخصائص العينة فقد تمّ دراستها في ضوء ثلاثة متغيرات أساسية هي: متغير المرحلة الدراسية، والحالة الصحية، وحجم العينة. وفيما يأتي عرض لهذه النتائج.

أولاً: الإجابة عن السؤالين الأول والثاني:

السؤال الأول: هل توجد فروق ذات دلالة إحصائية بين معاملات ثبات درجات الاختبار بالإعادة باختلاف طول الاختبار؟

السؤال الثاني: هل توجد فروق ذات دلالة إحصائية بين معاملات ثبات درجات الاختبار بالإعادة باختلاف حجم العينة؟

للإجابة عن السؤالين الأول والثاني تمَّ حساب معاملات ثبات الإعادة للمقاييس العيادية الأربعة: الاكتئاب، البارانويا، الفصام والهوس الخفيف، ببنودها الطويلة أي مقاييس هاثاواي وماكنللي، وبنودها القصيرة أي مقاييس هاريس ولنجوس، لأحجام العينات: (100%) و(75%) و(50%) من عدد بنود رائز مينيسوتا، والجدول الآتي يوضح قيم هذه المعاملات.

الجدول رقم 15: قيم معاملات ثبات الإعادة للمقاييس العيادية الطويلة والقصيرة لأحجام العينات المختلفة

المقياس العيادي / حجم العينة	الاكتئاب		البارانويا		الفصام		الهوس الخفيف	
	طويل	قصير	طويل	قصير	طويل	قصير	طويل	قصير
(284) %50	0.608**	0.572**	0.754**	0.675**	0.609**	0.575**	0.632**	0.504**
(425) %75	0.597**	0.559**	0.753**	0.674**	0.601**	0.579**	0.624**	0.509**
(567) %100	0.600**	0.566**	0.738**	0.654**	0.614**	0.584**	0.623**	0.505**

يتبين من الجدول السابق هبوط قيم معاملات ثبات الإعادة في المقاييس القصيرة عن مثيلاتها في المقاييس الطويلة، ولاختبار دلالة الفروق بين هذه المعاملات للمقاييس الطويلة والقصيرة* عند كل حجم من أحجام العينات الثلاث، تم تطبيق الاختبار الاحصائي (V) بعد تحويل معاملات الارتباط بالإعادة إلى قيم (Z) بطريقة فيشر، والجدول الآتي يبين نتائج الاختبار:

الجدول رقم 16: نتائج الاختبار الإحصائي (V) لاختبار الفروق في معاملات ثبات الإعادة بين المقاييس الطويلة والقصيرة عند أحجام العينات المختلفة

حجم العينة	المقياس العيادي	(R) للاختبار الطويل	(R) للاختبار القصير	قيمة V	كاي مربع **	قيمة المعنوية	القرار
284	الاكتئاب	0.608	0.572	0.436	3.841	0.51	غير دال
	الفصام	0.609	0.575	0.392	3.841	0.53	غير دال
	البارانويا	0.754	0.675	3.766	3.841	0.05	غير دال
	الهوس	0.632	0.504	5.166	3.841	0.02	دال
425	الاكتئاب	0.597	0.559	0.697	3.841	0.40	غير دال
	الفصام	0.601	0.579	0.243	3.841	0.62	غير دال
	البارانويا	0.753	0.674	5.603	3.841	0.02	دال
	الهوس	0.624	0.509	6.195	3.841	0.01	دال
567	الاكتئاب	0.600	0.566	0.748	3.841	0.39	غير دال
	الفصام	0.614	0.584	0.617	3.841	0.43	غير دال
	البارانويا	0.738	0.654	7.567	3.841	0.01	دال
	الهوس	0.623	0.505	8.527	3.841	0.00	دال

**** عدد درجات الحرية لاختبار كاي مربع هو (1)**

***المقصود بالمقاييس الطويلة: بنود المقاييس العيادية في النموذج الأصلي للرائز الذي أعدّه هاثاواي وماكنللي.**

والمقصود بالمقاييس القصيرة: بنود المقاييس العيادية في النموذج البحثي الذي أعدّه هاريس ولنجوس.

- يتبين من الجدول السابق ما يلي:

عدم وجود فروق دالة إحصائياً بين معاملي ثبات الإعادة لكلٍ من مقياسي الاكتئاب والفصام الطويل والقصير عند أحجام العينات الثلاث. كما يتبين عدم وجو فرق دال إحصائياً بين معاملي ثبات مقياس البارانويا الطويل والقصير عند حجم العينة (284) ولو أنه قارب الوصول إلى مستوى الدلالة الإحصائية، في حين كانت هناك فروق دالة إحصائياً بين معاملي ثبات المقياس الطويل والقصير لصالح معامل الثبات الطويل عند حجمي العينة (425) و(567). كما نلاحظ وجود فروق دالة إحصائياً بين معاملي ثبات الإعادة لمقياس الهوس الطويل والقصير عند أحجام العينات الثلاث، وهذه الفروق بمجملها لصالح معامل ارتباط المقياس الطويل، وقد يُعزى ذلك إلى زيادة نسبة الفرق في عدد البنود بين المقياسين الطويل والقصير في كلٍ من مقياسي البارانويا والهوس الخفيف، حيث تصل هذه النسبة في مقياس البارانويا إلى (15%)، وفي مقياس الهوس الخفيف إلى(26.08%).

ونلاحظ أنه مع انخفاض حجم العينة تقل قيم الفرق بين معاملي الثبات، وبالتالي تغدو الفروق بين معاملي ثبات المقياس الطويل والقصير غير دالة إحصائياً، مما يعني أن حجم العينة وعدد البنود يؤثران معاً في ارتفاع معامل الثبات وانخفاضه، وبالتالي كلما قل حجم العينة لابد أن يكون عدد البنود أكثر.

ولمزيد من التأكد تمَّ استخدام الاختبار الإحصائي (Z) عند مستوى الدلالة (0.05) لفحص الفروق بين معاملات الثبات بالإعادة، والجدول الآتي يبين قيم (Z):

الجدول رقم 17: قيم الاختبار (Z) لفحص الفروق في معاملات الثبات بين المقاييس الطويلة والقصيرة عند أحجام العينات المختلفة

المقياس العيادي	حجم العينة	قيمة معامل الارتباط بالإعادة	Z الجدولية	Z المحسوبة	القرار عند مستوى الدلالة (0.05)
الاكتئاب	567	(R) الطويل0.600	1.96	0.86533	غير دال
		(R) القصير0.566			
	425	(R) الطويل0.597		0.83523	غير دال
		(R) القصير0.559			
	284	(R) الطويل0.608		0.66072	غير دال
		(R) القصير0.572			
الفصام	567	(R) الطويل0.614	1.96	0.78599	غير دال
		(R) القصير0.584			
	425	(R) الطويل0.601		0.49378	غير دال
		(R) القصير0.579			

	284	(R) الطويل0.609		0.62625	غير دال
		(R) القصير0.575			
البارانويا	567	(R) الطويل 0.738	1.96	2.75094	دال
		(R) القصير0.654			
	425	(R) الطويل 0.753		2.36709	دال
		(R) القصير0.674			
	284	(R) الطويل0.754		1.94071	غير دال
		(R) القصير0.675			
الهوس	567	(R) الطويل0.623	1.96	2.92024	دال
		(R) القصير0.505			
	425	(R) الطويل0.624		2.48908	دال
		(R) القصير0.509			
	284	(R) الطويل0.632		2.27309	دال
		(R) القصير0.504			

يتبين من الجدول السابق ما يلي:

عدم وجود فروق دالة إحصائياً بين معاملي ثبات الإعادة لكلٍ من مقياسي الاكتئاب والفصام الطويل والقصير عند أحجام العينات الثلاث. كما يتبين عدم وجود فروق دالة إحصائياً بين معاملي ثبات مقياس البارانويا الطويل والقصير عند حجم العينة (284) على الرغم من أن قيمة (Z) المحسوبة بلغت (1.9407) أي أنها قاربت الوصول إلى مستوى الدلالة، في حين كانت هناك فروق دالة إحصائياً بين معاملي ثبات هذا المقياس الطويل والقصير عند حجمي العينة (425) و(567) لصالح معامل الثبات الطويل. كما يُلاحظ وجود فروق دالة إحصائياً بين معاملي ثبات الإعادة لمقياس الهوس الطويل والقصير عند أحجام العينات الثلاث، وهذه الفروق بمجملها لصالح معامل ارتباط المقياس الطويل.

ويُلاحظ اتفاق قيم الاختبار الاحصائي (V) مع قيم الاختبار الإحصائي (Z).

وللإجابة عن السؤال الثاني فيما إذا كانت هنالك فروق ذات دلالة إحصائية بين معاملات ثبات الإعادة لدرجات الاختبار باختلاف حجم العينة:

استخدمت الباحثة الاختبار الاحصائي (V) لاختبار الفروق في معاملات الثبات بين كل حجمي عينة من العينات الثلاث، ولكلٍ من الاختبارين الطويل والقصير على حدة، بعد تحويل معاملات الارتباط إلى قيم Z بطريقة فيشر، والجدولان رقم/17/ و/18/ يوضحان نتائج هذا الاختبار:

الجدول رقم 18: نتائج الاختبار الإحصائي (V) لفحص الفروق في معاملات ثبات الإعادة بين كل عينتين في المقاييس الطويلة

المقياس العيادي الطويل	حجم العينة	معامل الارتباط	قيمة V	كاي مربع*	Sig	القرار
الاكتئاب	284	0.608	0.05113	3.841	0.82	غير دال
	425	0.597				
	284	0.608	0.0301	3.841	0.86	غير دال
	567	0.600				
	425	0.597	0.00532	3.841	0.94	غير دال
	567	0.600				
البارانويا	284	0.754	0.00092	3.841	0.98	غير دال
	425	0.753				
	284	0.754	0.24715	3.841	0.62	غير دال
	567	0.738				
	425	0.753	0.27759	3.841	0.60	غير دال
	567	0.738				
الفصام	284	0.609	0.0273	3.841	0.87	غير دال
	425	0.601				
	284	0.609	0.0121	3.841	0.91	غير دال
	567	0.614				
	425	0.601	0.10332	3.841	0.75	غير دال
	567	0.614				
الهوس	284	0.632	0.02992	3.841	0.86	غير دال
	425	0.624				
	284	0.632	0.04183	3.841	0.84	غير دال
	567	0.623				
	425	0.624	0.00065	3.841	0.98	غير دال
	567	0.623				

* عدد درجات الحرية لاختبار كاي مربع هو (1)

الجدول رقم 19: نتائج الاختبار الاحصائي (V) لفحص الفروق في معاملات ثبات الإعادة بين كل عينتين في المقاييس القصيرة

المقياس العيادي الطويل	حجم العينة	معامل الارتباط	قيمة V	كاي مربع*	Sig	القرار
الاكتئاب	284	0.572	0.06263	3.841	0.80	غير دال
	425	0.559				
	284	0.572	0.01494	3.841	0.90	غير دال
	567	0.566				
	425	0.559	0.02552	3.841	0.87	غير دال
	567	0.566				
البارانويا	284	0.675	0.00058	3.841	0.98	غير دال
	425	0.674				
	284	0.675	0.268504	3.841	0.60	غير دال
	567	0.654				
	425	0.674	0.311531	3.841	0.58	غير دال
	567	0.654				
الفصام	284	0.575	0.00616	3.841	0.94	غير دال
	425	0.579				
	284	0.575	0.03485	3.841	0.85	غير دال
	567	0.584				
	425	0.579	0.01389	3.841	0.91	غير دال
	567	0.584				
الهوس	284	0.504	0.007754	3.841	0.93	غير دال
	425	0.509				
	284	0.504	0.000341	3.841	0.99	غير دال
	567	0.505				
	425	0.509	0.007054	3.841	0.93	غير دال
	567	0.505				

* عدد درجات الحرية لاختبار كاي مربع هو (1)

يتبين من الجدولين السابقين: عدم وجود فروق ذات دلالة إحصائية في معاملات الثبات بين كل عينتين من عينات البحث في المقاييس العيادية الطويلة والقصيرة، ويمكن أن يعود ذلك وبالرجوع إلى الجدول رقم (14) إلى اقتراب قيم الإحصاءات الوصفية لهذه العينات والتي تتضمن قيم المتوسط والانحراف المعياري والتباين والخطأ المعياري وأيضاً قيم الالتواء والتفرطح، كما يمكن أن يعود ذلك لعدم وجود فروقات في التجانس بين هذه العينات ذات الأحجام المختلفة.

حيث قامت الباحثة بحساب التجانس بين هذه العينات عند كل مقياس من المقاييس العيادية، باستخدام اختبار Leven للتجانس، وقد بينت النتائج أن قيمة ال Sig لهذا الاختبار بلغت (0.736) بالنسبة لمقياس الاكتئاب، و (0.787) بالنسبة لمقياس البارانويا، و(0.400) بالنسبة لمقياس الفصام، في حين بلغت(0.420) بالنسبة لمقياس الهوس الخفيف، وهذه القيم أكبر من (0.05) وبالتالي لاتوجد فروق دالة إحصائياً للتجانس بين هذه العينات بالنسبة للمقاييس العيادية الأربعة، مما يدل على تجانس هذه العينات.

ولمزيد من التأكُّد تمّ استخدام الاختبار الإحصائي (Z) عند مستوى الدلالة (0.05) لفحص الفروق في معاملات الثبات بين كل عينتين، لكل مقياس من المقاييس العيادية الطويلة والقصيرة، والجدولان الآتيان يوضحان نتائج الاختبار الإحصائي (Z):

الجدول رقم 20: نتائج الاختبار الإحصائي (Z) لفحص الفروق في معاملات ثبات الإعادة بين كل عينتين من العينات الثلاث في المقاييس الطويلة

المقياس العيادي الطويل	حجم العينة	قيمة معامل الارتباط	Z الجدولية	Z المحسوبة	القرار
الاكتئاب	284	0.608	1.96	0.226	غير دال
	425	0.597			
	284	0.608		0.173	غير دال
	567	0.600			
	425	0.597		-0.072	غير دال
	567	0.600			
البارانويا	284	0.754		0.030	غير دال
	425	0.753			
	284	0.754		0.497	غير دال
	567	0.738			
	425	0.753		0.526	غير دال
	567	0.738			
الفصام	284	0.609		0.165	غير دال
	425	0.601			
	284	0.609		-0.11	غير دال
	567	0.614			
	425	0.601		-0.321	غير دال
	567	0.614			

الهوس	284	0.632	1.96	0.172	غير دال
	425	0.624			
	284	0.632		0.204	غير دال
	567	0.623			
	425	0.624		0.025	غير دال
	567	0.623			

الجدول رقم 21: قيم الاختبار الإحصائي (Z) لفحص الفروق في معاملات ارتباط ثبات الإعادة بين كل عينتين من العينات الثلاث في المقاييس القصيرة

المقياس العيادي القصير	حجم العينة	قيمة معامل الارتباط بالإعادة	Z الجدولية	Z المحسوبة	القرار
الاكتئاب	284	0.572	1.96	0.25025	غير دال
	425	0.559			
	284	0.572		0.12223	غير دال
	567	0.566			
	425	0.559		-0.1597	غير دال
	567	0.566			
البارانويا	284	0.675	1.96	0.024	غير دال
	425	0.674			
	284	0.675		0.518	غير دال
	567	0.654			
	425	0.674		0.558	غير دال
	567	0.654			
الفصام	284	0.575	1.96	-0.0785	غير دال
	425	0.579			
	284	0.575		-0.1867	غير دال
	567	0.584			
	425	0.579		-0.1178	غير دال
	567	0.584			

الهوس	284	0.504	1.96	-0.088	غير دال
	425	0.509			
	284	0.504		-0.0185	غير دال
	567	0.505			
	425	0.509		0.0839	غير دال
	567	0.505			

يتبين من الجدولين السابقين: عدم وجود فروق ذات دلالة إحصائية في معاملات الثبات بين كل عينتين من عينات البحث، لكل مقياس من المقاييس العيادية الطويلة والقصيرة، وهذا ما يتفق مع نتائج الاختبار الإحصائي (V).

ثانياً: الإجابة عن السؤالين الثالث والرابع:

السؤال الثالث: هل توجد فروق ذات دلالة إحصائية بين معاملات ثبات الاتساق الداخلي باختلاف طول الاختبار؟

السؤال الرابع: هل توجد فروق ذات دلالة إحصائية بين معاملات ثبات الاتساق الداخلي باختلاف حجم العينة؟

للإجابة عن السؤالين الثالث والرابع تمّ حساب معاملات ثبات ألفا كرونباخ لكل مقياس من المقاييس العيادية الأربعة في الاختبارين الطويل والقصير، وعند كل حجم من أحجام العينات الآتية: (50%)، (100%)، (150%) و(200%) من عدد البنود الكلي للرائز، والجدول الآتي يوضح قيم هذه المعاملات.

الجدول رقم 22: قيم معاملات ثبات ألفا كرونباخ للمقاييس العيادية الطويلة والقصيرة وفق أحجام العينات المختلفة

المقياس العيادي / حجم العينة	الاكتئاب		البارانويا		الفصام		الهوس الخفيف	
	طويل	قصير	طويل	قصير	طويل	قصير	طويل	قصير
(284) %50	0.463	0.432	0.674	0.697	0.775	0.805	0.605	0.472
(567) %100	0.441	0.383	0.661	0.682	0.753	0.781	0.586	0.462
(851) %150	0.395	0.339	0.638	0.663	0.745	0.772	0.533	0.420
(1134) %200	0.372	0.312	0.631	0.657	0.740	0.767	0.526	0.421

يتبين من الجدول السابق هبوط قيم معاملات ثبات ألفا في المقاييس الطويلة والقصيرة مع ازدياد حجم العينة، ولاختبار دلالة الفروق في هذه المعاملات بين المقاييس الطويلة والقصيرة، عند كل

حجم من أحجام العينات الأربع، تم تطبيق الاختبار الإحصائي (M) الخاص بدراسة دلالة الفرق بين عدة معاملات ألفا كرونباخ لعينات مستقلة، وكانت النتائج على النحو الآتي:

أ-لمعرفة فيما إذا كانت هناك فروق ذات دلالة إحصائية بين معاملات ثبات الاتساق الداخلي باختلاف طول الاختبار:

قامت الباحثة بتطبيق الحالة الثالثة من الاختبار الإحصائي (M)، والتي تختبر الحالات التي يكون فيها عدد البنود مختلفاً وأحجام العينات متساوية، والجداول الآتية تبين نتائج هذا الاختبار:

الجدول رقم 23: نتائج الاختبار الإحصائي (M) لاختبار دلالة الفروق في معاملات ثبات ألفا بين المقاييس الطويلة والقصيرة عند حجم العينة (284)

المقياس العيادي	حجم العينة	عدد البنود	قيمة معامل ألفا	كاي مربع	درجات الحرية	قيمة M	Sig	القرار
الاكتئاب	284 (%50)	57	0.463	5.99	2	0.222	0.64	غير دال
	284 (%50)	49	0.432					
البارانويا	284 (%50)	40	0.674	5.99	2	0.415	0.52	غير دال
	284 (%50)	34	0,697					
الفصام	284 (%50)	78	0.775	5.99	2	1.407	0.24	غير دال
	284 (%50)	68	0,805					
الهوس الخفيف	284 (%50)	46	0.605	5.99	2	6.042	0.01	دال
	284 (%50)	34	0,472					

يتبين من الجدول السابق: عدم وجود فروق دالة إحصائياً عند مستوى الدلالة (0.05)، بين معاملات ثبات المقاييس الطويلة والقصيرة لكلٍ من مقياس: الاكتئاب والبارانويا والفصام، عند هذا الحجم من أحجام العينة، كما أظهرت نتائج الاختبار الاحصائي M وجود فرق دال إحصائياً بين معاملي ثبات ألفا الطويل والقصير، بالنسبة لمقياس الهوس الخفيف عند هذا الحجم من أحجام العينة؛ لصالح معامل ثبات المقياس الطويل.

الجدول رقم 24: نتائج الاختبار الاحصائي (M) لاختبار دلالة الفروق في معاملات ثبات ألفا بين المقاييس الطويلة والقصيرة عند حجم العينة (567)

المقياس العيادي	حجم العينة	عدد البنود	قيمة معامل ألفا	كاي مربع	درجات الحرية	قيمة M	Sig	القرار
الاكتئاب	567 (%100)	57	0.441	5.99	2	1.351	0.25	غير دال
	567 (%100)	49	0.383					
البارانويا	567 (%100)	40	0.661	5.99	2	0.643	0.42	غير دال
	567 (%100)	34	0,682					
الفصام	567 (%100)	78	0.753	5.99	2	1.931	0.16	غير دال
	567 (%100)	68	0,781					
الهوس الخفيف	567 (%100)	46	0.586	5.99	2	9.451	0.00	دال
	567 (%100)	34	0,462					

يتبين من الجدول السابق: عدم وجود فروق دالة إحصائياً عند مستوى الدلالة (0.05)، بين معاملات ثبات المقاييس الطويلة والقصيرة لكلٍ من مقياس: الاكتئاب والبارانويا والفصام، عند هذا الحجم من أحجام العينة، كما أظهرت نتائج الاختبار الإحصائي M وجود فرق دال إحصائياً بين معاملي ثبات ألفا الطويل والقصير، بالنسبة لمقياس الهوس الخفيف عند هذا الحجم من أحجام العينة؛ لصالح معامل ثبات المقياس الطويل.

الجدول رقم 25: نتائج الاختبار الاحصائي (M) لاختبار دلالة الفروق في معاملات ثبات ألفا بين المقاييس الطويلة والقصيرة عند حجم العينة (851)

المقياس العيادي	حجم العينة	عدد البنود	قيمة معامل ألفا	كاي مربع	درجات الحرية	قيمة M	Sig	القرار
الاكتئاب	851 (%150)	57	0.395	5.99	2	1.632	0.20	غير دال
	851 (%150)	49	0.339					
البارانويا	851 (%150)	40	0.638	5.99	2	1.179	0.28	غير دال
	851 (%150)	34	0,663					
الفصام	851 (%150)	78	0.745	5.99	2	2.493	0.11	غير دال
	851 (%150)	68	0,772					
الهوس الخفيف	851 (%150)	46	0.533	5.99	2	9.745	0.00	دال
	851 (%150)	34	0,420					

يتبين من الجدول السابق: عدم وجود فروق دالة إحصائياً عند مستوى الدلالة (0.05)، بين معاملات ثبات المقاييس الطويلة والقصيرة لكلٍ من مقياس: الاكتئاب والبارانويا والفصام، عند هذا الحجم من أحجام العينة، كما أظهرت نتائج الاختبار الإحصائي M وجود فرق دال إحصائياً بين معاملي ثبات ألفا الطويل والقصير، بالنسبة لمقياس الهوس الخفيف عند هذا الحجم من أحجام العينة؛ لصالح معامل ثبات المقياس الطويل.

الجدول رقم 26: نتائج الاختبار الإحصائي (M) لاختبار دلالة الفروق في معاملات ثبات ألفا بين المقاييس الطويلة والقصيرة عند حجم العينة (1134)

المقياس العيادي	حجم العينة	عدد البنود	قيمة معامل ألفا	كاي مربع	درجات الحرية	قيمة M	Sig	القرار
الاكتئاب	1134 (%200)	57	0.372	5.99	2	2.311	0.13	غير دال
	1134 (%200)	49	0.312					
البارانويا	1134 (%200)	40	0.631	5.99	2	1.632	0.20	غير دال
	1134 (%200)	34	0,657					
الفصام	1134 (%200)	78	0.740	5.99	2	3.183	0.07	غير دال
	1134 (%200)	68	0,767					
الهوس الخفيف	1134 (%200)	46	0.526	5.99	2	11.12	0.00	دال
	1134 (%200)	34	0,421					

يتبين من الجدول السابق: عدم وجود فروق دالة إحصائياً عند مستوى الدلالة (0.05)، بين معاملات ثبات المقاييس الطويلة والقصيرة لكلٍ من مقياس: الاكتئاب والبارانويا والفصام عند هذا الحجم من أحجام العينة، كما أظهرت نتائج الاختبار الإحصائي M وجود فرق دال إحصائياً بين معاملي ثبات ألفا الطويل والقصير بالنسبة لمقياس الهوس الخفيف عند هذا الحجم من أحجام العينة؛ لصالح معامل ثبات المقياس الطويل.

ويمكن أن تُعزى ظهور الفروق الدالة إحصائياً بين معاملي ثبات مقياس الهوس الطويل والقصير عند أحجام العينات الأربع، إلى زيادة الفرق في عدد البنود بين الاختبارين الطويل والقصير في هذا المقياس الموضح في الجدول ذي الرقم /13/ حيث تصل نسبة الفرق في عدد البنود إلى (26.08%) أي ما يقارب ربع بنود المقياس.

ب- لمعرفة فيما إذا كانت هناك فروق ذات دلالة إحصائية بين معاملات ثبات الاتساق الداخلي باختلاف حجم العينة:

قامت الباحثة بتطبيق الحالة الثانية من الاختبار الإحصائي (M)، والتي تدرس الحالات التي يكون فيها عدد البنود متساوٍ وأحجام العينات مختلفة، لاختبار الفروق في معاملات ثبات ألفا كرونباخ بين كل حجمين من أحجام العينات الأربع، لكلٍ من المقاييس الطويلة والقصيرة على حدة، والجدولان الآتيان يوضحان نتائج هذا الاختبار:

الجدول رقم 27: نتائج الاختبار الإحصائي (M) لاختبار دلالة الفروق بين معاملات ثبات ألفا للمقاييس الطويلة عند أحجام العينات المختلفة

المقياس العيادي	حجم العينة	عدد البنود	قيمة معامل ألفا	كاي مربع*	قيمة M	مستوى الدلالة	القرار
الاكتئاب	284 (%50)	57	0.463	7.81	4.264	0.12	غير دال
	567 (%100)	57	0.441				
	851 (%150)	57	0.395				
	1134 (%200)	57	0.372				
البارا نويا	284 (%50)	40	0.674	7.81	2.583	0.27	غير دال
	567 (%100)	40	0,661				
	851 (%150)	40	0,638				
	1134 (%200)	40	0,631				
الفصام	284 (%50)	78	0.775	7.81	2.565	0.28	غير دال
	567 (%100)	78	0,753				
	851 (%150)	78	0,745				
	1134 (%200)	78	0,740				
الهوس الخفيف	284 (%50)	46	0.605	7.81	6.415	0.04	دال
	567 (%100)	46	0,586				
	851 (%150)	46	0,533				
	1134 (%200)	46	0,526				

* عدد درجات الحرية لاختبار كاي مربع هو (3)

يتبين من الجدول السابق عدم وجود فروق دالة إحصائياً عند مستوى الدلالة (0.05)، بين معاملات ثبات المقاييس الطويلة لأحجام العينات الأربع المدروسة، بالنسبة للمقاييس العيادية: الاكتئاب والبارانويا والفصام، بينما أظهرت نتائج الاختبار الاحصائي M وجود فروق دالة إحصائياً عند مستوى الدلالة (0.05) بين معاملات ثبات ألفا بالنسبة لمقياس الهوس الطويل لأحجام العينات الأربع، وهذه الفروق لصالح معامل ارتباط العينة الأصغر (284) والتي تبلغ نسبتها (50%) من عدد البنود الكلي للرائز.

الجدول رقم 28: نتائج الاختبار الإحصائي (M) لاختبار دلالة الفروق بين معاملات ثبات ألفا للمقاييس القصيرة عند أحجام العينات المختلفة

المقياس العيادي	حجم العينة	عدد البنود	قيمة ألفا	كاي مربع	قيمة M	مستوى الدلالة	القرار
الاكتئاب	284 (50%)	49	0.432	7.81	5.215	0.07	غير دال
	567 (100%)	49	0.383				
	851 (150%)	49	0.339				
	1134 (200%)	49	0.312				
البارا نويا	284 (50%)	34	0.697	7.81	2.361	0.31	غير دال
	567 (100%)	34	0,682				
	851 (150%)	34	0,663				
	1134 (200%)	34	0,657				
الفصام	284 (50%)	68	0.805	7.81	3.918	0.14	غير دال
	567 (100%)	68	0,781				
	851 (150%)	68	0,772				
	1134 (200%)	68	0,767				
الهوس الخفيف	284 (50%)	34	0.472	7.81	1.917	0.38	غير دال
	567 (100%)	34	0,462				
	851 (150%)	34	0,420				
	1134 (200%)	34	0,421				

* عدد درجات الحرية لاختبار كاي مربع هو (3).

يتبين من الجدول السابق:

عدم وجود فروق دالة إحصائياً عند مستوى الدلالة (0.05) بين معاملات ثبات ألفا كرونباخ لأحجام العينات الأربع المدروسة، بالنسبة للمقاييس الأربعة: الاكتئاب، البارانويا، الفصام والهوس الخفيف.

ثالثاً: الإجابة عن السؤالين الخامس والسادس:

السؤال الخامس: هل توجد فروق ذات دلالة إحصائية بين معاملات الصدق المحكي باختلاف طول الاختبار بدلالة محك قائمة الأعراض؟

السؤال السادس: هل توجد فروق ذات دلالة إحصائية بين معاملات الصدق المحكي باختلاف حجم العينة بدلالة محك قائمة الأعراض؟

للإجابة عن السؤالين الخامس والسادس تمّ حساب معاملات ارتباط المقاييس العيادية: الاكتئاب والبارانويا والفصام الطويلة والقصيرة مع مثيلاتها في قائمة شطب الأعراض، لأحجام عينات (100%) و(75%) و(50%) من عدد البنود، في حين تمّ استثناء مقياس الهوس الخفيف وذلك لعدم وجود مقياس في قائمة الأعراض يقيس النواحي الإكلينيكية التي يقيسها هذا المقياس، والجدول الآتي يوضح قيم هذه المعاملات:

الجدول رقم 29: قيم معاملات الصدق المحكي للمقاييس العيادية الطويلة والقصيرة مع مثيلاتها في قائمة شطب الأعراض وفقاً لأحجام العينات المختلفة

المقياس العيادي / حجم العينة	الاكتئاب		البارانويا		الفصام	
	طويل	قصير	طويل	قصير	طويل	قصير
(284) %50	0.504**	0.438**	0.439**	0.380**	0.582**	0.490**
(425) %75	0.498**	0.460**	0.483**	0.384**	0.593**	0.498 **
(567) %100	0.528**	0.488**	0.495**	0.396**	0.610**	0.522**

يتبين من الجدول السابق: هبوط قيم معاملات الصدق المحكي للمقاييس القصيرة عن مثيلاتها للمقاييس الطويلة، ولاختبار دلالة الفروق بين هذه المعاملات للمقاييس الطويلة والقصيرة عند كل حجم من أحجام العينات الثلاث. تمّ تطبيق الاختبار الاحصائي (V) بعد تحويل معاملات الارتباط إلى قيم Z بطريقة فيشر، والجدول الآتي يبين نتائج الاختبار:

الجدول رقم 30: نتائج الاختبار الاحصائي (V) لفحص الفروق بين معاملات الصدق المحكي للمقاييس الطويلة والقصيرة عند أحجام العينات المختلفة

حجم العينة	المقياس	معامل الارتباط		قيمة V	كاي مربع*	قيمة المعنوية	القرار
567	الاكتئاب	0.528	0.488	1.0307	3.841	0.31	غير دال
	البارانويا	0.495	0.396	4.3196	3.841	0.04	دال
	الفصام	0.610	0.522	4.7538	3.841	0.03	دال
425	الاكتئاب	0.498	0.460	0.8204	3.841	0.37	غير دال
	البارانويا	0.483	0.384	3.1928	3.841	0.07	غير دال
	الفصام	0.593	0.498	4.337	3.841	0.04	دال
284	الاكتئاب	0.504	0.438	0.5207	3.841	0.47	غير دال
	البارانويا	0.439	0.380	0.7194	3.841	0.40	غير دال
	الفصام	0.582	0.490	2.3952	3.841	0.12	غير دال

* عدد درجات الحرية لاختبار كاي مربع هو (1)

يتبين من الجدول السابق: عدم وجود فروق دالة إحصائياً بين معاملي الصدق المحكي لمقياس الاكتئاب الطويل والقصير عند أحجام العينات الثلاث، في حين يتبين وجود فروق دالة إحصائياً بين معاملي الصدق المحكي لمقياس البارانويا الطويل والقصير عند حجم العينة (567) لصالح معامل صدق المقياس الطويل، ولم تصل هذه الفروق لمستوى الدلالة عند حجمي العينة (425) و(284)، ويُلاحظ وجود فروق دالة إحصائياً بين معاملي الصدق المحكي لمقياس الفصام الطويل والقصير عند حجمي العينة (425) و(567)، لصالح معامل صدق المقياس الطويل، بينما لم تصل هذه الفروق لمستوى الدلالة عند حجم العينة (284).

كما يُلاحظ أنه مع انخفاض حجم العينة تقل الفروق بين معاملات الصدق المحكي، وتصبح غير دالة إحصائياً؛ مما يعني أن حجم العينة وعدد البنود يؤثران معاً في ارتفاع معامل الصدق وانخفاضه، وبالتالي كلما قل حجم العينة لابد أن يكون عدد البنود أطول.

ولمزيد من التأكد تمَّ استخدام الاختبار الاحصائي (Z) عند مستوى الدلالة (0.05) لفحص الفروق بين معاملات الصدق المحكي للمقاييس الطويلة والقصيرة، عند كل حجم من أحجام العينات الثلاث، والجدول الآتي يبين نتائج هذا الاختبار:

الجدول رقم 31: قيم الاختبار (Z) لفحص الفروق في معاملات الارتباط المحكي للمقاييس الطويلة والقصيرة عند أحجام العينات المختلفة

حجم العينة	المقياس	معامل الارتباط		Zالجدولية	Z المحسوبة	القرار
567	الاكتئاب	0.528	0.488	1.96	1.01526	غير دال
	البارانويا	0.495	0.396	1.96	2.07838	دال
	الفصام	0.610	0.522	1.96	2.18032	دال
425	الاكتئاب	0.498	0.460	1.96	0.90574	غير دال
	البارانويا	0.483	0.384	1.96	1.78685	غير دال
	الفصام	0.593	0.498	1.96	1.98421	دال
284	الاكتئاب	0.504	0.438	1.96	0.72166	غير دال
	البارانويا	0.439	0.380	1.96	0.84822	غير دال
	الفصام	0.582	0.490	1.96	1.54766	غير دال

يتبين من الجدول السابق: عدم وجود فروق دالة إحصائياً بين معاملي الصدق المحكي لمقياس الاكتئاب الطويل والقصير عند أحجام العينات الثلاث، في حين يتبين وجود فروق دالة إحصائياً بين معاملي الصدق المحكي لمقياس البارانويا الطويل والقصير عند حجم العينة (567) لأن قيمة Z الجدولية أعلى من قيمة Z المحسوبة وهذا الفرق لصالح معامل صدق المقياس الطويل، بينما لم تصل هذه الفروق لمستوى الدلالة عند حجمي العينة (425) و(284)، ويُلاحظ وجود فروق دالة إحصائياً بين معاملي الصدق المحكي لمقياس الفصام الطويل والقصير عند حجمي العينة (567) و(425)، لصالح معامل صدق المقياس الطويل، بينما لم تصل هذه الفروق لمستوى الدلالة عند حجم العينة (284).

كما يُلاحظ اتفاق قيم الاختبار الإحصائي (Z) مع قيم الاختبار الإحصائي (V).

وللإجابة عن السؤال السادس فيما إذا كانت هناك فروق ذات دلالة إحصائية بين معاملات الصدق المحكي باختلاف حجم العينة بدلالة محك قائمة الأعراض:

قامت الباحثة باستخدام الاختبار الإحصائي (V) لاختبار الفروق في معاملات الصدق المحكي بين كل حجمي عينة من أحجام العينات الثلاث المدروسة، ولكلٍ من المقاييس الطويلة والقصيرة على حدة، وذلك بعد تحويل معاملات الارتباط إلى قيم Z بطريقة فيشر. والجدولان رقم 31 و 32 يوضحان نتائج هذا الاختبار:

الجدول رقم 32: نتائج الاختبار الاحصائي (V) لفحص الفروق في معاملات الصدق المحكي بين كل عينتين في المقاييس الطويلة

المقياس العيادي الطويل	حجم العينة	معامل الارتباط	قيمة V	كاي مربع*	Sig	القرار
الاكتئاب	284	0.504	0.011002	3.841	0.92	غير دال
	425	0.498				
	284	0.504	0.203095	3.841	0.65	غير دال
	567	0.528				
	425	0.498	0.40358	3.841	0.53	غير دال
	567	0.528				
البارانويا	284	0.439	0.53568	3.841	0.46	غير دال
	425	0.483				
	284	0.439	0.97477	3.841	0.32	غير دال
	567	0.495				
	425	0.483	0.06053	3.841	0.81	غير دال
	567	0.495				
الفصام	284	0.582	0.04838	3.841	0.83	غير دال
	425	0.593				
	284	0.582	0.35809	3.841	0.55	غير دال
	567	0.610				
	425	0.593	0.17271	3.841	0.68	غير دال
	567	0.610				

* عدد درجات الحرية لاختبار كاي مربع هو (1)

الجدول رقم 33: نتائج الاختبار الإحصائي (V) لفحص الفروق في معاملات الصدق المحكي بين كل عينتين في المقاييس القصيرة

المقياس العيادي القصير	حجم العينة	معامل الارتباط	قيمة V	كاي مربع*	Sig	القرار
الاكتئاب	284	0.438	0.1302	3.841	0.72	غير دال
	425	0.460				
	284	0.438	0.7695	3.841	0.38	غير دال
	567	0.488				
	425	0.460	0.3174	3.841	0.57	غير دال
	567	0.488				
البارانويا	284	0.380	0.00376	3.841	0.95	غير دال
	425	0.384				
	284	0.380	0.06733	3.841	0.80	غير دال
	567	0.396				
	425	0.384	0.04874	3.841	0.83	غير دال
	567	0.396				
الفصام	284	0.490	0.0192	3.841	0.89	غير دال
	425	0.498				
	284	0.490	0.35128	3.841	0.55	غير دال
	567	0.522				
	425	0.498	0.25611	3.841	0.61	غير دال
	567	0.522				

* عدد درجات الحرية لاختبار كاي مربع هو (1)

يتبين من الجدولين السابقين: عدم وجود فروق ذات دلالة إحصائية في معاملات الصدق المحكي بين كل عينتين من العينات الثلاث في المقاييس العيادية الطويلة والقصيرة، ويمكن أن يعود ذلك وبالرجوع إلى الجدول رقم (14) إلى اقتراب قيم الإحصاءات الوصفية لهذه العينات والتي تتضمن قيم المتوسط والانحراف المعياري والتباين والخطأ المعياري وأيضاً قيم الالتواء والتفرطح، كما يمكن أن يعود ذلك لعدم وجود فروقات في التجانس بين هذه العينات ذات الأحجام المختلفة.

حيث قامت الباحثة بحساب التجانس بين هذه العينات عند كل مقياس من المقاييس العيادية، باستخدام اختبار Leven للتجانس، وقد بينت النتائج أن قيمة ال Sig لهذا الاختبار بلغت

(0.736) بالنسبة لمقياس الاكتئاب، و (0.787) بالنسبة لمقياس البارانويا، و(0.400) بالنسبة لمقياس الفصام، وهذه القيم أكبر من (0.05) وبالتالي لاتوجد فروق دالة إحصائياً للتجانس بين هذه العينات بالنسبة للمقاييس العيادية الأربعة، مما يدل على تجانس هذه العينات.

ولمزيد من التأكُّد تمّ استخدام الاختبار الإحصائي (Z) عند مستوى الدلالة (0.05)، لفحص الفروق في معاملات الصدق المحكي بين كل عينتين، لكلٍ من المقاييس العيادية الطويلة والقصيرة المدروس ارتباطها مع مثيلاتها في قائمة الشطب، والجدولان الآتيان يوضحان نتائج هذا الاختبار:

الجدول رقم 34: نتائج الاختبار الإحصائي (Z) لفحص الفروق في معاملات الصدق المحكي بين كل عينتين من العينات الثلاث في المقاييس الطويلة

المقياس العيادي الطويل	حجم العينة	قيمة معامل الارتباط	Z الجدولية	Z المحسوبة	القرار عند مستوى الدلالة (0.05)
الاكتئاب	284	0.504	1.96	0.1048	غير دال
	425	0.498			
	284	0.504		-0.4507	غير دال
	567	0.528			
	425	0.498		-0.6353	غير دال
	567	0.528			
البارانويا	284	0.439	1.96	-0.7319	غير دال
	425	0.483			
	284	0.439		-0.9873	غير دال
	567	0.495			
	425	0.483		-0.246	غير دال
	567	0.495			
الفصام	284	0.582	1.96	-0.22	غير دال
	425	0.593			
	284	0.582		-0.5984	غير دال
	567	0.610			
	425	0.593		-0.4156	غير دال
	567	0.610			

الجدول رقم 35: نتائج الاختبار الإحصائي (Z) لفحص الفروق في معاملات الصدق المحكي بين كل عينتين من العينات الثلاث في المقاييس القصيرة

المقياس العيادي القصير	حجم العينة	قيمة معامل الارتباط	Z الجدولية	Z المحسوبة	القرار عند مستوى الدلالة (0.05)
الاكتئاب	284	0.438	1.96	-0.3608	غير دال
	425	0.460			
	284	0.438		-0.8338	غير دال
	567	0.488			
	425	0.460		-0.5635	غير دال
	567	0.488			
البارانويا	284	0.38	1.96	-0.0613	غير دال
	425	0.384			
	284	0.38		-0.2595	غير دال
	567	0.396			
	425	0.384		-0.2208	غير دال
	567	0.396			
الفصام	284	0.490	1.96	-0.1386	غير دال
	425	0.498			
	284	0.490		-0.5927	غير دال
	567	0.522			
	425	0.498		-0.5061	غير دال
	567	0.522			

يتبين من الجدولين (33) و(34): عدم وجود فروق ذات دلالة إحصائية في معاملات الصدق المحكي بين كل عينتين من العينات الثلاث المدروسة، لكل مقياس من المقاييس العيادية الطويلة والقصيرة، وهذا ما يتفق مع نتائج الاختبار الإحصائي (V).

رابعاً: الإجابة عن السؤال السابع: هل توجد فروق ذات دلالة إحصائية بين معاملات صدق الاختبار باستخدام طريقة المجموعات المتضادة /بين عينتي الأسوياء، والمضطربين نفسياً/ باختلاف طول الاختبار؟ ويتفرع عن هذا السؤال الأسئلة الفرعية الآتية:

- هل توجد فروق ذات دلالة إحصائية بين معاملات صدق الاختبار باستخدام طريقة المجموعات المتضادة، باختلاف طول مقياس الاكتئاب العيادي؟

- هل توجد فروق ذات دلالة إحصائية بين معاملات صدق الاختبار باستخدام طريقة المجموعات المتضادة، باختلاف طول مقياس البارانويا العيادي؟
- هل توجد فروق ذات دلالة إحصائية بين معاملات صدق الاختبار باستخدام طريقة المجموعات المتضادة، باختلاف طول مقياس الفصام العيادي؟
- هل توجد فروق ذات دلالة إحصائية بين معاملات صدق الاختبار باستخدام طريقة المجموعات المتضادة، باختلاف طول مقياس الهوس العيادي؟

للإجابة عن السؤال الفرعي الأول: قامت الباحثة باختبار الفروق بين متوسطات درجات مقياس الاكتئاب، لكلٍ من عينتي الأسوياء والمضطربين نفسياً، لنوعي بنود الاختبار الطويل والقصير، من خلال إجراء تحليل التباين الأحادي (ANOVA)، والجدول الآتي يبين نتائج هذا التحليل:

الجدول رقم 36: نتائج تحليل التباين الأحادي لمتوسطات درجات مقياس الاكتئاب لكلٍ من عينتي الأسوياء والمضطربين نفسياً، لنوعي بنود الاختبار الطويل والقصير

العينة	حجم العينة	المتوسط الحسابي	الانحراف المعياري	مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	اختبار F	Sig
مضطربون نفسياً طويل	150	27.86	4.78	بين المجموعات	3976.77	3	1325.59	69.853	0.000
أسوياء طويل	150	31.73	4.66	ضمن المجموعات	11310.18	596	18.97		
أسوياء قصير	150	28.21	4.09	المجموع	15286.95	599			
مضطربون نفسياً قصير	150	24.44	3.81						

يتبين من الجدول السابق: وجود فروق ذات دلالة إحصائية عند مستوى الدلالة (0.05)، في متوسطات درجات المقياس العيادي الاكتئاب، بين كلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير، وقد قامت الباحثة بحساب حجم التأثير للفروق بين المتوسطات باستخدام إحصاء إيتا تربيع كما يلي:

قيمة إحصاء إيتا تربيع= مجموع المربعات البينية / المجموع الكلي للتربيع (بالانت، 2006، ص 246)

وهكذا كانت قيمة إيتا تربيع = (0.26)، يُلاحَظ أن هذه القيمة – وبالرجوع إلى إرشادات بالانت لتفسير قوة قيم إيتا تربيع– تدل على حجم تأثير كبير.

وللكشف عن هذه الفروق تمّ إجراء عمليات المقارنات البعدية، وفق اختبار Dunnett's (C) للمجموعات غير المتجانسة، لأن قيمة Sig لاختبار Leven بلغت (0.018)، أي توجد فروق دالة إحصائياً للتجانس بين العينات، وبالتالي العينات غير متجانسة، والجدول الآتي يوضح نتائج هذه المقارنات:

الجدول رقم 37: نتائج المقارنات البعدية بين متوسطات درجات مقياس الاكتئاب لكلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير

المقياس	العينة (I)	العينة (j)	الفروق بين المتوسطات	الخطأ المعياري	مستوى الثقة 95%	
					الحد الأدنى	الحد الأعلى
الاكتئاب	مضطربون نفسياً طويل	أسوياء طويل	-3.87154*	0.54624	-5.2909	-2.4522
		أسوياء قصير	-0.35854	0.51338	-1.6924	0.9753
		مضطربون نفسياً قصير	3.41333*	0.49936	2.1159	4.7108
	أسوياء طويل	أسوياء قصير	3.51300*	0.50710	2.1954	4.8306
		مضطربون نفسياً قصير	7.28488*	0.49290	6.0041	8.5656
	أسوياء قصير	مضطربون نفسياً قصير	3.77188*	0.45620	2.5866	4.9572

The mean difference is significant at the (0.05) level

يتبين من الجدول السابق: وجود فروق دالة إحصائياً بين متوسط درجات عينة المضطربين نفسياً على الاختبارين الطويل والقصير، لصالح متوسط عينة المضطربين نفسياً على الاختبار الطويل، وكذلك وجود فروق دالة إحصائياً بين متوسط درجات عينة الأسوياء على الاختبارين الطويل والقصير، لصالح متوسط عينة الأسوياء على الاختبار الطويل، ما يعني أن هناك أثراً دالاً لطول الاختبار على متوسط درجات العينتين، وبالتالي على معاملات صدق الاختبار.

للإجابة عن السؤال الفرعي الثاني: قامت الباحثة باختبار الفروق بين متوسطات درجات مقياس البارانويا، لكلٍ من عينتي الأسوياء والمضطربين نفسياً، لنوعي بنود الاختبار الطويل والقصير، من خلال إجراء تحليل التباين الأحادي (ANOVA)، والجدول الآتي يبين نتائج هذا التحليل:

الجدول رقم 38: نتائج تحليل التباين الأحادي لمتوسطات درجات مقياس البارانويا لكلٍ من عينتي الأسوياء والمضطربين نفسياً، لنوعي بنود الاختبار الطويل والقصير

العينة	حجم العينة	المتوسط الحسابي	الانحراف المعياري	مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	اختبار F	Sig
مضطربون نفسياً طويل	150	23.89	4.97	بين المجموعات	6186.82	3	2062.27	97.713	0.000
أسوياء طويل	150	18.51	4.42	ضمن المجموعات	12578.80	596	21.105		
أسوياء قصير	150	15.25	4.16	المجموع	18765.62	599			
مضطربون نفسياً قصير	150	21.25	4.76						

يتبين من الجدول السابق: وجود فروق ذات دلالة إحصائية عند مستوى الدلالة (0.05)، في متوسطات درجات المقياس العيادي البارانويا، بين كلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير، وقد قامت الباحثة بحساب حجم التأثير للفروق بين المتوسطات باستخدام إحصاء إيتا تربيع كما يلي:

قيمة إحصاء إيتا تربيع= مجموع المربعات البينية / المجموع الكلي للتربيع (المرجع السابق، ص 246).

وهكذا كانت قيمة إيتا تربيع = (0.33)، يُلاحَظ أن هذه القيمة – وبالرجوع إلى إرشادات بالانت لتفسير قوة قيم إيتا تربيع– تدل على حجم تأثير كبير.

وللكشف عن هذه الفروق تمّ إجراء عمليات المقارنات البعدية وفق اختبار LSD للمجموعات المتجانسة، لأن قيمة Sig لاختبار Leven بلغت (0.622)، أي لا توجد فروق دالة إحصائياً في التجانس بين المجموعات، وبالتالي العينات متجانسة، والجدول الآتي يوضح نتائج هذه المقارنات:

الجدول رقم 39: نتائج المقارنات البعدية بين متوسطات درجات مقياس البارنويا لكلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير

المقياس	العينة (I)	العينة (j)	الفروق بين المتوسطات	الخطأ المعياري	مستوى الثقة 95%	
					الحد الأدنى	الحد الأعلى
البارانويا	مضطربون نفسياً طويل	أسوياء طويل	5.37655*	0.53137	4.3330	6.4201
		أسوياء قصير	8.63506*	0.52960	7.5950	9.6752
		مضطربون نفسياً قصير	2.64000*	0.53048	1.5982	3.6818
	أسوياء طويل	أسوياء قصير	3.25850*	0.53049	2.2166	4.3004
		مضطربون نفسياً قصير	-2.73655*	0.53137	-3.7801	-1.6930
	أسوياء قصير	مضطربون نفسياً قصير	-5.99506*	0.52960	-7.0352	-4.9550

The mean difference is significant at the (0.05) level

يتبين من الجدول السابق: وجود فروق دالة إحصائياً بين متوسط درجات عينة المضطربين نفسياً على الاختبارين الطويل والقصير، لصالح متوسط عينة المضطربين نفسياً على الاختبار الطويل، وكذلك وجود فروق دالة إحصائياً بين متوسط درجات عينة الأسوياء على الاختبارين الطويل والقصير، لصالح متوسط عينة الأسوياء على الاختبار الطويل، ما يعني أن هناك أثراً دالاً لطول الاختبار على متوسط درجات العينتين، وبالتالي على معاملات صدق الاختبار.

للإجابة عن السؤال الفرعي الثالث: قامت الباحثة باختبار الفروق بين متوسطات درجات مقياس الفصام، لكلٍ من عينتي الأسوياء والمضطربين نفسياً، لنوعي بنود الاختبار الطويل والقصير، من خلال إجراء تحليل التباين الأحادي (ANOVA)، والجدول الآتي يبين نتائج هذا التحليل:

الجدول رقم 40: نتائج تحليل التباين الأحادي لمتوسطات درجات مقياس الفصام لكلٍ من عينتي الأسوياء والمضطربين نفسياً، لنوعي بنود الاختبار الطويل والقصير

العينة	حجم العينة	المتوسط الحسابي	الانحراف المعياري	مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	اختبار F	Sig
مضطربون نفسياً طويل	150	44.66	11.21	بين المجموعات	28638.40	3	9546.13	114.611	0.000
أسوياء طويل	150	32.28	7.47	ضمن المجموعات	49641.66	596	83.29		
أسوياء قصير	150	26.33	7.11	المجموع	78280.07	599			
مضطربون نفسياً قصير	150	38.98	10.04						

يتبين من الجدول السابق: وجود فروق ذات دلالة إحصائية عند مستوى الدلالة (0.05)، في متوسطات درجات المقياس العيادي الفصام، بين كلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير، وقد قامت الباحثة بحساب حجم التأثير للفروق بين المتوسطات باستخدام إحصاء إيتا تربيع كما يلي:

قيمة إحصاء إيتا تربيع= مجموع المربعات البينية / المجموع الكلي للتربيع (المرجع السابق، ص 246).

وهكذا كانت قيمة إيتا تربيع = (0.36)، يُلاحَظ أن وهذه القيمة – وبالرجوع إلى إرشادات بالانت لتفسير قوة قيم إيتا تربيع– تدل على حجم تأثير كبير.

وللكشف عن هذه الفروق تمّ إجراء عمليات المقارنات البعدية وفق اختبار Dunnett's (C) للمجموعات غير المتجانسة، لأن قيمة Sig لاختبار Leven بلغت (0.000) أي توجد فروق دالة إحصائياً للتجانس بين العينات، وبالتالي العينات غير متجانسة، والجدول الآتي يوضح نتائج هذه المقارنات:

الجدول رقم 41: نتائج المقارنات البعدية بين متوسطات درجات مقياس الفصام لكلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير

المقياس	العينة (I)	العينة (J)	الفروق بين المتوسطات	الخطأ المعياري	مستوى الثقة 95%	
					الحد الأدنى	الحد الأعلى
الفصام	مضطربون نفسياً طويل	أسوياء طويل	12.37141*	1.10184	9.5085	15.2344
		أسوياء قصير	18.32887*	1.08338	15.5140	21.1437
		مضطربون نفسياً قصير	5.67333*	1.22921	2.4795	8.8671
	أسوياء طويل	أسوياء قصير	5.95746*	0.84301	3.7671	8.1478
		مضطربون نفسياً قصير	-6.69808*	1.02367	-9.3579	-4.0382
	أسوياء قصير	مضطربون نفسياً قصير	-12.6555*	1.00377	-15.2635	-10.0475

The mean difference is significant at the (0.05) level

يتبين من الجدول السابق: وجود فروق دالة إحصائياً بين متوسط درجات عينة المضطربين نفسياً على الاختبارين الطويل والقصير، لصالح متوسط عينة المضطربين نفسياً على الاختبار الطويل، وكذلك وجود فروق دالة إحصائياً بين متوسط درجات عينة الأسوياء على الاختبارين الطويل والقصير، لصالح متوسط عينة الأسوياء على الاختبار الطويل، ما يعني أن هناك أثراً دالاً لطول الاختبار على متوسط درجات العينتين، وبالتالي على معاملات صدق الاختبار.

وللإجابة عن السؤال الفرعي الرابع: قامت الباحثة باختبار الفروق بين متوسطات درجات مقياس الهوس، لكلٍ من عينتي الأسوياء والمضطربين نفسياً، لنوعي بنود الاختبار الطويل والقصير، من خلال إجراء تحليل التباين الأحادي (ANOVA)، والجدول الآتي يبين نتائج هذا التحليل:

الجدول رقم 42: نتائج تحليل التباين الأحادي لمتوسطات درجات مقياس الهوس الخفيف لكلٍ من عينتي الأسوياء والمضطربين نفسياً، لنوعي بنود الاختبار الطويل والقصير

العينة	حجم العينة	المتوسط الحسابي	الانحراف المعياري	مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	اختبار F	Sig
مضطربون نفسياً طويل	150	30.16	5.98	بين المجموعات	11284.04	3	3763.01	164.159	0.000
أسوياء طويل	150	22.77	4.82	ضمن المجموعات	13662.07	596	22.92		
أسوياء قصير	150	18.09	3.73	المجموع	24951.11	599			
مضطربون نفسياً قصير	150	22.50	4.33						

يتبين من الجدول السابق: وجود فروق ذات دلالة إحصائية عند مستوى الدلالة (0.05)، في متوسطات درجات المقياس العيادي الهوس الخفيف، بين كلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير، وقد قامت الباحثة بحساب حجم التأثير للفروق بين المتوسطات باستخدام إحصاء إيتا تربيع كما يلي:

قيمة إحصاء إيتا تربيع= مجموع المربعات البينية / المجموع الكلي للتربيع (المرجع السابق، ص 246).

وهكذا كانت قيمة إيتا تربيع = (0.45)، يُلاحَظ أن هذه القيمة – وبالرجوع إلى إرشادات بالانت لتفسير قوة قيم إيتا تربيع– تدل على حجم تأثير كبير.

وللكشف عن هذه الفروق تمّ إجراء عمليات المقارنات البعدية وفق اختبار Dunnett's (C) للمجموعات غير المتجانسة، لأن قيمة Sig لاختبار Leven بلغت (0.000)، أي توجد فروق دالة إحصائياً للتجانس بين العينات، وبالتالي العينات غير متجانسة، والجدول الآتي يوضح نتائج هذه المقارنات:

الجدول رقم 43: نتائج المقارنات البعدية بين متوسطات درجات مقياس الهوس الخفيف لكلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير

المقياس	العينة (I)	العينة (j)	الفروق بين المتوسطات	الخطأ المعياري	مستوى الثقة 95%	
					الحد الأدنى	الحد الأعلى
الهوس الخفيف	مضطربون نفسياً طويل	أسوياء طويل	7.39485*	0.62833	5.7623	9.0275
		أسوياء قصير	12.06733*	0.57513	10.5730	13.5616
		مضطربون نفسياً قصير	7.66000*	0.60314	6.0929	9.2271
	أسوياء طويل	أسوياء قصير	4.67247*	0.49835	3.3776	5.9673
		مضطربون نفسياً قصير	0.26515	0.53043	-1.1131	1.6434
	أسوياء قصير	مضطربون نفسياً قصير	-4.40733*	0.46619	-5.6186	-3.1961

The mean difference is significant at the (0.05) level

يتبين من الجدول السابق: وجود فروق دالة إحصائياً بين متوسط درجات عينة المضطربين نفسياً على الاختبارين الطويل والقصير، لصالح متوسط عينة المضطربين نفسياً على الاختبار الطويل، وكذلك وجود فروق دالة إحصائياً بين متوسط درجات عينة الأسوياء على الاختبارين الطويل والقصير، لصالح متوسط عينة الأسوياء على الاختبار الطويل، ما يعني أن هناك أثراً دالاً لطول الاختبار على متوسط درجات العينتين، وبالتالي على معاملات صدق الاختبار.

<u>خامساً: الإجابة عن السؤال الثامن:</u> هل توجد فروق ذات دلالة إحصائية بين متوسطات الدرجات على المقاييس العيادية استناداً لمتغير المرحلة الدراسية /مرحلة التعليم الجامعي، ومرحلة التعليم ما بعد الجامعي – الماجستير/ باختلاف طول الاختبار؟

ويتفرع عن هذا السؤال الأسئلة الفرعية الآتية:

- هل توجد فروق ذات دلالة إحصائية بين متوسطات درجات مقياس الاكتئاب العيادي استناداً لمتغير المرحلة الدراسية، باختلاف طول المقياس؟

- هل توجد فروق ذات دلالة إحصائية بين متوسطات درجات مقياس البارانويا العيادي استناداً لمتغير المرحلة الدراسية، باختلاف طول المقياس؟
- هل توجد فروق ذات دلالة إحصائية بين متوسطات درجات مقياس الفصام العيادي استناداً لمتغير المرحلة الدراسية، باختلاف طول المقياس؟
- هل توجد فروق ذات دلالة إحصائية بين متوسطات درجات مقياس الهوس العيادي استناداً لمتغير المرحلة الدراسية، باختلاف طول المقياس؟

للإجابة عن السؤال الفرعي الأول: قامت الباحثة باختبار الفروق بين متوسطات درجات مقياس الاكتئاب، لكلٍ من عينتي طلبة الماجستير والطلبة الجامعيين، لنوعي بنود الاختبار الطويل والقصير، من خلال إجراء تحليل التباين الأحادي (ANOVA)، والجدول الآتي يبين نتائج هذا التحليل:

الجدول رقم 44: نتائج تحليل التباين الأحادي لمتوسطات درجات مقياس الاكتئاب لكلٍ من عينتي طلبة الماجستير والطلبة الجامعيين لنوعي بنود الاختبار الطويل والقصير

العينة	حجم العينة	المتوسط الحسابي	الانحراف المعياري	مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	اختبار F	Sig
طلبة ماجستير طويل	206	31.868	3.72	بين المجموعات	3084.28	3	1028.09	66.824	0.000
طلبة جامعيون طويل	206	31.762	4.51	ضمن المجموعات	12615.88	820	15.38		
طلبة ماجستير قصير	206	27.694	3.36	المجموع	15700.17	823			
طلبة جامعيون قصير	206	28.237	3.99						

يتبين من الجدول السابق: وجود فروق ذات دلالة إحصائية عند مستوى الدلالة (0.05)، في متوسطات درجات المقياس العيادي الاكتئاب، بين كلٍ من عينتي طلبة الماجستير والطلبة الجامعيين لنوعي بنود الاختبار الطويل والقصير، وقد قامت الباحثة بحساب حجم التأثير للفروق بين المتوسطات باستخدام إحصاء إيتا تربيع كما يلي:

قيمـــة إحصـــاء إيتـــا تربيـــع= مجمـــوع المربعـــات البينيـــة / المجمـــوع الكلـــي للتربيـــع (بالانت، 2006، ص 246).

وهكذا كانت قيمة إيتا تربيع = (0.19)، ويُلاحَظ أن هذه القيمة – وبالرجوع إلى إرشادات بالانت لتفسير قوة قيم إيتا تربيع– تدل على حجم تأثير كبير.

وللكشف عن هذه الفروق تمّ إجراء عمليات المقارنات البعدية، وفق اختبار Dunnett's (C) للمجموعات غير المتجانسة، لأن قيمة Sig لاختبار Leven بلغت (0.000)، أي توجد فروق دالة إحصائياً للتجانس بين العينات، وبالتالي العينات غير متجانسة، والجدول الآتي يوضح نتائج هذه المقارنات:

الجدول رقم 45: نتائج المقارنات البعدية بين متوسطات درجات مقياس الاكتئاب لكلٍ من عينتي طلبة الماجستير والطلبة الجامعيين، لنوعي بنود الاختبار الطويل والقصير

المقياس	العينة (I)	العينة (j)	الفروق بين المتوسطات	الخطأ المعياري	مستوى الثقة 95%	
					الحد الأدنى	الحد الأعلى
الاكتئاب	طلبة ماجستير طويل	طلبة جامعيون طويل	0.10680	0.40768	-0.9492	1.1628
		طلبة ماجستير قصير	4.17476*	0.35000	3.2682	5.0813
		طلبة جامعيون قصير	3.63107*	0.38052	2.6454	4.6167
	طلبة جامعيون طويل	طلبة ماجستير قصير	4.06796*	0.39236	3.0517	5.0843
		طلبة جامعيون قصير	3.52427*	0.41981	2.4369	4.6117
	طلبة ماجستير قصير	طلبة جامعيون قصير	-0.54369	0.36406	-1.4867	0.3993

The mean difference is significant at the (0.05) level

يتبين من الجدول السابق: وجود فروق دالة إحصائياً بين متوسط درجات عينة طلبة الماجستير على الاختبارين الطويل والقصير، لصالح متوسط عينة طلبة الماجستير على الاختبار الطويل، كما يتبين وجود فروق دالة إحصائياً بين متوسط درجات عينة الطلبة الجامعيين على الاختبارين

الطويل والقصير، لصالح متوسط عينة الطلبة الجامعيين على الاختبار الطويل، ما يعني أن هناك أثراً دالاً لطول الاختبار في متوسطات درجات العينتين.

للإجابة عن السؤال الفرعي الثاني: قامت الباحثة باختبار الفروق بين متوسطات درجات مقياس البارانويا، لكلٍ من عينتي طلبة الماجستير والطلبة الجامعيين، لنوعي بنود الاختبار الطويل والقصير، من خلال إجراء تحليل التباين الأحادي(ANOVA)، والجدول الآتي يبين نتائج هذا التحليل:

الجدول رقم 46: نتائج تحليل التباين الأحادي لمتوسطات درجات مقياس البارانويا لكلٍ من عينتي طلبة الماجستير والطلبة الجامعيين لنوعي بنود الاختبار الطويل والقصير

العينة	حجم العينة	المتوسط الحسابي	الانحراف المعياري	مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	اختبار F	Sig
طلبة ماجستير طويل	206	17.568	3.36	بين المجموعات	2536.80	3	845.600	57.832	0.000
طلبة جامعيون طويل	206	18.679	4.43	ضمن المجموعات	11989.79	820	14.622		
طلبة ماجستير قصير	206	14.242	3.13	المجموع	14526.59	823			
طلبة جامعيون قصير	206	15.344	4.20						

يتبين من الجدول السابق: وجود فروق ذات دلالة إحصائية عند مستوى الدلالة (0.05)، في متوسطات درجات المقياس العيادي البارانويا، بين كلٍ من عينتي طلبة الماجستير والطلبة الجامعيين لنوعي بنود الاختبار الطويل والقصير، وقد قامت الباحثة بحساب حجم التأثير للفروق بين المتوسطات باستخدام إحصاء إيتا تربيع كما يأتي:

قيمة إحصاء إيتا تربيع= مجموع المربعات البينية / المجموع الكلي للتربيع (بالانت، 2006، ص 246)

وهكذا كانت قيمة إيتا تربيع = (0.17)، حيث يُلاحَظ أن هذه القيمة – وبالرجوع إلى إرشادات بالانت لتفسير قوة قيم إيتا تربيع– تدل على حجم تأثير كبير.

وللكشف عن هذه الفروق تمّ إجراء عمليات المقارنات البعدية وفق اختبار Dunnett's (C) للمجموعات غير المتجانسة، لأن قيمة Sig لاختبار Leven بلغت (0.000)، أي توجد فروق دالة إحصائياً للتجانس بين العينات، وبالتالي العينات غير متجانسة، والجدول الآتي يوضح نتائج هذه المقارنات:

الجدول رقم 47: نتائج المقارنات البعدية بين متوسطات درجات مقياس البارانويا لكلٍ من عينتي طلبة الماجستير والطلبة الجامعيين، لنوعي بنود الاختبار الطويل والقصير

المقياس	العينة (I)	العينة (j)	الفروق بين المتوسطات	الخطأ المعياري	مستوى الثقة 95%	
					الحد الأدنى	الحد الأعلى
البارانويا	طلبة ماجستير طويل	طلبة جامعيون طويل	-1.11165*	0.38794	-2.1165	-0.1068
		طلبة ماجستير قصير	3.32524*	0.32008	2.4962	4.1543
		طلبة جامعيون قصير	2.22330*	0.37508	1.2518	3.1948
	طلبة جامعيون طويل	طلبة ماجستير قصير	4.43689*	0.37846	3.4566	5.4172
		طلبة جامعيون قصير	3.33495*	0.42599	2.2315	4.4384
	طلبة ماجستير قصير	طلبة جامعيون قصير	-1.10194*	-0.36527	-2.0481	-0.1558

The mean difference is significant at the (0.05) level

يتبين من الجدول السابق: وجود فروق دالة إحصائياً بين متوسط درجات عينة طلبة الماجستير على الاختبارين الطويل والقصير، لصالح متوسط عينة طلبة الماجستير على الاختبار الطويل، كما يتبين وجود فروق دالة إحصائياً بين متوسط درجات عينة الطلبة الجامعيين على الاختبارين الطويل والقصير، لصالح متوسط عينة الطلبة الجامعيين على الاختبار الطويل، ما يعني أن هناك أثراً دالاً لطول الاختبار في متوسطات درجات العينتين.

للإجابة عن السؤال الفرعي الثالث: قامت الباحثة باختبار الفروق بين متوسطات درجات مقياس الفصام، لكلٍ من عينتي طلبة الماجستير والطلبة الجامعيين، لنوعي بنود الاختبار الطويل

والقصير، من خلال إجراء تحليل التباين الأحادي (ANOVA)، والجدول الآتي يبين نتائج هذا التحليل:

الجدول رقم 48: نتائج تحليل التباين الأحادي لمتوسطات درجات مقياس الفصام لكلٍ من عينتي طلبة الماجستير والطلبة الجامعيين لنوعي بنود الاختبار الطويل والقصير

العينة	حجم العينة	المتوسط الحسابي	الانحراف المعياري	مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	اختبار F	Sig
طلبة ماجستير طويل	206	30.830	5.04	بين المجموعات	8177.33	3	2725.77	70.572	0.000
طلبة جامعيون طويل	206	32.640	7.38	ضمن المجموعات	31671.58	820	38.62		
طلبة ماجستير قصير	206	24.771	4.91	المجموع	39848.91	823			
طلبة جامعيون قصير	206	26.650	7.09						

يتبين من الجدول السابق: وجود فروق ذات دلالة إحصائية عند مستوى الدلالة (0.05)، في متوسطات درجات المقياس العيادي الفصام، بين كلٍ من عينتي طلبة الماجستير والطلبة الجامعيين لنوعي بنود الاختبار الطويل والقصير، وقد قامت الباحثة بحساب حجم التأثير للفروق بين المتوسطات باستخدام إحصاء إيتا تربيع كما يلي:

قيمة إحصاء إيتا تربيع= مجموع المربعات البينية / المجموع الكلي للتربيع (بالانت، 2006، ص 246)

وهكذا كانت قيمة إيتا تربيع = (0.20)، ويُلاحَظ أن هذه القيمة – وبالرجوع إلى إرشادات بالانت لتفسير قوة قيم إيتا تربيع– تدل على حجم أثر كبير.

وللكشف عن هذه الفروق تمّ إجراء عمليات المقارنات البعدية، وفق اختبار Dunnett's (C) للمجموعات غير المتجانسة، لأن قيمة Sig لاختبار Leven بلغت (0.000)، أي توجد فروق دالة إحصائياً للتجانس بين العينات، وبالتالي العينات غير متجانسة، والجدول الآتي يوضح نتائج هذه المقارنات:

الجدول رقم 49: نتائج المقارنات البعدية بين متوسطات درجات مقياس الفصام لكلٍ من عينتي طلبة الماجستير والطلبة الجامعيين، لنوعي بنود الاختبار الطويل والقصير

المقياس	العينة (I)	العينة (J)	الفروق بين المتوسطات	الخطأ المعياري	مستوى الثقة 95%	
					الحد الأدنى	الحد الأعلى
الفصام	طلبة ماجستير طويل	طلبة جامعيون طويل	-1.81068*	0.62326	-3.4251	-0.1963
		طلبة ماجستير قصير	6.05825*	0.49038	4.7880	7.3285
		طلبة جامعيون قصير	4.17961*	0.60650	2.6085	5.7506
	طلبة جامعيون طويل	طلبة ماجستير قصير	7.86893*	0.61817	6.2677	9.4701
		طلبة جامعيون قصير	5.99029*	0.71379	4.1414	7.8392
	طلبة ماجستير قصير	طلبة جامعيون قصير	-1.87864*	0.60127	-3.4361	-0.3212

The mean difference is significant at the (0.05) level

يتبين من الجدول السابق: وجود فروق دالة إحصائياً بين متوسط درجات عينة طلبة الماجستير على الاختبارين الطويل والقصير، لصالح متوسط عينة طلبة الماجستير على الاختبار الطويل، كما يتبين وجود فروق دالة إحصائياً بين متوسط درجات عينة الطلبة الجامعيين على الاختبارين الطويل والقصير، لصالح متوسط عينة الطلبة الجامعيين على الاختبار الطويل، ما يعني أن هناك أثراً دالاً لطول الاختبار على متوسطات درجات العينتين.

وللإجابة عن السؤال الفرعي الرابع: قامت الباحثة باختبار الفروق بين متوسطات درجات مقياس الهوس الخفيف، لكلٍ من عينتي طلبة الماجستير والطلبة الجامعيين، لنوعي بنود الاختبار الطويل والقصير، من خلال إجراء تحليل التباين الأحادي (ANOVA)، والجدول الآتي يبين نتائج هذا التحليل:

الجدول رقم 50: نتائج تحليل التباين الأحادي لمتوسطات درجات مقياس الهوس الخفيف لكلٍ من عينتي طلبة الماجستير والطلبة الجامعيين لنوعي بنود الاختبار الطويل والقصير

العينة	حجم العينة	المتوسط الحسابي	الانحراف المعياري	مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	اختبار F	Sig
طلبة ماجستير طويل	206	22.165	3.87	بين المجموعات	4597.25	3	1532.41	98.689	0.000
طلبة جامعيون طويل	206	22.912	4.77	ضمن المجموعات	12732.70	820	15.52		
طلبة ماجستير قصير	206	17.577	3.17	المجموع	17329.96	823			
طلبة جامعيون قصير	206	18.145	3.77						

يتبين من الجدول السابق: وجود فروق ذات دلالة إحصائية عند مستوى الدلالة (0.05)، في متوسطات درجات المقياس العيادي الهوس الخفيف، بين كلٍ من عينتي طلبة الماجستير والطلبة الجامعيين لنوعي بنود الاختبار الطويل والقصير، وقد قامت الباحثة بحساب حجم التأثير للفروق بين المتوسطات باستخدام إحصاء إيتا تربيع كما يلي:

قيمة إحصاء إيتا تربيع= مجموع المربعات البينية / المجموع الكلي للتربيع (بالانت، 2006، ص 246).

وهكذا كانت قيمة إيتا تربيع = (0.26)، ويُلاحَظ أن هذه القيمة – وبالرجوع إلى إرشادات بالانت لتفسير قوة قيم إيتا تربيع– تدل على حجم أثر كبير.

وللكشف عن هذه الفروق تمّ إجراء عمليات المقارنات البعدية، وفق اختبار Dunnett's (C) للمجموعات غير المتجانسة، لأن قيمة Sig لاختبار Leven بلغت (0.000)، أي توجد فروق دالة إحصائياً للتجانس بين العينات، وبالتالي العينات غير متجانسة، والجدول الآتي يوضح نتائج هذه المقارنات:

الجدول رقم 51: نتائج المقارنات البعدية بين متوسطات درجات مقياس الهوس الخفيف لكلٍ من عينتي طلبة الماجستير والطلبة الجامعيين، لنوعي بنود الاختبار الطويل والقصير

المقياس	العينة (I)	العينة (j)	الفروق بين المتوسطات	الخطأ المعياري	مستوى الثقة 95%	
					الحد الأدنى	الحد الأعلى
الهوس	طلبة ماجستير طويل	طلبة جامعيون طويل	-0.74757	0.42841	-1.8573	0.3621
		طلبة ماجستير قصير	4.58738*	0.34896	3.6837	5.4911
		طلبة جامعيون قصير	4.01942*	0.37662	3.0439	4.9949
	طلبة جامعيون طويل	طلبة ماجستير قصير	5.33495*	0.39958	4.2999	6.3700
		طلبة جامعيون قصير	4.76699*	0.42400	3.6687	5.8653
	طلبة ماجستير قصير	طلبة جامعيون قصير	-0.56796	0.34347	-1.4576	0.3217

The mean difference is significant at the (0.05) level

يتبين من الجدول السابق: وجود فروق دالة إحصائياً بين متوسط درجات عينة طلبة الماجستير على الاختبارين الطويل والقصير، لصالح متوسط عينة طلبة الماجستير على الاختبار الطويل، كما يتبين وجود فروق دالة إحصائياً بين متوسط درجات عينة الطلبة الجامعيين على الاختبارين الطويل والقصير، لصالح متوسط عينة الطلبة الجامعيين على الاختبار الطويل، ما يعني أن هناك أثراً دالاً لطول الاختبار على متوسطات درجات العينتين.

<u>سادساً: الإجابة عن السؤال التاسع:</u> هل توجد فروق ذات دلالة إحصائية بين درجات قطع المقاييس العيادية لكلٍ من عينتي الأسوياء والمضطربين نفسياً، باختلاف طول الاختبار؟

للإجابة عن هذا السؤال تمّ حساب كلٍّ من درجات القطع الخام والتائية، العليا والدنيا، لنوعي المقاييس العيادية الطويلة والقصيرة، لكلٍ من عينتي الأسوياء والمضطربين نفسياً، وذلك لإمكانية المقارنة بينها.

والجداول الآتية توضح هذه الدرجات المحسوبة بطريقة المتوسط والانحراف المعياري*:

الجدول رقم 52: درجات القطع الخام لمقياس الاكتئاب العيادي، لكلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير

نوع المقياس	حجم العينة	المتوسط الحسابي	الانحراف المعياري	درجة القطع العليا	درجة القطع الدنيا
مضطربون نفسياً طويل	150	27.86	4.78	37.42	18.3
مضطربون نفسياً قصير	150	24.44	3.81	32.06	16.82
أسوياء طويل	150	31.73	4.66	41.05	22.41
أسوياء قصير	150	28.21	4.09	36.39	20.03

الجدول رقم 53: درجات القطع التائية لمقياس الاكتئاب العيادي، لكلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير

نوع المقياس	حجم العينة	المتوسط الحسابي	الانحراف المعياري	درجة القطع العليا		درجة القطع الدنيا	
				ذكر	أنثى	ذكر	أنثى
مضطربون نفسياً طويل	150	59.33	11.11	83	71	42	36
مضطربون نفسياً قصير	150	52.13	8.85	72	62	37	32
أسوياء طويل	150	64.14	8.69	91	79	50	43
أسوياء قصير	150	57.84	9.17	80	70	46	40

*** تم حساب درجتي القطع العليا والدنيا باستخدام المعادلتين الآتيتين:**

- **درجة القطع العليا= متوسط الدرجات + 2 × الانحراف معياري.**
- **درجة القطع الدنيا= متوسط الدرجات – 2 × الانحراف معياري.**

الجدول رقم 54: درجات القطع الخام للمقياس البارانويا العيادي، لكلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير

نوع المقياس	حجم العينة	المتوسط الحسابي	الانحراف المعياري	درجة القطع العليا	درجة القطع الدنيا
مضطربون نفسياً طويل	150	23.89	4.97	33.83	13.95
مضطربون نفسياً قصير	150	21.25	4.76	30.77	11.73
أسوياء طويل	150	18.51	4.42	27.35	9.67
أسوياء قصير	150	15.25	4.16	23.57	6.93

الجدول رقم 55: درجات القطع التائية لمقياس البارانويا العيادي، لكلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير

نوع المقياس	حجم العينة	المتوسط الحسابي	الانحراف المعياري	درجة القطع العليا		درجة القطع الدنيا	
				ذكر	أنثى	ذكر	أنثى
مضطربون نفسياً طويل	150	78.20	13.07	105	97	53	47
مضطربون نفسياً قصير	150	71.43	12.73	97	89	47	42
أسوياء طويل	150	62.95	10.72	89	82	42	37
أسوياء قصير	150	54.65	10.45	79	72	34	29

الجدول رقم 56: درجات القطع الخام للمقياس الفصام العيادي، لكلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير

نوع المقياس	حجم العينة	المتوسط الحسابي	الانحراف المعياري	درجة القطع العليا	درجة القطع الدنيا
مضطربون نفسياً طويل	150	44.66	11.21	67.08	22.24
مضطربون نفسياً قصير	150	38.98	10.04	59.06	18.9
أسوياء طويل	150	32.28	7.47	47.22	17.34
أسوياء قصير	150	26.33	7.11	40.55	12.11

الجدول رقم 57: درجات القطع التائية لمقياس الفصام العيادي، لكلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير

نوع المقياس	حجم العينة	المتوسط الحسابي	الانحراف المعياري	درجة القطع العليا		درجة القطع الدنيا	
				ذكر	أنثى	ذكر	أنثى
مضطربون نفسياً طويل	150	72.42	11.95	96	93	53	49
مضطربون نفسياً قصير	150	67.43	9.60	88	85	49	46
أسوياء طويل	150	60.32	6.94	77	73	48	45
أسوياء قصير	150	54.63	6.59	70	66	43	40

الجدول رقم 58: درجات القطع الخام لمقياس الهوس العيادي، لكلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير

نوع المقياس	حجم العينة	المتوسط الحسابي	الانحراف المعياري	درجة القطع العليا	درجة القطع الدنيا
مضطربون نفسياً طويل	150	30.16	5.98	42.12	18.2
مضطربون نفسياً قصير	150	22.50	4.33	31.16	13.84
أسوياء طويل	150	22.77	4.82	32.41	13.13
أسوياء قصير	150	18.09	3.73	25.55	10.63

الجدول رقم 59: درجات القطع التائية لمقياس الهوس العيادي، لكلٍ من عينتي الأسوياء والمضطربين نفسياً لنوعي بنود الاختبار الطويل والقصير

نوع المقياس	حجم العينة	المتوسط الحسابي	الانحراف المعياري	درجة القطع العليا		درجة القطع الدنيا	
				ذكر	أنثى	ذكر	أنثى
مضطربون نفسياً طويل	150	73.65	13.15	97	101	48	46
مضطربون نفسياً قصير	150	57.01	9.63	75	76	38	35
أسوياء طويل	150	57.29	10.92	77	78	38	35
أسوياء قصير	150	46.86	8.42	63	62	31	28

يتبين من الجداول السابقة: هبوط قيم درجات القطع في الاختبارات القصيرة بمجملها عن مثيلاتها في الاختبارات الطويلة، ونلاحظ أيضاً في معظم الأحيان أن المدى بين درجات القطع العليا والدنيا في الاختبارات الطويلة كان أوسع من نظيره في الاختبارات القصيرة، وبالتالي يمكن القول بأن درجات القطع في الاختبارات الطويلة يمكن أن تكون أكثر دقة من حيث قدرتها على إعطاء مؤشرات إكلينيكية من مثيلاتها في الاختبارات القصيرة.

ولمعرفة فيما إذا كانت هذه الفروق دالة إحصائياً بين درجات قطع المقاييس الطويلة ومثيلاتها للمقاييس القصيرة، قامت الباحثة باستخدام الاختبار الاحصائي (t) ستيودنت، للمقارنة بين متوسطات درجات القطع للاختبارات الطويلة والقصيرة، والجدول الآتي يبين نتائج هذا الاختبار:

الجدول رقم 60: الفروقات بين متوسطات درجات القطع لعينتي الأسوياء والمضطربين نفسياً، لكلٍ من المقاييس الطويلة والقصيرة

المقياس	نوع المقياس	المتوسط	الانحراف المعياري	الخطأ المعياري	قيمة t *	مستوى الدلالة	القرار
الاكتئاب	المقياس الطويل	29.79	11.12	5.56	0.47	0.65	غير دال
	المقياس القصير	26.32	9.38	4.69			
البارانويا	المقياس الطويل	21.20	11.29	5.64	0.37	0.72	غير دال
	المقياس القصير	18.25	10.88	5.44			
الفصام	المقياس الطويل	38.47	23.12	11.56	0.36	0.72	غير دال
	المقياس القصير	32.65	21.37	10.68			
الهوس الخفيف	المقياس الطويل	26.46	13.24	6.62	0.75	0.48	غير دال
	المقياس القصير	20.29	9.67	4.83			

* عدد درجات الحرية هو (1)

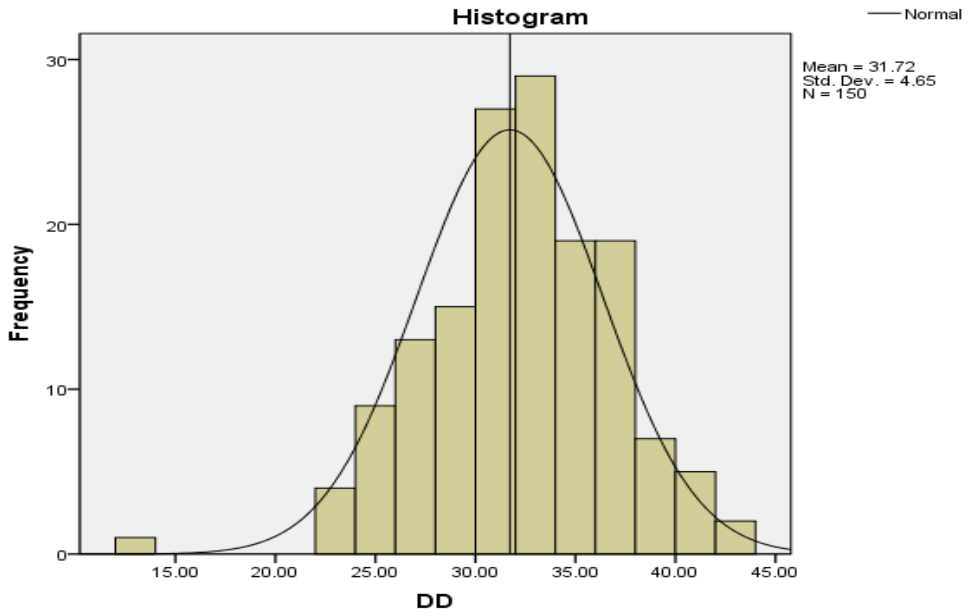
يتبين من الجدول السابق أنَّه على الرغم من هبوط قيم درجات القطع في الاختبارات القصيرة عن مثيلاتها في الاختبارات الطويلة، إلا أن قيمة هذه الفروق لم تصل إلى مستوى الدلالة الإحصائية.

والجدول الآتي يبين الإحصاءات الوصفية لدرجات كل من عينتي الأسوياء والمضطربين نفسياً، على المقاييس العيادية الأربعة، لنوعي بنود الاختبار الطويل والقصير:

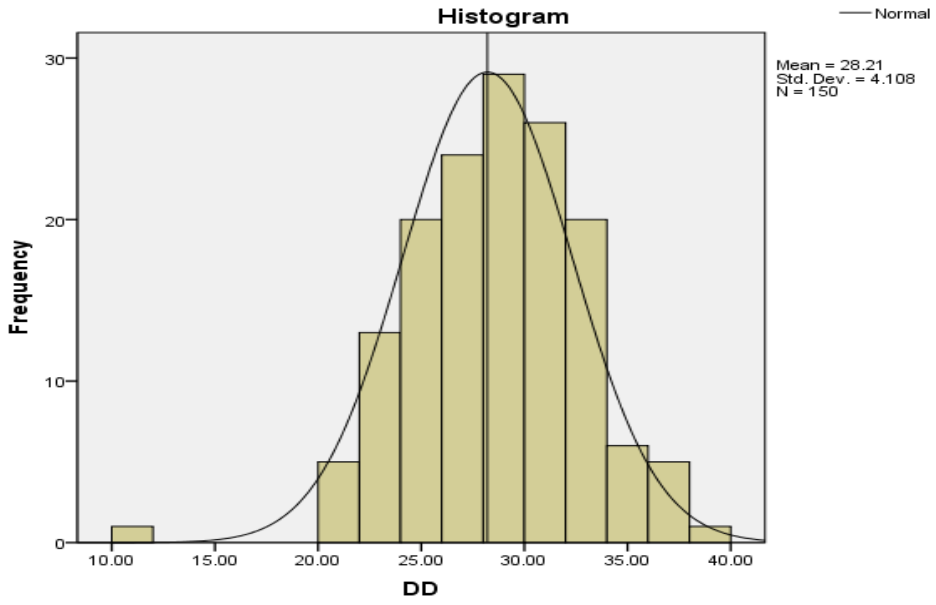
الجدول رقم 61: الإحصاءات الوصفية لدرجات كل من عينتي الأسوياء والمضطربين نفسياً، على المقاييس العيادية الأربعة، لنوعي بنود الاختبار الطويل والقصير

نوع المقياس	العينة	المتوسط	الوسيط	المنوال	الانحراف المعياري	الالتواء	التفرطح
الاكتئاب الطويل	الأسوياء	31.72	32.00	32.00	4.65	-0.364	0.895
الاكتئاب القصير		28.20	28.00	27.00	4.108	-0.378	1.213
الاكتئاب الطويل	المرضى	27.86	29.00	29.00	4.78	-0.103	-0.574
الاكتئاب القصير		24.44	25.00	25.00	3.81	0.084	-0.318
البارانويا الطويل	الأسوياء	18.52	18.50	17.00	4.41	0.148	-0.564
البارانويا القصير		15.22	15.00	14.00	4.161	0.217	-0.491
البارانويا الطويل	المرضى	23.89	24.00	25.00	4.978	-0.604	0.562
البارانويا القصير		21.25	22.00	22.00	4.76	-0.870	0.959
الفصام الطويل	الأسوياء	32.28	31.00	29.00	7.45	0.426	0.405
الفصام القصير		26.29	25.00	24.00	7.12	0.514	0.215
الفصام الطويل	المرضى	44.66	42.00	42.00	11.21	0.226	-1.035
الفصام القصير		38.98	37.00	36.00	10.04	0.124	-1.020
الهوس الطويل	الأسوياء	22.76	23.00	22.00	4.81	-0.143	0.346
الهوس القصير		18.08	18.00	17.00	3.73	-0.179	0.965
الهوس الطويل	المرضى	30.166	31.00	32.00	5.982	-0.132	-0.656
الهوس القصير		22.50	23.00	24.00	4.33	-0.135	-0.570

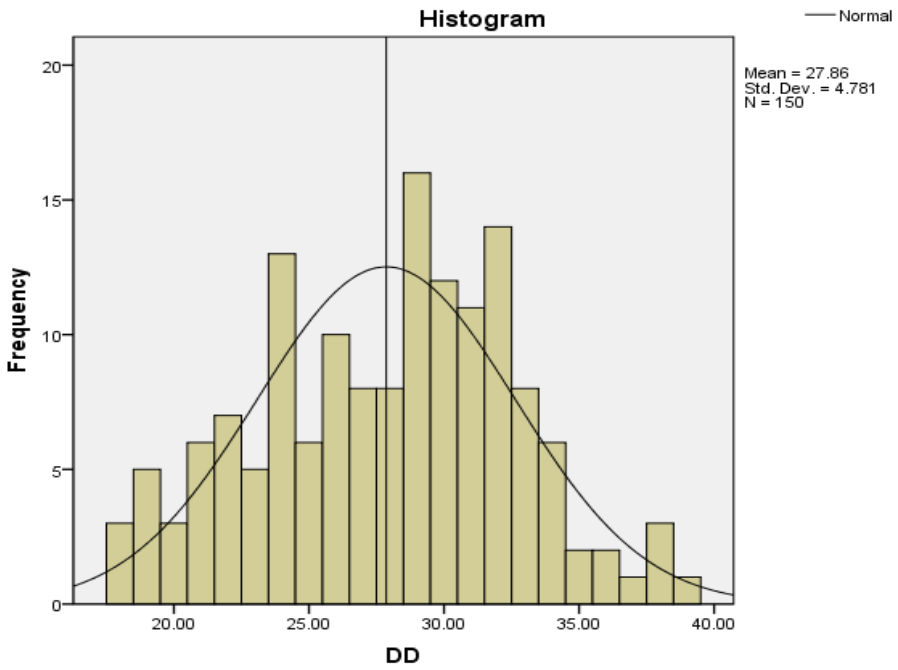
كما تمثل الأشكال البيانية الآتية المدرجات والمنحنيات التكرارية لدرجات كل من عينتي الأسوياء والمضطربين نفسياً، على المقاييس العيادية الأربعة، لنوعي بنود الاختبار الطويل والقصير:



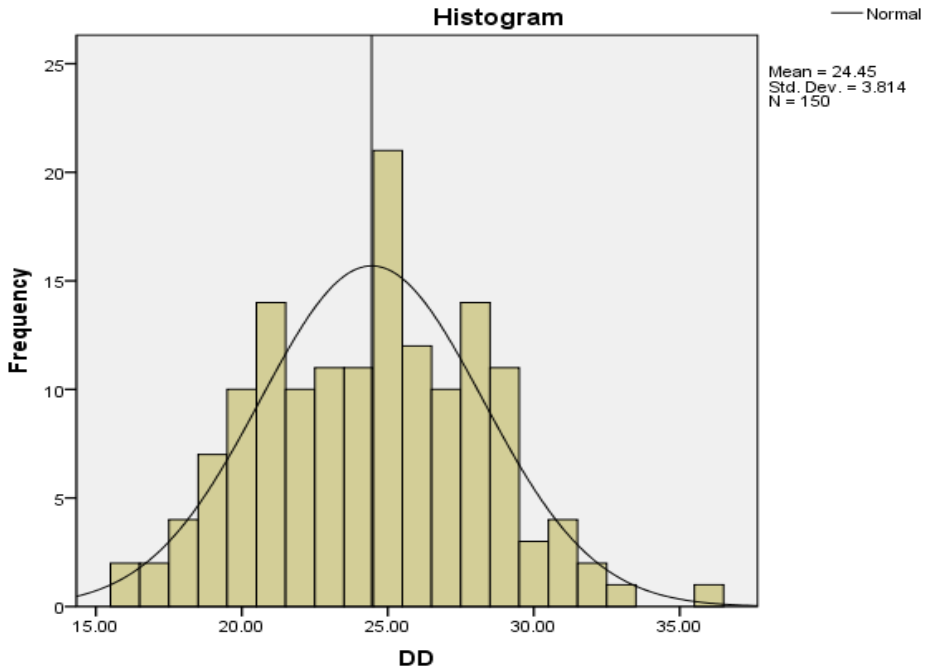
الشكل رقم 3: المدرج التكراري لدرجات مقياس الاكتئاب الطويل لعينة الأسوياء



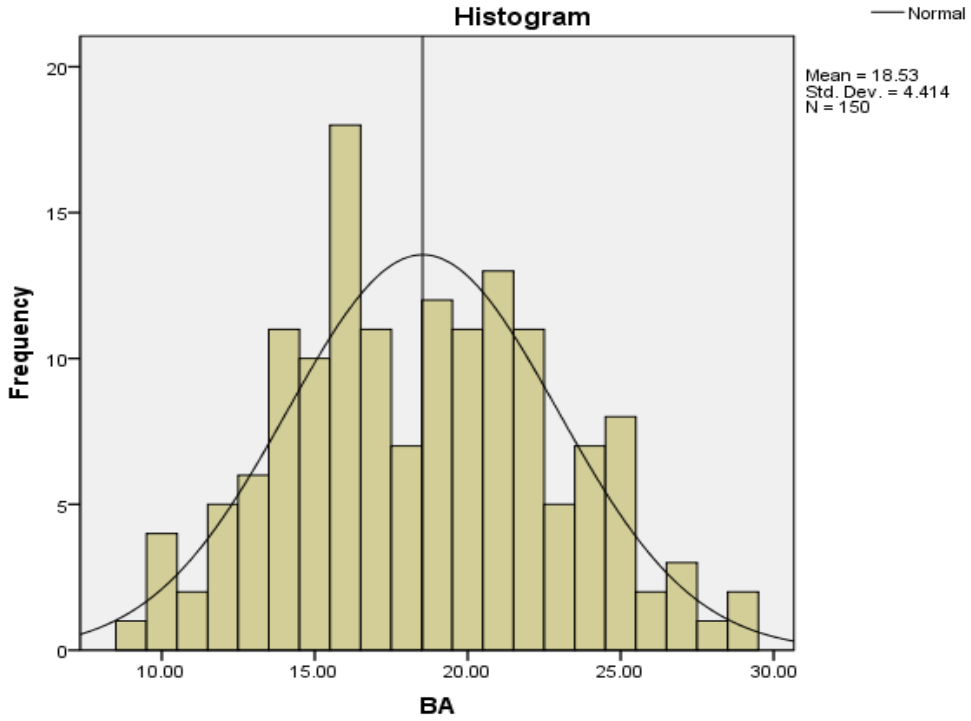
الشكل رقم 4: المدرج التكراري لدرجات مقياس الاكتئاب القصير لعينة الأسوياء



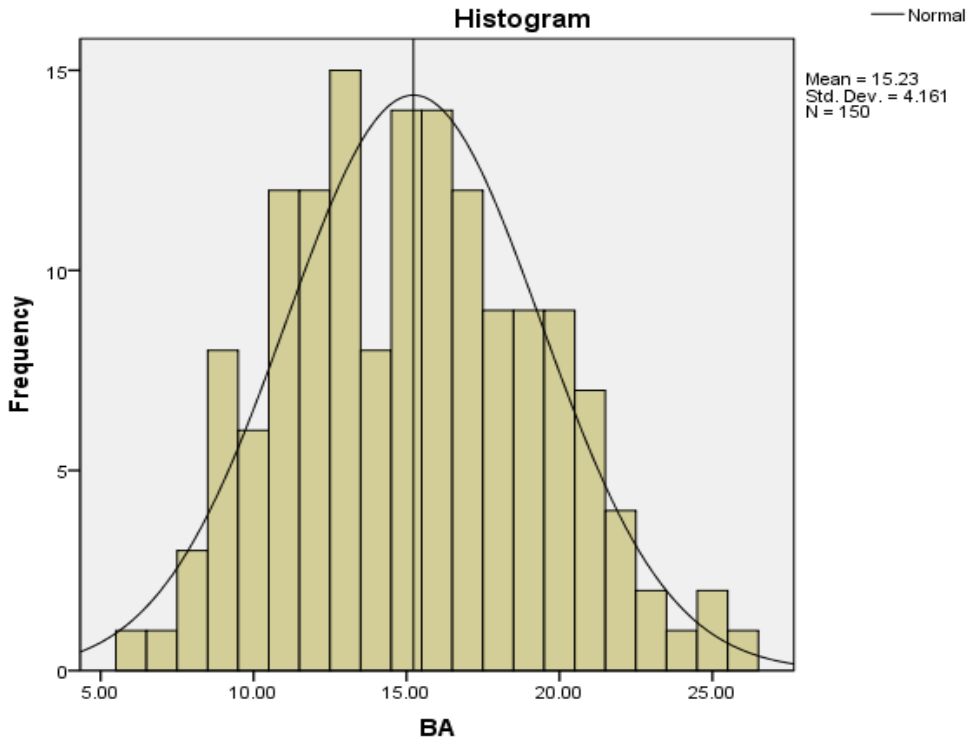
الشكل رقم 5: المدرج التكراري لدرجات مقياس الاكتئاب الطويل لعينة المضطربين نفسياً



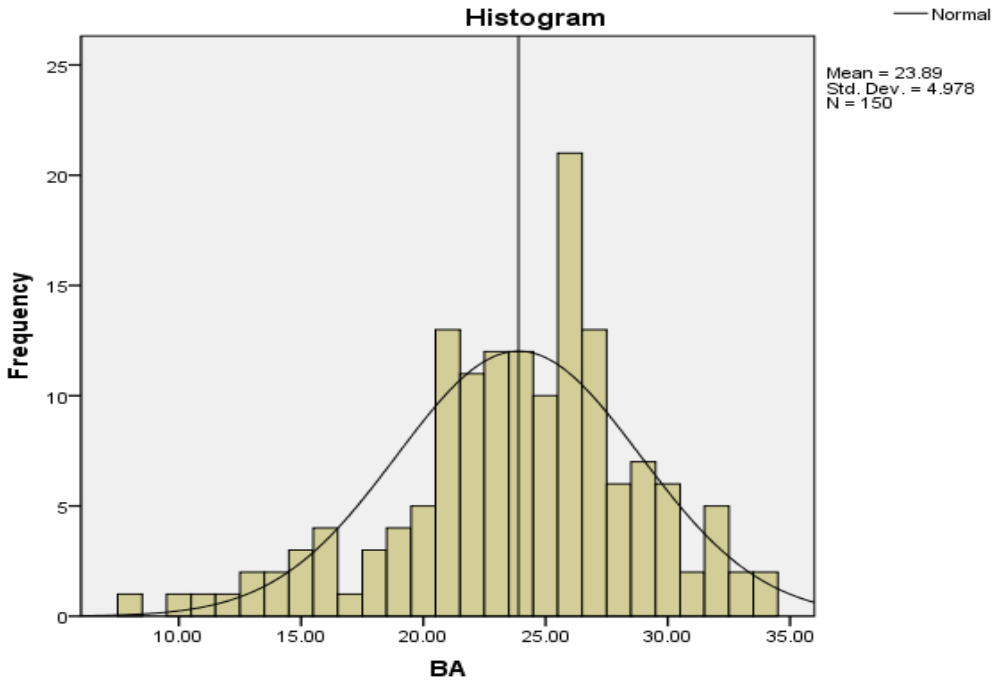
الشكل رقم 6: المدرج التكراري لدرجات مقياس الاكتئاب القصير لعينة المضطربين نفسياً



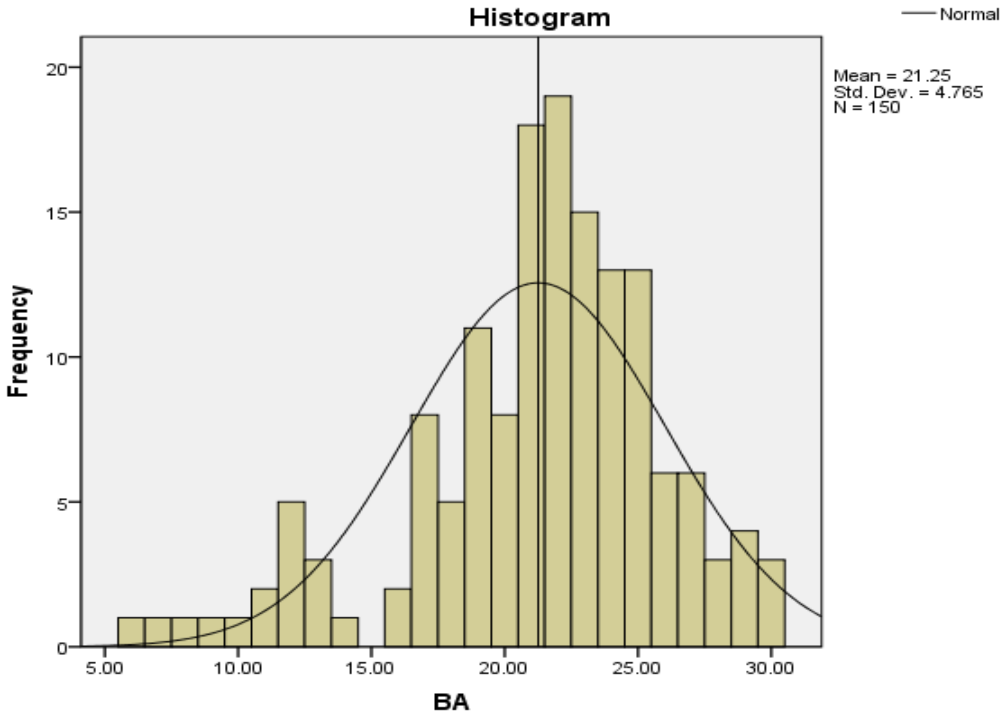
الشكل رقم 7: المدرج التكراري لدرجات مقياس البارانويا الطويل لعينة الأسوياء



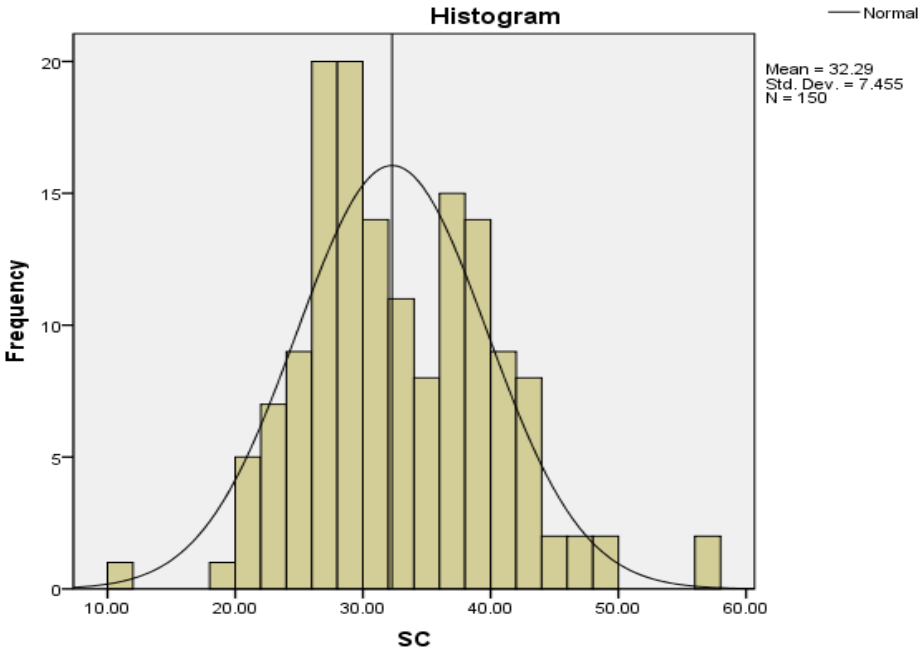
الشكل رقم 8: المدرج التكراري لدرجات مقياس البارانويا القصير لعينة الأسوياء



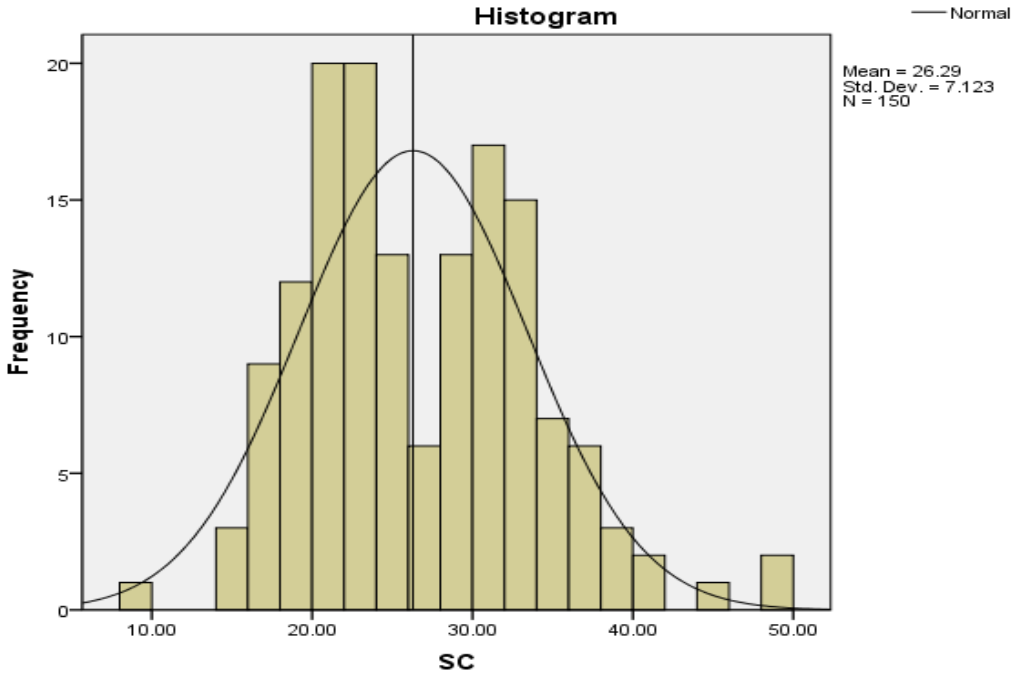
الشكل رقم 9: المدرج التكراري لدرجات مقياس البارانويا الطويل لعينة المضطربين نفسياً



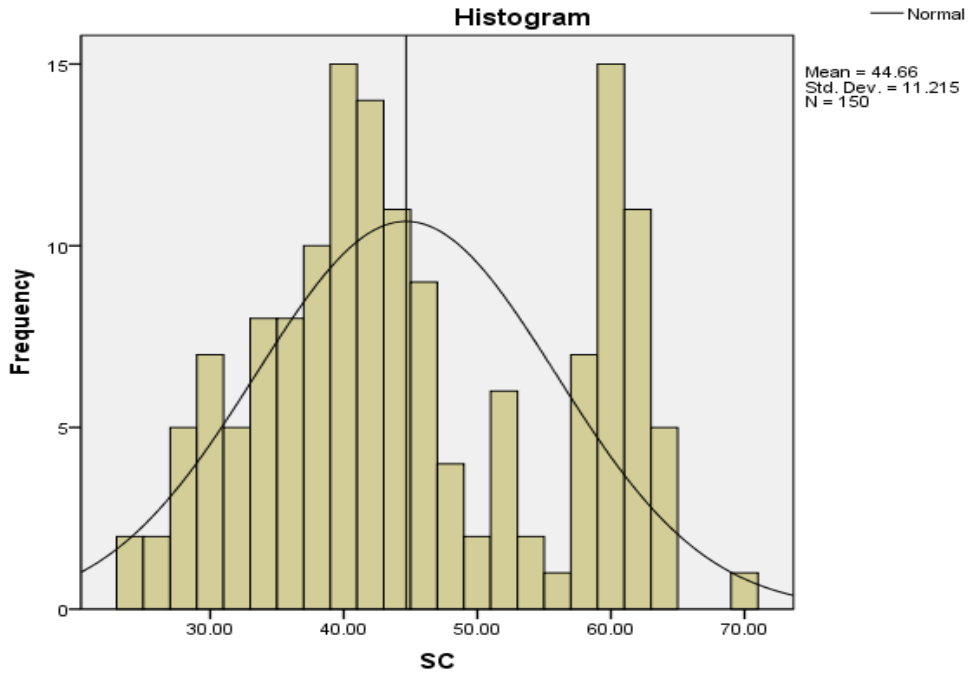
الشكل رقم 10: المدرج التكراري لدرجات مقياس البارانويا القصير لعينة المضطربين نفسياً



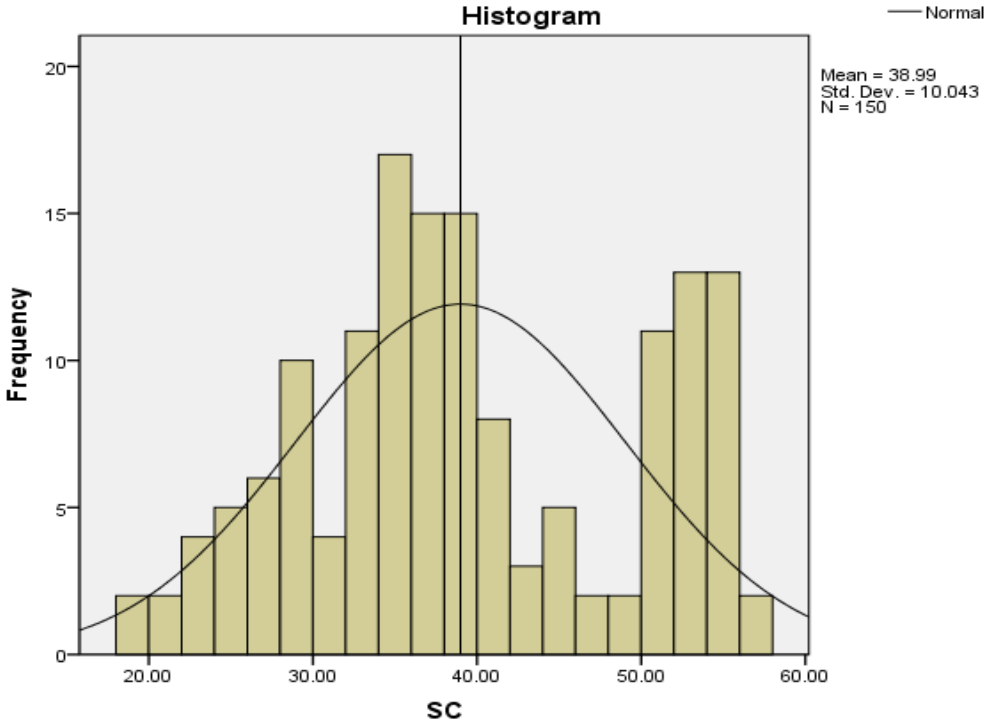
الشكل رقم 11: المدرج التكراري لدرجات مقياس الفصام الطويل لعينة الأسوياء



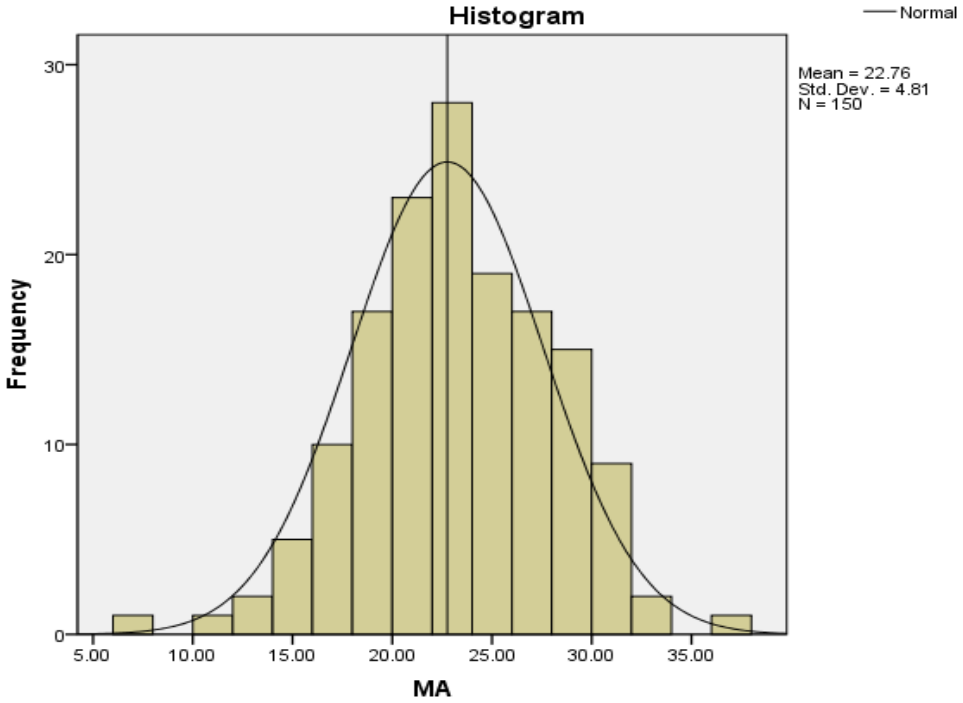
الشكل رقم 12: المدرج التكراري لدرجات مقياس الفصام القصير لعينة الأسوياء



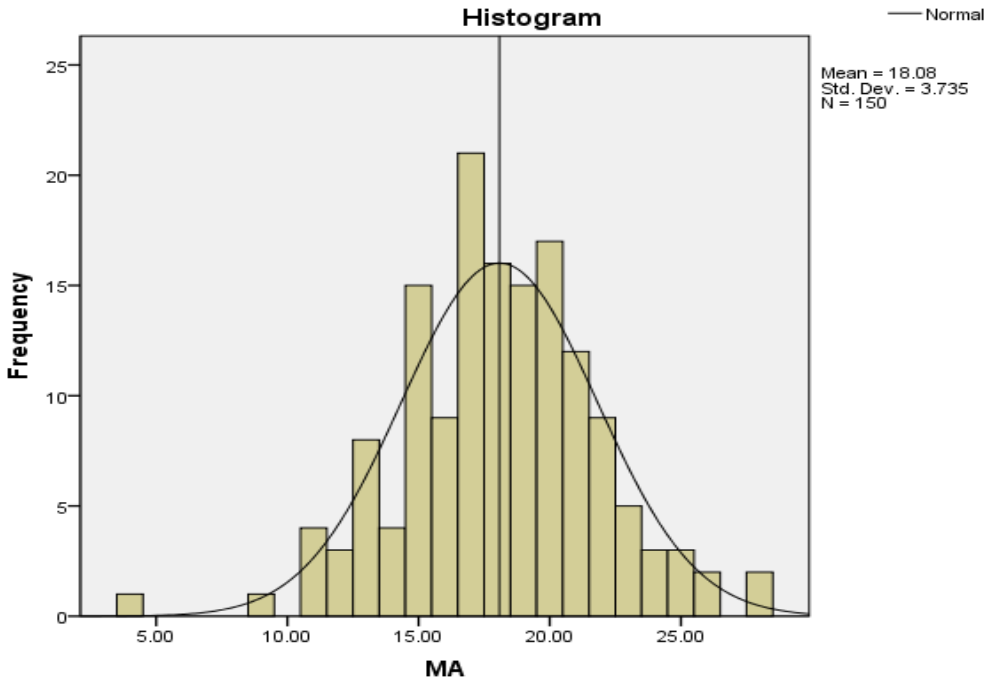
الشكل رقم 13: المدرج التكراري لدرجات مقياس الفصام الطويل لعينة المضطربين نفسياً



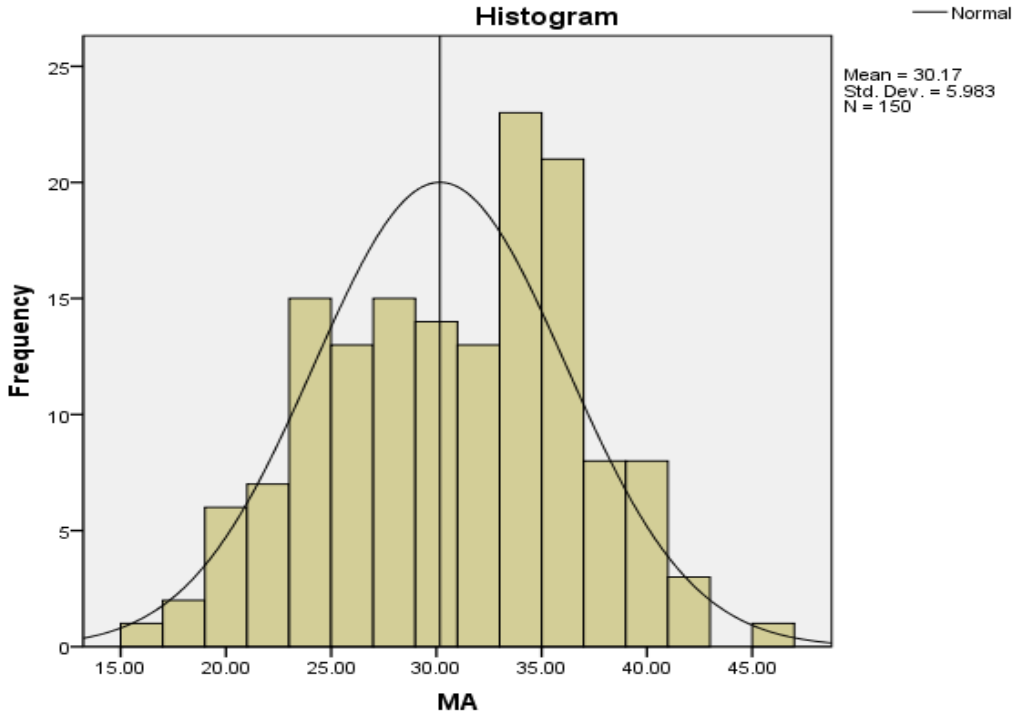
الشكل رقم 14: المدرج التكراري لدرجات مقياس الفصام القصير لعينة المضطربين نفسياً



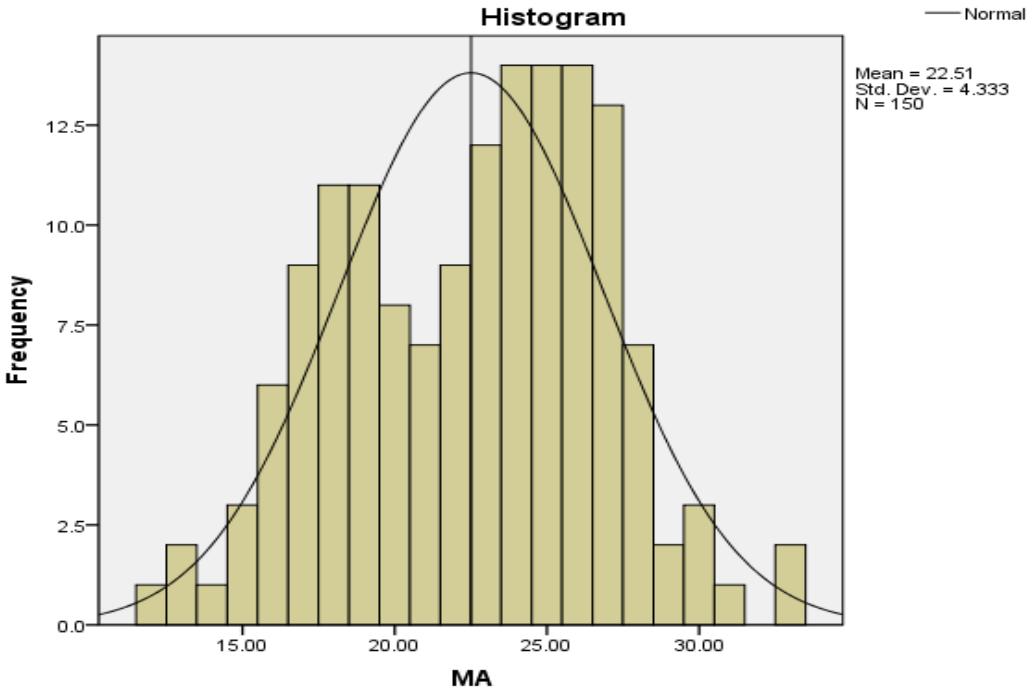
الشكل رقم 15: المدرج التكراري لدرجات مقياس الهوس الطويل لعينة الأسوياء



الشكل رقم 16: المدرج التكراري لدرجات مقياس الهوس القصير لعينة الأسوياء



الشكل رقم 17: المدرج التكراري لدرجات مقياس الهوس الطويل لعينة المضطربين نفسياً



الشكل رقم 18: المدرج التكراري لدرجات مقياس الهوس القصير لعينة المضطربين نفسياً

يُلاحظ من الجدول والمدرجات التكرارية السابقة: هبوط قيم مقاييس النزعة المركزية في المقاييس القصيرة عن مثيلاتها في المقاييس الطويلة، كما يُلاحظ اقتراب قيم مقاييس النزعة المركزية،

(المتوسط والوسيط والمنوال) بالنسبة لكل عينة عند كل مقياس من المقاييس العيادية الطويلة والقصيرة، مما يدل على أن توزع درجات المقاييس العيادية تقترب من التوزيع الطبيعي إلى حدٍ كبير، كما يُلاحظ أن جميع الالتواءات كانت قريبة من الصفر، حيث وقعت هذه الالتواءات ضمن الحدود الطبيعية التي تتراوح بين (+1، −1)، ومعامل التفرطح داخل المجال من (+3، −3)، مما سبق يمكن القول أن درجات المقاييس تتوزع بين أفراد العينة توزعاً طبيعياً اعتدالياً.

سابعاً: الإجابة عن السؤال العاشر: ما أثر التفاعل بين طول الاختبار وخصائص العينة في الخصائص السيكومترية للاختبار؟

تمّت الإجابة عن معظم جوانب هذا السؤال من خلال الإجابة عن الأسئلة السابقة، وفيما يلي عرض أهم النتائج المتعلقة به:

أظهرت نتائج الاختبار الإحصائي (V) لدراسة دلالة الفروق بين معاملات ثبات الإعادة ما يلي:

وجود فروق دالة إحصائياً بين معاملي ثبات الإعادة لمقياس البارانويا الطويل مع نظيره القصير لصالح معامل الثبات الطويل عند حجمي العينة (425) و(567). وعدم وجود فروق دالة إحصائياً بين معاملي ثبات هذا المقياس عند حجم العينة (284)، كما أظهرت نتائج هذا الاختبار وجود فروق دالة إحصائياً بين معاملي ثبات الإعادة لمقياس الهوس الطويل مع نظيره القصير عند أحجام العينات الثلاث، وهذه الفروق بمجملها لصالح معامل ثبات المقياس الطويل، أما بالنسبة لمقياسي الاكتئاب والفصام فقد بينت نتائج هذا الاختبار عدم وجود فروق دالة إحصائياً بين معاملات ثبات هذه المقاييس الطويلة والقصيرة عند أحجام العينات المختلفة.

يُلاحظ أن عدد البنود وحجم العينة يؤثران معاً في ارتفاع معامل الثبات وانخفاضه، حيث أن الاختبار الطويل مع حجم العينة الأكبر أعطى دلالة إحصائية هامة للفروق بين معاملات الثبات.

كما أظهرت نتائج الاختبار الإحصائي(V) لدراسة دلالة الفروق بين معاملات الصدق المحكي مايلي:

وجود فروق دالة إحصائياً بين معاملي الصدق المحكي لمقياس البارانويا الطويل والقصير لصالح معامل المقياس الطويل عند حجم العينة (567)، وعدم وجود فروق دالة إحصائياً بين معاملي الصدق المحكي لهذا المقياس عند حجمي العينة (425) و (284). وكذلك بينت نتائج هذا الاختبار وجود فروق دالة إحصائياً بين معاملات الصدق المحكي لمقياس الفصام الطويل مع نظيره القصير عند حجمي العينة (567) و(425)، وهذه الفروق دالة لصالح معاملات المقياس الطويل، في حين لم تكن هنالك فروق دالة إحصائياً بين معاملي ارتباط هذا المقياس الطويل والقصير عند حجم العينة (284)، أما بالنسبة لمقياس الاكتئاب فقد بينت النتائج عدم وجود فروق دالة إحصائياً بين معاملات الصدق المحكي لهذه المقياس الطويل والقصير عند أحجام العينات

المختلفة. ويُلاحظ أنه مع انخفاض حجم العينة تقل قيم الفرق بين معاملات الارتباط التي تمثل معاملات الصدق المحكي ، وتغدو الفروق بينها غير دالة إحصائياً؛ مما يعني أن حجم العينة وعدد البنود يؤثران معاً في ارتفاع معامل الصدق وانخفاضه.

وأظهرت نتائج الاختبار الإحصائي (M) الخاص بمعرفة دلالة الفروق بين عدة معاملات ألفا كرونباخ ما يلي:

وجود فروق دالة إحصائياً بين معاملات ثبات ألفا لمقياس الهوس الطويل والقصير عند أحجام العينات 50% و100% و150% و200% من عدد بنود الرائز، كانت هذه الفروق بمجملها لصالح معامل الثبات الطويل، في حين لم تكن هنالك فروق دالة إحصائياً بين معاملات ثبات المقاييس الطويلة والقصيرة بالنسبة للمقاييس العيادية: الاكتئاب والبارانويا والفصام عند أحجام العينات المختلفة، كما أوضحت نتائج هذا الاختبار وجود فروق دالة إحصائياً بالنسبة لمقياس الهوس الطويل بين أحجام العينات المختلفة، وهذه الفروق لصالح معامل ثبات العينة الأصغر (284) والتي تبلغ نسبتها (50%) من عدد بنود الاختبار، كما يُلاحظ ارتفاع قيمة معامل ثبات ألفا مع انخفاض حجم العينة.

يُلاحظ هنا وبالنسبة لمعامل ثبات ألفا بأن الاختبار الطويل مع حجم العينة الأصغر أعطى العلاقة ذات الدلالة الإحصائية تتفق هذه النتيجة مع النتائج التي توصلت لها دراسة لوزانو (2008) والتي أوضحت أن أفضل نتائج ثبات لمقاييس التصنيف المُعتمدة من قبل الباحث كانت عند حجم العينة (50)، وبدأت بالتناقص عند أحجام العينات (200)، و(300)، و(500). وقد تمّ الإشارة في العديد من أدبيات القياس والتقويم إلى أن استخدام عينات ذات حجوم صغيرة في بعض البحوث ربما كان أفضل من استخدام عينات ذات أحجام كبيرة مثل الدراسات التي تعتمد على التحليل النفسي وأدوات القياس الإسقاطي، وأن العينات كبيرة الحجم في الدراسات التجريبية تُعد ضعيفة لأنه سيكون هناك عدد من المتغيرات غير المضبوطة (درويش ورحمة، 2011، ص55).

أظهرت نتائج اختبار تحليل التباين وحساب حجم الأثر بين نوعي بنود الاختبار الطويل والقصير، بالنسبة لعينتي الأسوياء والمضطربين نفسياً ما يلي:

وجود فروق دالة إحصائياً بين متوسط درجات عينة المضطربين نفسياً على الاختبارين الطويل والقصير، لصالح متوسط عينة المضطربين نفسياً على الاختبار الطويل، وكذلك وجود فروق دالة إحصائياً بين متوسط درجات عينة الأسوياء على الاختبارين الطويل والقصير، لصالح متوسط عينة الأسوياء على الاختبار الطويل، ما يعني أن هناك أثراً دالاً لطول الاختبار في متوسط درجات العينتين، وبالتالي على معاملات صدق الاختبار.

أظهرت نتائج اختبار تحليل التباين وحساب حجم الأثر بين نوعي بنود الاختبار الطويل والقصير بالنسبة لمتغير المرحلة الدراسية / طلبة التعليم الجامعي، وطلبة الماجستير/ ما يلي:

وجود فروق دالة إحصائياً بين متوسط درجات عينة طلبة الماجستير على الاختبارين الطويل والقصير، لصالح متوسط عينة طلبة الماجستير على الاختبار الطويل، وكذلك وجود فروق دالة إحصائياً بين متوسط درجات عينة الطلبة الجامعيين على الاختبارين الطويل والقصير، لصالح متوسط عينة الطلبة الجامعيين على الاختبار الطويل، ما يعني أن هناك أثراً دالاً لطول الاختبار في متوسطات درجات العينتين.

ثامناً: الإجابة عن السؤال الحادي عشر: ما أثر التفاعل بين طول الاختبار وخصائص العينة* في درجات القطع؟

أظهرت النتائج أن الفروق بين درجات قطع الاختبارات الطويلة والقصيرة لعينتي الأسوياء والمضطربين نفسياً لم تصل لمستوى الدلالة الإحصائية، أي أنه لم يترتب على زيادة طول الاختبار ارتفاع أو تدنّ دال إحصائياً في درجات القطع.

تاسعاً: مناقشة نتائج البحث:

مناقشة النتائج المتعلقة بالسؤال الأول: هل توجد فروق ذات دلالة إحصائية بين معاملات ثبات درجات الاختبار بالإعادة باختلاف طول الاختبار؟

أظهرت نتائج الاختبار الإحصائي (V) وجود فروق دالة إحصائياً بين معاملي ثبات الإعادة لمقياس البارانويا الطويل مع نظيره القصير لصالح معامل الثبات الطويل عند حجمي العينة (425) و(567). وعدم وجود فروق دالة عند حجم العينة (284)، كما أظهرت نتائج هذا الاختبار وجود فروق دالة إحصائياً بين معاملي ثبات الإعادة لمقياس الهوس الطويل مع نظيره القصير عند أحجام العينات الثلاث، وهذه الفروق بمجملها لصالح معامل ثبات المقياس الطويل، في حين بينت النتائج بالنسبة لمقياسي الاكتئاب والفصام عدم وجود فروق دالة إحصائياً بين معاملات ثبات هذه المقاييس الطويلة والقصيرة عند أحجام العينات المختلفة. وقد يُعزى ذلك إلى زيادة نسبة الفرق في عدد البنود بين كلٍ من مقياسي البارانويا والهوس الخفيف الطويل ونظيره القصير، حيث تصل هذه النسبة في مقياس البارانويا إلى (15%)، وفي مقياس الهوس الخفيف إلى(26.08%)، وبالتالي أدى ذلك إلى زيادة نسبة الفرق في معاملات الارتباط بين الاختبارات الطويلة والقصيرة لهذين المقياسين.

*** خاصية العينة المدروسة في هذا السؤال هي متغير الحالة الصحية؛ أي أن الباحثة درست أثر التفاعل بين طول الاختبار والحالة الصحية في درجات القطع.**

يُلاحظ مما سبق أن حجم العينة وعدد البنود يؤثران معاً في ارتفاع معامل الثبات وانخفاضه، حيث أن الاختبار الطويل مع حجم العينة الأكبر أعطى دلالة إحصائية هامة للفروق بين معاملات الثبات، ويمكن تفسير هذه النتيجة بأن زيادة حجم العينة أعطى دلالة إحصائية، لأن زيادة الحجم يؤدي إلى الحصول على معلومات أكثر، وبالتالي يقود إلى دقة أكبر في تلك المعلومات مما يُقلص مقدار الخطأ المعياري للقياس (ملحم، 2001، ص155).

كما أن زيادة عدد البنود يؤدي إلى الحصول على عينة أكبر من السلوك، وبالتالي كلما كانت العينة أكبر كان من المتوقع أن يُمثل فيها بشكل مستقر العدد الأكبر من مكونات السلوك أو القدرة المقيسة، والقابل للظهور في مرتي التطبيق (فرج، 2007، ص347). ووفقاً لمعادلة سبيرمان براون فإن ثبات درجات الاختبار يزداد بازدياد طول الاختبار بشرط أن تقيس أسئلته ومفرداته التي أدّت إلى زيادة الطول السمة نفسها التي تقيسها بقية مفردات الاختبار (علام، 2006، ص97).

هنا نجد العلاقة الارتباطية بين عدد البنود الكبير وحجم العينة الكافي من حيث علاقتهما برفع قيمة معامل الثبات، حيث أن عدد البنود الأكبر مع حجم العينة الأكبر أنتج علاقة ارتباطية ذات أهمية ودلالة إحصائية.

تتفق هذه النتيجة مع دراسة كلٍ من السكت (2013) التي بينت أن الصورة الطويلة (R) ذات ال 100 بنداً وحجم العينة (100) و(200) أعلى قيم لمعاملات الثبات، كما تتفق مع نتيجة دراسة هيلمان (2006) التي بينت أن أما مقياس (Spos) الأصلي الأطول ذو البعد الواحد يحقق ثباتاً أعلى من النسخة المختصرة ، كما خلصت نتائج هاتين الدراستين إلى أن عدد البنود عندما أدخل مع حجم العينة إلى الأدوات أنتجت علاقة ارتباطية ذات أهمية إحصائية، وتتفق أيضاً مع نتائج دراسة كل من الأحمدي (2008) التي أوضحت وجود حجم تأثير دال لطول الاختبار وحجم العينة في دقة واتساق النتائج، ودراسة عوض (2004) التي بينت وجود فروق ذات دلالة إحصائية عند مستوى الدلالة (0،05) في قيمة معامل الثبات المحسوبة لكل من صورتي الاختبار التحصيلي المُعد لهذه الدراسة لصالح الصورة الأطول.

مناقشة النتائج المتعلقة بالسؤال الثاني: هل توجد فروق ذات دلالة إحصائية بين معاملات ثبات الإعادة لدرجات الاختبار باختلاف حجم العينة؟

أظهرت النتائج عدم وجود فروق دالة إحصائياً في معاملات الثبات بين كل عينتين من عينات البحث في المقاييس العيادية الطويلة والقصيرة، حيث يمكن أن يعود ذلك وبالرجوع إلى الجدول رقم (14) إلى اقتراب قيم الإحصاءات الوصفية لهذه العينات والتي تتضمن قيم المتوسط والانحراف المعياري والتباين والخطأ المعياري وأيضاً قيم الالتواء والتفرطح، كما يمكن أن يعود ذلك لعدم وجود فروقات في التجانس بين هذه العينات ذات الأحجام المختلفة.

حيث قامت الباحثة بحساب التجانس بين هذه العينات عند كل مقياس من المقاييس العيادية، باستخدام اختبار Leven للتجانس، وقد بينت النتائج أن قيمة ال Sig لهذا الاختبار بلغت (0.736) بالنسبة لمقياس الاكتئاب، و (0.787) بالنسبة لمقياس البارانويا، و(0.400) بالنسبة لمقياس الفصام، في حين بلغت(0.420) بالنسبة لمقياس الهوس الخفيف، وهذه القيم أكبر من (0.05) وبالتالي لاتوجد فروق دالة إحصائياً للتجانس بين هذه العينات بالنسبة للمقاييس العيادية الأربعة، مما يدل على تجانس هذه العينات، تتعارض هذه النتيجة مع نتائج دراسة الأحمدي (2008) التي أظهرت وجود حجم تأثير دال لحجم العينة على دقة واتساق النتائج، كما تتعارض مع نتيجة دراسة السكت (2013) التي بينت أن حجم العينة الكبير يزيد من الثبات كما أوضحت وجود فروق دالة إحصائياً في معامل الثبات للبعد (N) عند المقارنة بين حجمي العينة (24 و 48) لصالح العينة الأكبر.

مناقشة النتائج المتعلقة بالسؤال الثالث: هل توجد فروق ذات دلالة إحصائية بين معاملات ثبات الاتساق الداخلي لدرجات الاختبار باختلاف طول الاختبار؟

أظهرت نتائج الاختبار الإحصائي (M) عدم وجود فروق دالة إحصائياً بين معاملات ثبات المقاييس الطويلة والقصيرة عند مستوى الدلالة (0.05)، بالنسبة للمقاييس العيادية: الاكتئاب والبارانويا والفصام عند أحجام العينات المختلفة.

بينما أظهرت نتائج الاختبار الاحصائي (M) وجود فرق دال إحصائياً بين معاملي ثبات ألفا للمقياس الطويل والقصير عند مستوى الدلالة (0.05)، بالنسبة لمقياس الهوس الخفيف عند أحجام العينات المختلفة لصالح معامل ثبات المقياس الطويل. ويمكن أن يُعزى ذلك إلى زيادة الفرق في عدد البنود بين الاختبارين الطويل والقصير في هذا المقياس والموضح في الجدول ذي الرقم /13/ حيث تصل نسبة الفرق في عدد البنود إلى (26.08%) أي ما يقارب ربع بنود المقياس.

ويمكن تفسير هذه النتيجة بأن زيادة عدد البنود يؤدي إلى رفع قيمة معامل الثبات، لأنه كلما زاد عدد البنود أصبح المقياس أقدر على كشف الفروق الحقيقية بين الأفراد، حيث أن الاختبارات الطويلة ذات العدد الأكبر من البنود تتيح فرصة كافية أمام المفحوص لتحديد ما يشعر به وما يعبر عن شخصيته بدقة، وهذا بدوره يزيد من كمية التباين الحقيقي على حساب تباين الخطأ، مما يؤدي إلى ارتفاع قيمة معامل الاتساق الداخلي وهذا بدوره يزيد من الثبات، على عكس الاختبارات الأقصر ذات العدد الأقل من البنود، والتي يضطر المفحوص فيها للإجابة عن عدد محدود من البنود، ربما لا يلبي احتياجاته ولا ينسجم مع رأيه، وربما يكون هذا العدد المحدود من البنود غير كاف للتعبير عن شخصيته بدقة، مما يؤدي إلى وقوعه بالأخطاء المقصودة وغير المقصودة للقياس، وبالتالي يزيد تباين الخطأ على التباين الحقيقي، وهذا يؤدي إلى خفض قيمة معامل

الثبات. ذلك لأن زيادة عدد البنود يسمح للأخطاء العشوائية الموجبة والسالبة أن يلغي بعضها الآخر، وبذلك تقترب درجة الفرد الملاحظة من درجته الحقيقية، وبنفس الطريقة تعطي الاختبارات القصيرة أو الفرعية درجات أقل ثباتاً (علام، 2000، ص 181).

تتفق هذه النتيجة مع نتائج دراسة الأحمدي (2008) التي أظهرت وجود حجم تأثير دال لطول الاختبار وحجم العينة على دقة واتساق النتائج، ودراسة عوض (2004) التي بينت وجود فروق ذات دلالة إحصائية عند مستوى الدلالة (0،05) في قيمة معامل الثبات المحسوبة لكل من صورتي الاختبار التحصيلي المُعد لهذه الدراسة لصالح الصورة الأطول، ودراسة هيلمان (2006) التي أظهرت أن أما مقياس (Spos) الأصلي الأطول ذو البعد الواحد فيتميز باتساق داخلي عالٍ أكثر، ويحقق ثباتاً أعلى من النسخة المختصرة لهذا المقياس ، وكذلك تتفق مع نتائج دراسة السكت (2013) التي أظهرت أن الصورة المختصرة (A) لمقياس أيزنك للشخصية أعطت معاملات ثبات أدنى من الصورتين الطويلة (R) والقصيرة (S)، كما أظهرت أن الخصائص السيكومترية لأدوات القياس تتأثر بعدد البنود وحجم العينة وهي ليست مستقلة عنها، وإذا ازداد عدد البنود وحجم العينة تزداد قيم معاملات الصدق والثبات وبفروق دالة إحصائياً عند مستوى الدلالة (0,05).

مناقشة النتائج المتعلقة بالسؤال الرابع: هل توجد فروق ذات دلالة إحصائية بين معاملات ثبات الاتساق الداخلي لدرجات الاختبار باختلاف حجم العينة في كلٍ من المقاييس القصيرة والطويلة؟

أظهرت نتائج الاختبار الاحصائي (M) عدم وجود فروق دالة إحصائياً عند مستوى الدلالة (0.05)، في معاملات ثبات المقاييس الطويلة والقصيرة بين أحجام العينات الأربع المدروسة، وهي (50%)، (100%)، (150%) و(200%) من عدد البنود الكلي للاختبار، فيما عدا وجود فروق دالة إحصائياً عند مستوى الدلالة (0.05) في معاملات ثبات ألفا بالنسبة لمقياس الهوس الطويل بين أحجام العينات الأربع، وهذه الفروق لصالح معامل ارتباط العينة الأصغر (284) والتي تبلغ نسبتها (50%) من عدد بنود الرائز الكلي.

تتعارض هذه النتيجة مع النتائج التي توصلت إليها دراسة السكت (2013) التي بينت وجود فروق دالة إحصائياً في معامل الثبات للبعد (N) عند المقارنة بين حجمي العينة (24 و 48) لصالح العينة الأكبر، في حين تتفق مع النتائج التي توصلت إليها دراسة لوزانو (2008) والتي أوضحت أن أفضل نتائج ثبات لمقاييس التصنيف المُعتمدة من قبل الباحث كانت عند حجم العينة (50)، وبدأت بالتناقص عند أحجام العينات (200)، و(300)، و(500).

وقد تمّ الإشارة في العديد من أدبيات القياس والتقويم إلى أن استخدام عينات ذات حجوم صغيرة في بعض البحوث ربما كان أفضل من استخدام عينات ذات أحجام كبيرة مثل الدراسات التي تعتمد على التحليل النفسي وأدوات القياس الإسقاطي، وأن العينات كبيرة الحجم في الدراسات التجريبية تُعد ضعيفة لأنه سيكون هناك عدد من المتغيرات غير المضبوطة (درويش ورحمة، 2011، ص55).

مناقشة النتائج المتعلقة بالسؤال الخامس: هل توجد فروق ذات دلالة إحصائية بين معاملات الصدق المحكي باختلاف طول الاختبار بدلالة محك قائمة الأعراض؟

أظهرت نتائج الاختبار الإحصائي (V) عدم وجود فروق دالة إحصائياً بين معاملي الصدق المحكي لمقياس الاكتئاب الطويل والقصير عند أحجام العينات الثلاث. في حين بينت وجود فروق دالة إحصائياً بين معاملي الصدق المحكي لمقياس البارانويا الطويل والقصير عند حجم العينة (567) لصالح معامل ارتباط المقياس الطويل. بينما لم تصل هذه الفروق لمستوى الدلالة عند حجمي العينة (425) و(284)، كما بينت النتائج بالنسبة لمقياس الفصام وجود فروق دالة إحصائياً بين معاملي الصدق المحكي لهذا المقياس الطويل والقصير عند حجمي العينة (425) و(567)، وهذه الفروق لصالح معامل ارتباط المقياس الطويل، بينما لم تصل هذه الفروق لمستوى الدلالة عند حجم العينة (284).

ويُلاحظ أنه مع انخفاض حجم العينة تقل الفروق بين معاملات الصدق المحكي، وتغدو غير دالة إحصائياً؛ مما يعني أن حجم العينة وعدد البنود يؤثران معاً في ارتفاع معامل الصدق وانخفاضه،

هنا أيضاً نجد العلاقة الارتباطية بين عدد البنود الكبير وحجم العينة الكافي من حيث علاقتهما برفع قيمة معامل الصدق، حيث أن عدد البنود الأكبر مع حجم العينة الأكبر أنتج علاقة ارتباطية ذات أهمية ودلالة إحصائية.

يمكن تفسير هذه النتيجة أنه كلما زاد عدد بنود الاختبار زاد تمثيله للسمة المقيسة، وبالتالي ارتفع معامل صدقه، مما يرجح قوته وصلاحيته للاستخدام من حيث تقديمه لمعلومات دقيقة عن المفحوصين، أما إذا كان الاختبار قصيراً وتضمن عدداً ضئيلاً من البنود يضعف تمثيله، وبالتالي يضعف صدقه. (ميخائيل، 2008، ص ص 175- 178).

تتفق هذه النتيجة مع نتيجة دراسة السكت (2013) التي بينت أن الصورة الطويلة (R) ذات ال 100 بنداً، وحجم العينة (100) و(200)، أعطت أعلى قيم لمعاملات الصدق بطريقة الارتباطات البينية.

مناقشة النتائج المتعلقة بالسؤال السادس: هل توجد فروق ذات دلالة إحصائية بين معاملات الصدق المحكي لدرجات الاختبار باختلاف حجم العينة بدلالة محك قائمة الأعراض؟

أظهرت نتائج الاختبار الإحصائي (V)عدم وجود فروق دالة إحصائياً في معاملات الصدق المحكي بين كل عينتين من عينات البحث، لكل مقياس من المقاييس العيادية الطويلة والقصيرة، ويمكن أن يعود ذلك وبالرجوع إلى الجدول رقم (14) إلى اقتراب قيم الإحصاءات الوصفية لهذه العينات والتي تتضمن قيم المتوسط والانحراف المعياري والتباين والخطأ المعياري وأيضاً قيم الالتواء والتفرطح، كما يمكن أن يعود ذلك لعدم وجود فروقات في التجانس بين هذه العينات ذات الأحجام المختلفة.

حيث قامت الباحثة بحساب التجانس بين هذه العينات عند كل مقياس من المقاييس العيادية، باستخدام اختبار Leven للتجانس، وقد بينت النتائج أن قيمة ال Sig لهذا الاختبار بلغت (0.736) بالنسبة لمقياس الاكتئاب، و (0.787) بالنسبة لمقياس البارانويا، و(0.400) بالنسبة لمقياس الفصام، وهذه القيم أكبر من (0.05) وبالتالي لاتوجد فروق دالة إحصائياً للتجانس بين هذه العينات بالنسبة للمقاييس العيادية الأربعة، مما يدل على تجانس هذه العينات.

مناقشة النتائج المتعلقة بالسؤالين السابع والثامن:

السؤال السابع: هل توجد فروق ذات دلالة إحصائية بين معاملات صدق الاختبار باستخدام طريقة المجموعات المتضادة بين عينتي/الأسوياء، والمضطربين نفسياً/ باختلاف طول الاختبار؟

السؤال الثامن: هل توجد فروق ذات دلالة إحصائية بين متوسطات الدرجات على المقاييس العيادية استناداً لمتغير المرحلة الدراسية /مرحلة التعليم الجامعي، ومرحلة التعليم ما بعد الجامعي – الماجستير/ باختلاف طول الاختبار؟

أوضحت نتائج اختبار تحليل التباين والمقارنات البعدية وجود فروق دالة إحصائياً بين متوسط درجات عينة المضطربين نفسياً على الاختبارين الطويل والقصير، لصالح متوسط عينة المضطربين نفسياً على الاختبار الطويل، وكذلك وجود فروق دالة إحصائياً بين متوسط درجات عينة الأسوياء على الاختبارين الطويل والقصير، لصالح متوسط عينة الأسوياء على الاختبار الطويل، ما يعني أن هناك أثراً دالاً لطول الاختبار في متوسط درجات العينتين، وبالتالي في معاملات صدق الاختبار.

كما أوضحت نتائج اختبار تحليل التباين والمقارنات البعدية وجود فروق دالة إحصائياً بين متوسط درجات عينة طلبة الماجستير على الاختبارين الطويل والقصير، لصالح متوسط عينة طلبة الماجستير على الاختبار الطويل، وكذلك وجود فروق دالة إحصائياً بين متوسط درجات عينة

الطلبة الجامعيين على الاختبارين الطويل والقصير، لصالح متوسط عينة الطلبة الجامعيين على الاختبار الطويل، ما يعني أن هناك أثراً دالاً لطول الاختبار على متوسطات درجات العينتين.

وأن المقاييس الطويلة كانت أقدر على إظهار الفروق في الدرجات بين العينات من المقاييس القصيرة ، حيث تعزو الباحثة ذلك إلى أن زيادة عدد البنود يتيح فرصة كافية أمام المفحوص لتحديد ما يشعر به ويختار بالتالي ما يعبر عن شخصيته بدقة أعلى، وهذا بدوره يزيد من مستوى التباين الحقيقي بين الأفراد، على عكس المقاييس القصيرة التي لها عدد بنود أقل، حيث يكون المفحوص مضطراً للإجابة عن عدد بنود أقل وربما لا يلبي احتياجاته، وبالتالي يكون للمقياس قدرة أقل على إظهار التباين في السمة موضع القياس.

تتفق هذه النتيجة مع نتيجة دراسة السكت (2013) التي بينت أن الصورة الطويلة (R) ذات ال 100 بنداً وحجم العينة (100) و(200) أعطت أعلى قيم لمعاملات الصدق.

مناقشة النتائج المتعلقة بالسؤال التاسع: هل توجد فروق ذات دلالة إحصائية بين درجات القطع الناتجة عن الاختبار باختلاف طول الاختبار؟

أوضحت النتائج هبوط قيم درجات القطع في الاختبارات القصيرة بمجملها عن مثيلاتها في الاختبارات الطويلة، وفي معظم الأحيان كان المدى بين درجات القطع العليا والدنيا في الاختبارات الطويلة أوسع من نظيره في الاختبارات القصيرة، وبالتالي يمكن القول بأن درجات القطع في الاختبارات الطويلة يمكن أن تكون أكثر دقة من حيث قدرتها على إعطاء مؤشرات إكلينيكية من مثيلاتها في الاختبارات القصيرة.

يمكن أن يعود ذلك وكما تمّ الإشارة آنفاً إلى أن عدد البنود الكافي يتيح الفرصة أمام المفحوص لتحديد ما يعبر عن شخصيته بدقة أعلى، على عكس المقاييس القصيرة التي لها عدد بنود أقل حيث يكون المفحوص مضطراً للإجابة عن عدد بنود أقل وربما لا يلبي احتياجاته، وبالتالي يكون للمقاييس الأقصر قدرة أقل على إعطاء مؤشرات إكلينيكية حول السمة موضع القياس.

كما أوضحت نتائج الاختبار الإحصائي (t) ستيودنت أنَّه على الرغم من هبوط قيم درجات القطع في الاختبارات القصيرة عن مثيلاتها في الاختبارات الطويلة، إلا أن قيمة هذه الفروق لم تصل إلى مستوى الدلالة الإحصائية. تتفق هذه النتيجة مع نتيجة دراسة التميمي (1999) التي بينت تساوي أو تقارب نسب درجات قطع الاختبار الطويل والمتوسط والقصير وفق الطرق المختلفة لحساب درجات القطع باستثناء طريقة المجموعات المتضادة، حيث تساوت نسب درجات قطع الاختبار الطويل والمتوسط والقصير وفق طريقتي ندلسكاي والمجموعات المحكية وبلغت هذه النسبة (57.1)، في حين تقاربت هذه النسب ولم تصل إلى مستوى الدلالة وفق طريقة أنجوف مما يدل على عدم وجود أثر لمتغير طول الاختبار في تحديد درجات القطع، أي لا يترتب على زيادة

عدد المفردات تدني أو ارتفاع في نسب درجات القطع، كذلك تتفق مع نتيجة دراسة ذيب (2013) التي أظهرت تساوي نسب درجات قطع الاختبار الطويل والمتوسط والقصير وفق طريقتي نيدلسكاي والمجموعات المحكية وبلغت هذه النسبة (64.3)، كما أظهرت أن الاختبار الطويل حقق أعلى نسبة لدرجة القطع وفقاً لطريقة المجموعات المتضادة، حيث بلغت هذه النسبة (72.6).

مناقشة النتائج المتعلقة بالسؤالين العاشر والحادي عشر:

السؤال العاشر: ما أثر التفاعل بين طول الاختبار وخصائص العينة في الخصائص السيكومترية للاختبار؟

السؤال الحادي عشر: ما أثر التفاعل بين طول الاختبار وخصائص العينة على درجات القطع؟

تمّت مناقشة نتائجهما من خلال مناقشة نتائج الأسئلة السابقة.

<u>عاشراً: مقترحات البحث:</u>

اعتماداً على نتائج البحث توصلت الباحثة إلى المقترحات الآتية:

- العمل على إجراء المزيد من الأبحاث والدراسات التي تتناول القوانين الخاصة بمنهج ما وراء التحليل الإحصائي وتتبع مراحل إجرائه.
- العودة إلى البروتوكولات المنشورة حول أبحاث ما وراء التحليل الإحصائي، وإجراء بحوث ودراسات متعلقة بها بشكل أعمق.
- العمل على إجراء المزيد من الأبحاث والدراسات التي تتناول القانون الإحصائي لمعامل ألفا كرونباخ، والعوامل المؤثرة فيه.
- إجراء دراسات وبحوث حول إمكانية إيجاد نماذج مختصرة من الاختبارات والمقاييس التربوية بشكل عام، والنفسية على وجه الخصوص، بحيث تحافظ على الخصائص السيكومترية الجيدة للاختبار، تعتمد على تقنيات التحليل العاملي ومنهج ماوراء التحليل الإحصائي.
- إجراء المزيد من الدراسات على المقاييس المشتقة من رائز MMPI-2 كاختباري MMPI-RFو MMPI-A والتي تختلف بنودها بشكل كامل عن اختبار MMPI-2.
- الاهتمام بالأساليب الإحصائية والرياضية لتقدير حجم العينة.

ملخص البحث باللغة العربية:

هدف البحث لمعرفة أثر كل من طول الاختبار وحجم العينة على الخصائص السيكومترية ودرجات القطع لاختبار الشخصية المتعدد الأوجه مينيسوتا 2 – MMPI.

ولجمع البيانات وتحقيق أهداف البحث قامت الباحثة بتطبيق استخبار الشخصية مينيسوتا على مجموعة عينات على النحو الآتي:

- **عينة طلبة التعليم الجامعي:** تكونت الدراسة من طلبة السنة الرابعة في جامعة دمشق بأقسامها الطبية، والتطبيقية، والتربوية، والعلوم الإنسانية، والبالغ عددهم (1134) طالب وطالبة من الأقسام التي تم السحب منها، وفقاً لإحصائيات مديرية شؤون الطلاب في وزارة التعليم العالي للعام الدراسي 2012/ 2013 م، حيث جرى سحب هذه العينة بطريقة العينة العشوائية، تراوحت نسب السحب بين 28.57% و 10.34%.
- **عينة طلبة التعليم ما بعد الجامعي /طلبة الماجستير/:** بلغ عدد أفرادها (206) طالب وطالبة.
- **عينة المضطربين نفسياً:** تم سحبها من مشفى ابن رشد للأمراض النفسية ومركز الشهيد زهير حبي / العيادة النفسية /.

وكانت أحجام العينات المدروس أثرها على الخصائص السيكومترية، (50%) و(75%) و(100%) و(150%) و(200%) من عدد بنود الاستخبار، اعتمدت الباحثة المنهج الوصفي التحليلي لملاءمته طبيعة وموضوع البحث.

وقد توصل البحث إلى النتائج التالية:

- وجود علاقة ارتباطية بين عدد البنود الكبير وحجم العينة الكافي، من حيث تأثيرهما على رفع قيمة معامل ارتباط بيرسون حيث أن عدد البنود الأكبر مع حجم العينة الأكبر أنتج علاقة ارتباطية ذات أهمية ودلالة إحصائية.
- عدم وجود فروق دالة إحصائياً في معاملات الارتباط بين كل حجمي عينة من عينات البحث لكل مقياس من المقاييس العيادية الطويلة والقصيرة.
- وجود فروق دالة إحصائياً بين معاملات ثبات ألفا للمقياس الطويل ومعاملات ثبات ألفا للمقياس القصير عند مستوى الدلالة (0.05)، بالنسبة لمقياس الهوس الخفيف عند أحجام العينات الأربع؛ وهذا الفرق لصالح معاملات ثبات المقياس الطويل.
- عدم وجود فروق دالة إحصائياً بين معاملات ثبات ألفا كرونباخ بالنسبة للمقاييس القصيرة بين أحجام العينات الأربع، للمقاييس العيادية كافة.

- عدم وجود فروق دالة إحصائياً بين معاملات ثبات ألفا كرونباخ عند مستوى الدلالة (0.05) للمقاييس الثلاثة الاكتئاب والبارانويا والفصام الطويلة، بين أحجام العينات الأربع.
- وجود فروق دالة إحصائياً بين معاملات ثبات ألفا كرونباخ للاختبار الطويل لمقياس الهوس الخفيف بين أحجام العينات الأربع، وهذه الفروق لصالح معامل ارتباط العينة الأصغر.
- أظهرت نتائج البحث هبوط قيم معاملات الارتباط المحكي في المقاييس القصيرة عن مثيلاتها في المقاييس الطويلة.
- وجود أثر دال لطول الاختبار في متوسط درجات العينات المختلفة.
- وعلى الرغم من وجود فروقات بين درجات القطع للاختبارات الطويلة والقصيرة، إلا أن هذه الفروق لم تصل إلى مستوى الدلالة الإحصائية.

وتوصل البحث اعتماداً على نتائجه إلى المقترحات الآتية:

- العمل على إجراء المزيد من الأبحاث والدراسات التي تتناول القوانين الخاصة بمنهج ما وراء التحليل الإحصائي وتتبع مراحل إجرائه.
- التركيز على استخدام القوانين الاحصائية والرياضية التي يستند إليها هذا المنهج.
- العودة إلى البروتوكولات المنشورة حول أبحاث منهج ما وراء التحليل الإحصائي، وإجراء بحوث ودراسات متعلقة بها بشكل أعمق.
- العمل على إجراء المزيد من الأبحاث والدراسات التي تتناول القانون الإحصائي لمعامل ألفا كرونباخ، والعوامل المؤثرة فيه.
- إجراء دراسات وبحوث حول إمكانية إيجاد نماذج مختصرة من الاختبارات والمقاييس التربوية بشكل عام، والنفسية على وجه الخصوص، بحيث تحافظ على الخصائص السيكومترية الجيدة للاختبار، بزمن أداء وجهد وتكاليف أقل.
- الاهتمام بالأساليب الإحصائية والرياضية لتقدير حجم العينة.

المراجع

أولاً: المراجع العربية

- الأحمدي، شرف بنت حامد عبد الله. (2008). *أثر اختلاف طول الاختبار وحجم العينة على دقة واتساق تقدير بارامترات المفردة والفرد في نظرية الاستجابة المفردة*، رسالة ماجستير غير منشورة، السعودية، جامعة أم القرى.
- الأنصاري. بدر محمد. (2000). *قياس الشخصية*. كلية العلوم الاجتماعية: جامعة الكويت.
- أبو حطب، فؤاد وعثمان، أحمد. (2003). *التقويم النفسي*. (ط.3). القاهرة: مكتبة الأنجلو المصرية.
- أبو علام، رجاء محمود. (2005). *تقويم التعلم*. (ط.1). عمّان: دار المسيرة للنشر والتوزيع.
- أبو النيل، محمود السيد. (1987). *الإحصاء النفسي والاجتماعي والتربوي*. جامعة عين شمس: كلية الآداب.
- أبو لبدة، سبع. (1985). مبادئ القياس النفسي والتقييم التربوي. (ط.3). الجامعة الأردنية: كلية التربية.
- أحمد الشريم، يوسف سوالمة.(2006). تحديد درجة القطع لاختبار محكي المرجع في الرياضيات باستخدام نموذجي أنجوف وندلسكي: دراسة مقارنة بمعرفة صعوبة الفقرات وعدم معرفتها. *المجلة الأردنية في العلوم التربوية*، جامعة اليرموك.
- أحمد عودة، أحمد الشريم. (2010). تطوير أسلوب قائم على الأهداف السلوكية لتحديد درجة القطع: دراسة مقارنة مع أسلوب أنجوف. *المجلة الأردنية في العلوم التربوية*، جامعة اليرموك.
- إيمان، عبد الكريم ذيب. (2013). أثر طول الاختبار على تحديد درجة القطع لاختبار تحصيلي محكي المرجع غي مادة علم نفس الشخصية، الجامعة العراقية، كلية التربية.
- بالانت، جولي. (2006). التحليل الإحصائي باستخدام برنامج ال SPSS. عمان: دار الفاروق. (ترجمة خالد العامري).
- البقاعي، هيفاء. (2007). *تعيير اختبار روتر لتكملة الجمل الناقصة وفاعلية استخدامه في تشخيص بعض الاضطرابات النفسي*، رسالة ماجستير غير منشورة، جامعة دمشق.

- البهي السيد، فؤاد. (2005). *علم النفس الإحصائي وقياس العقل البشري*. (ط.3). القاهرة: دار الفكر.
- التميمي، خالد بن حسن. (1999). *أثر كل من نوع المحكم وطول الاختبار على تحديد درجة القطع لاختبار محكي المرجع يقيس الكفايات الرياضية في العمليات الحسابية على الاعداد بالصف السادس الابتدائي بمدينة جدة*، رسالة ماجستير غير منشورة، السعودية، جامعة أم القرى.
- الجادري، عدنان حسين. (2007). *الإحصاء الوصفي في العلوم التربوية*. (ط.2). الأردن: دار المسيرة للنشر والتوزيع والطباعة.
- حريري، عبد الرحمن. (2009). *تحليل البيانات والاختبارات الإحصائية*.
- خيري، السيد محمد. (1997). *الإحصاء النفسي*. مصر: دار الفكر العربي.
- خليل، شرف الدين. (2015). *الإحصاء الوصفي*. شبكة الأبحاث والدراسات الاقتصادية.
- درويش، رمضان ورحمة، عزيزة. (2012). *الإحصاء الوصفي*. دمشق: منشورات جامعة دمشق.
- رحمة، عزيزة. (1999). *التحليل الإحصائي لصدق وثبات رائز وكسلر لذكاء الراشدين*. رسالة ماجستير غير منشورة، دمشق.
- رزوق وسناد، إبراهيم وجلال. (2008). *مناهج البحث*. جامعة دمشق: كلية التربية.
- رضوان، رانيا. (2011). *أثر طرائق تقدير الدرجات في الخصائص السيكومترية لاختبار تحصيلي ذي الاختيار من متعدد في مادة الرياضيات للصف الرابع من التعليم الأساسي*. رسالة ماجستير غير منشورة، جامعة دمشق.
- الروسان، فاروق. (1996). *أساليب القياس والتشخيص في التربية الخاصة*. (ط.1). الأردن: عمّان.
- السكت، فاديا. (2013). *أثر حجم العينة وعدد البنود على الخصائص السيكومترية لمقياس آيزنك للشخصية (EPQ-R) دراسة لدى عينة من طلبة جامعة دمشق* ، رسالة ماجستير غير منشورة، كلية التربية، جامعة دمشق.
- السيد عبده، عبد الهادي. (2002). *القياس والاختبارات النفسية*. (ط.1). دار الفكر العربي.
- الصياد، جلال ،وربيع، عبد الحميد.(1983). مبادئ الطرق الاحصائية.(ط.1). المملكة العربية السعودية: جدة.
- الطيب، ملاذ. (2011). *تعيير رائز مينيسوتا متعدد الأوجه للشخصية الخاص بالمراهقين*. رسالة ماجستير غير منشورة، جامعة دمشق.

- عبد الخالق، أحمد محمد. (2005). *استخبارات الشخصية*. (ط.3). جامعتا الإسكندرية والكويت: دار المعرفة الجامعية.
- عبد الرحمن، سعد. (1998). *القياس النفسي، النظرية والتطبيق*. (ط.3). القاهرة: دار الفكر العربي.
- *عبد الرحمن، محمد السيد*. (2000). *علم الأمراض النفسية والعقلية. (الأسباب- الأعراض- التشخيص- العلاج)*. الجزء الأول. القاهرة: دار قباء للطباعة.
- عبيدات، ذوقان وآخرون. (1993). *البحث العلمي، مفهومه، أدواته وأساليبه*. الرياض: دار أسامة للنشر.
- عز، إيمان (قيد النشر): تقنين رائز مينيسوتا MMPI-2، *مجلة اتحاد عمداء كلية التربية*، جامعة دمشق: كلية التربية.
- عز، إيمان (1990): *رائز بيرنروتر للشخصية، دراسة الرائز وتعييره في القطر العربي السوري. رسالة ماجستير غير منشورة*، كلية التربية، جامعة دمشق.
- العزاوي ، رحيم يونس. (2008). *المنهل في العلوم التربوية*. (ط.1). المملكة الأردنية الهاشمية: دار دجلة.
- علام ، صلاح الدين. (2002). *القياس والتقويم التربوي والنفسي أساسياته وتطبيقاته وتوجهاته المعاصرة*. القاهرة: دار الفكر العربي.
- علام، صلاح الدين. (2005). *الأساليب الإحصائية الاستدلالية في تحليل بيانات البحوث النفسية والتربوية والاجتماعية – البارامترية واللابارامترية-*. القاهرة: دار الفكر العربي.
- علام، صلاح الدين. (2006). *الاختبارات والمقاييس التربوية والنفسية*. عمّان: دار الفكر.
- علام، صلاح الدين. (1991). دراسة مقارنة لبعض طرق تحديد مستويات الأداء في اختبار محكي المرجع. *المجلة المصرية للدراسات النفسية*. العدد 1.
- العلي، ديالا صالح. (2011). *تدريج اختبار مينيسوتا للشخصية (MMPI-2) باستخدام نموذج راش. رسالة ماجستير غير منشورة*، كلية التربية، جامعة دمشق.
- عنبر، أحمد (1990): *تعيير رائز مينيسوتا متعدد الأوجه للشخصية ر. م. م. ش*، رسالة دكتوراه غير منشورة، كلية التربية، جامعة دمشق.
- عوض، رضا. (2004). *دراسة أثر عدد المفردات وموقع المفردة من الاختبار وحجم العينة وعدد البدائل على قيمة معامل ثبات المقياس*. رسالة ماجستير غير منشورة، معهد الدراسات والبحوث التربوية، مصر، جامعة القاهرة.
- عيسوي، عبد الرحمن. (2000). *الإحصاء السيكولوجي والتطبيقي*. جامعة الإسكندرية: كلية الآداب.

- الغريب، رمزية. (1996). *التقويم والقياس النفسي والتربوي*. القاهرة: مكتبة الأنجلو المصرية.
- فرج، صفوت. (2007). *القياس النفسي*. (ط.6). القاهرة: مكتبة الأنجلو المصرية.
- فهمي، محمد شامل. (2005). *الإحصاء بلا معاناة، المفاهيم مع التطبيقات باستخدام برنامج الSPSS*. الجزء الأول. الرياض: مكتبة فهد الوطنية.
- القصاص، مهدي محمد. (2007). *مبادئ الإحصاء والقياس الاجتماعي*. جامعة المنصورة: كلية الآداب.
- الكيلاني، عبد الله زيد. (1989). *القياس والتقويم في علم النفس والتربية*. (ط.4). مركز الكتب الاردني.
- محمود، جودت شاكر. (2007). *البحث العلمي في العلوم السلوكية*. مصر: مكتبة الأنجلو المصرية.
- مجيد، سوسن شاكر. (2007).*أسس بناء الاختبارات والمقاييس النفسية والتربوية*. (ط.1). عمّان: دار ديبونو للنشر والتوزيع.
- مجيد، سوسن شاكر. (2008). *اضطرابات الشخصية، أنماطها، قياسها*. (ط.1). عمان: دار صفاء للنشر والتوزيع.
- *مراد، صلاح وسليمان، أمين. (2002). الاختبارات والمقاييس في العلوم النفسية والتربوية، خطوات إعدادها وخصائصها*. القاهرة: دار الكتاب الحديث.
- ملحم، سامي محمد. (2005). *القياس والتقويم في التربية وعلم النفس*. (ط.3). عمّان: دار المسيرة للنشر والتوزيع.
- ملحم، سامي. (2001). *القياس والتقويم في التربية وعلم النفس*. (ط.3). عمّان: دار المسيرة للنشر والتوزيع.
- المنيزل، عبدالله وغرايبة، عايش. (2006). *الإحصاء التربوي*. (ط.1). الأردن: دار المسيرة.
- ميخائيل، امطانيوس. (1997). *القياس والتقويم في التربية الحديثة*. منشورات جامعة دمشق: كلية التربية.
- ميخائيل، امطانيوس. (2006). *القياس النفسي (الجزء الأول)*. منشورات جامعة دمشق: كلية التربية.
- ميخائيل، امطانيوس .(2008). *القياس النفسي (الجزء الثاني)*. منشورات جامعة دمشق: كلية التربية.

-النبهان، موسى. (2004). *أساسيات القياس في العلوم السلوكية*. (ط.1). الأردن: دار الشروق للنشر والتوزيع.

ثانياً: المراجع الأجنبية

- Socha, A, Barclay, R. (2005). *Cut Score Analysis and Validity for Writing Prompts*. The College of New Jersey.

-Bunny, P. (1990). *Application of items response theory to criterion – referenced measurement: An investigation of the effects of model choice*. Educational and Psychological Measurement.

- Crask, R. (1999). *The perfect design optimization between reliability, validity , redundancy in scale items and response rates*. University of Georgia.

- Drost , A. (2011). *Validity and Reliability in Social Science Research*. USA: California State University.

- Dean, L. (2004). *Psychometric Properties and Predictive Validity of the Mt Scale of the MMPI-2*. Louisiana State: Eastern Michigan University.

- Evan, P. (2010). *MMPI to MMPI-2 to MMPI-2 FR: An examination of the psychometric properties in a sample of patients with medically unexplained physical symptoms*. Fielding Graduate University.

- Garson, G. (2009). *Reliability Analysis*. USA: State notes from- north Carolina State University.

- Hakstian, A, Whalen, R . (1976). *AK- Sample significance test for independent alpha coefficients*. Psychometrika, 219 – 231.

- Halpin, g. (1983).*minimum competency standard set by three divergent groups of raters using three judgmental procedures: implications for validity*. Educational and Psychological Measurement.

- Hays, W. (1977). *Statistics for the Social Sciences (*2^{nd} *edition)*. New York: Holt, Rinehart, and Winston, Inc

- Hellman, CM. (2006). *A reliability generalization study on the survey of perceived organizational support: The effects of mean age and number of items on score reliability*. Oklahoma: university of Oklahoma.

- Glenn, D. (2009). *Determining Sample Size*. USA: University of Florida.

- Horrocks, J. (1965). *Assessment Of Behavior*. Ohio: The Ohio State University.

- Lee, J, Cronbach, A, & Richard, J. (2004). *My current thoughts on coefficient Alpha and successor procedures*. Educational and Psychological Measurement. Stanford University, 391- 418

- Leonard, S, Feldt, D, Woodruff, J. (1987). *Statistical Inference for Coefficient Alpha. Applied Psychological Measurement.* The University of Lowa, 93- 103.

- Lozano, M. (2008). *Effect of the number of response and sample size on the reliability and validity of rating scales*. Spain: University of Jaen.

- Meca, R, Martinez, S. (2010). *Meta- analysis in Psychological Research*. International Journal of Psychological Research. Spain:

University of Murcia.

- Phelan, D, Wren, A. (2005). *Exploring Reliability In Academic Assessment*. USA.

- Scott, W. (2009). *A simulation study provided sample size guidance for differential items functioning (DIF) studies using short scales*. University of Aberdeen.

- Slutske, S. (1989). *A Real- Data Simulation of Computerized Adaptive Administration of MMPI*. University of Minnesota.

- Timothy, M. (2008),*The MMPI-2 Clinical Scales and Restructured Clinical (RC) SCALES, Comparative Psychometric Properties and Relative Diagnostic Efficiency in Young Adult*. A department of psychology. Niagara University.

- Woher, D. (1991). *An empirical comparison of cut off score methods for content related and criterion related Validity settings*. Educational and Psychological Measurement.

ثالثاً: مراجع الانترنت

- www.http://occmed.oxfordjournals.org/cgi/content/full/59

- telc.tanta.edu.eg/hosting/pro16/contain/l3.htm

الملحق رقم 1: تحديد حجم العينة (من 500- >100000)

حجم العينة عند مستويات الدقة ±3%، ±5%، ±7% و ±10% حيث مستوى الثقة 95% و P = 5				
حجم المجتمع	حجم العينة			
	±3%	±5%	±7%	±10%
500	a	222	145	83
600	a	240	152	86
700	a	255	158	88
800	a	267	163	89
900	a	277	166	90
1000	a	286	169	91
2000	714	333	185	95
3000	811	353	191	97
4000	870	364	194	98
5000	909	370	196	98
6000	938	375	197	98
7000	959	378	198	99
8000	967	381	199	99
9000	989	383	200	99
10000	1000	385	200	99
15000	1034	390	201	99
20000	1053	392	204	100
25000	1064	394	204	100
50000	1087	397	204	100
100000	1099	398	204	100
100000<	1111	400	204	100

الملحق رقم 2: تحديد حجم العينة (من 100– 450)

حجم العينة عند مستويات الدقة ±3%، ±5%، ±7% و ±10% حيث مستوى الثقة 95% و P = 5			
حجم المجتمع	حجم العينة		
	±5%	±7%	±10%
100	81	67	51
125	96	78	56
150	110	86	61
175	122	94	64
200	134	101	67
225	144	107	70
250	154	112	72
275	163	117	74
300	172	121	76
325	180	125	77
350	187	129	78
375	194	132	80
400	201	135	81
425	207	138	82
450	212	140	82

The research summary

The present study aimed to determine the effect of each of the test length and the size of the sample on the psychometric characteristics and cut of degrees to Multi-Faceted Minnesota Personality Inventory MMPI -2.

To collect data and achieve the goals of research, the researcher applying the scale on samples group as follows:

- **a sample of university education students**: study consisted of fourth year students at Damascus University Departments including: medical sections, applied sections, education and humanities sections, total number of sample is (1134) students, according to the statistics of the Directorate of Student in the Ministry of Higher Education 2012/2013, where This sample was withdrawn with random sampling method, drag ratios ranged between (28.57%) and (10.34%) of each section.
- **sample of education students after the university /Masters students/:** the number of members reached to 206 students.
- **sample of psychologists patients**: was withdrawn from the hospital Ibn Rushd psychiatric, and center of /Zuhair Hubbi/ Psychiatric Clinic.

The sizes of the samples studied their impact on the psychometric properties were: (50%) and (75%) and (100%) and (150%) and (200%) of the number of scale items, the researcher adopted the descriptive analytical approach to the nature and relevance of the research topic.

The research has found the following results:

- The existence of correlation between the large number of items and the enough sample size, in terms of their impact on raising the value of the Pearson correlation coefficient, as the large number of items with the enough sample size produced a correlation with statistical significance.

- There is no statistically significant differences in the correlation coefficients between every two sample size from research samples, to each clinical scale long and short.

- There are significant differences between the alpha coefficients reliability for long scale and the alpha coefficients reliability short scale at the significance level (0.05), for Hypomania scale to the four sample sizes; and this difference in favor of the long-scale coefficient.

- There is no statistically significant differences between Cronbach's alpha coefficients for short-scales between the four sample sizes, in all clinical measurements.

- There is no statistically significant differences between Cronbach Alpha coefficients at the significance level (0.05) for the three measures of depression, paranoia and schizophrenia long scales, at the four sample sizes.

- There are significant differences between Cronbach's alpha coefficient for Hypomania long scale among the four sample sizes, and these differences in favor of the smaller sample correlation coefficient.

- Finds showed reduction in the values of correlation coefficients criterion for the short scales from those in long scales.

- The existence of significant effect to the test length on the average scores of different samples.

- In spite of the existence of differences between the cut of scores of long scales and the cut of scores of short scales, but these differences did not reach the level of statistical significance.

The research found depending on its findings to the following proposals

- Work on doing further research and studies dealing with special laws approach beyond the meta- analyzes approach and following its applications.

- Focus on the use of statistical and mathematical laws underlying this approach

- Go to the published protocols about meta- analyzes approach research and conducting studies related to them deeper.

- Work on further research and studies dealing with the statistical law Cronbach's alpha coefficient, and the factors affecting on it.

- Conduct studies and research on the possibility of a brief educational models of tests and scales in general, and in particular psychological, so that keeps the good psychometric properties of the test, a time of performance, effort and costs less

More Interesting in statistical and mathematician methods to estimate . the sample size.

Damascus University
Faculty of Education
Educational and Psychological Evaluation and Measuremen

The impact of interaction between the test length and the sample properties on the psychometric characteristics and cut- off scores to Multiphasic Personality Inventory Minnesota MMPI - 2

A research submitted to attain Master Degree in Educational and Psychological Evaluation and Measurement Department

Presented By

Sumaiah Houssin Swaid

Supervised By

Dr.Eman Ezz

Assistant professor in Evaluation and Measurement Department

1436 – 1437 H

2015 – 2016 C