صدق المحتوى فى البحوث التربوية : الواقع والتطوير

د/ محمد ابراهيم محمد محمد

مدرس علم النفس التربوي – كلية التربية – جامعة المنيا

ملخص البحث :

يهدف البحث الحالى الى د راسة واقع ممارسات صدق المحتوى فى البحوث التربوية من خلال تحليل البحوث التربوية من ناحية واراء الخبراء من ناحية اخرى والمقارنة بين طريقتى مؤشر صدق المحتوى ونسب الاتفاق بالاضافة لتطوير جداول مؤشر صدق المحتوى والتعرف على أثر حجم المحكمين على قبول أو رفض المفردات بطريقتى مؤشر صدق المحتوى ونسب الاتفاق وللتحقق من ذلك قام الباحث بتحليل عدد (70) رسالة دكتوراه تم إجازتها بكلية التربية جامعة المنيا وتطبيق أداة البحث على عدد (30) محكما واوضحت النتائج وجود مشاكل فى ممارسات صدق المحتوى نتيجة لضعف المام الباحثين بمفهوم صدق المحتوى وعدم وجود معايير واضحة لتحديد المحكم ومعيار الحكم على صدق الادوات والاقتصار على طريقة نسب الاتفاق رغم تأثرها بعامل صدفة الاتفاق وتضخم تقديرتها واوضحت النتائج اهمية معامل صدق المحتوى كطريقة كمية تعبر عن تمثيل الاختبار للمحتوى حيث تقدم مؤشر لصدق محتوى المفردات ومؤشر لصدق محتوى الاختبار ككل عند مستوى دلالة (0.05) علاوة على أنها تستخدم متوسط المفردة لتقليل تأثير حجم عينة التحكيم كما توصل الباحث لمعادلة مبسطة لحساب مؤشر صدق المحتوى وتطوير جداول لطريقة نسب الاتفاق عند مستوى دلالة (0.05) .

صدق المحتوى فى البحوث التربوية : الواقع والتطوير

د/ محمد ابراهيم محمد محمد

مدرس علم النفس التربوي – كلية التربية – جامعة المنيا

مقدمة :

يعد صدق القياسات موضوعا مثيرا للاهتمام في مجال القياس النفسي والتربوي، حيث يهدف إلى التحقق من قياس الأدوات لما وضعت لقياسه من خلال تركيزه على تفسير البيانات والنتائج المستخرجة من تجريب أداة القياس، وقد تضمن جوانب التحقق من مناسبة ومماثلة المفردات والأداة ككل للبناء المستهدف وهو ما يطلق عليه صدق المحتوى، والتحقق من قياس الأداة للظاهرة نفسها من خلال مقارنتها باختبار قديم كمحك وهو ما أطلق عليه الصدق المرتبط بالمحك، والتحقق من الصدق التجريبي للمفهوم والبناء المستهدف وهو ما يطلق عليه صدق البناء، وتلك الأنماط الثلاثة التي اعتمدت عليها نظرية الصدق التقليدية لدراسة الصدق لأدوات القياس .

ويهدف صدق المحتوى إلى التحقق من أربعة جوانب أساسية لأدوات القياس هي : تعريف المجال ومناسبة المجال وتمثيل المجال وملاءمة إجراءات بناء المجال وهذه الجوانب الأربعة تمثل جودة الاختبار، حيث يشمل صدق المحتوى التحقق من عناصر الاختبار قبل التطبيق وهو ما يحدد جودة الاختبار من حيث البناء اللغوي ومناسبة المفردات للمرحلة العمرية والمرحلة الدراسية والأهداف المستهدفة وتمثيلها للبناء المستهدف والتعريفات الإجرائية فهو يشمل جميع عناصر أداة القياس مما يساعد على تجويد الاختبار وإخراجه في صورة مقبولة علميا ولغويا وفنيا .

ومن ثم فان صدق المحتوى يقدم أدلة حول صدق البناء لأداة القياس (Anastasi,1988) فصدق البناء هو الدرجة التي تقيس بها أداة القياس البناء المستهدف ، فطبقا لنظرية القياس التقليدية فإن صدق إجراءات التحقق من صدق البناء تتضمن أربعة مراحل هي :

1- التعريف والوصف للبناء المستهدف وكيفيه تناوله.

2- العمليات والإجراءات لترجمة البناء إلى سلوكيات فعلية.

3- تقويم الصدق التقاربي والصدق التمييزي

4- تقويم الصدق المنطقي ويتضح من ذلك أن صدق المحتوى يستخدم كمفتاح لقياس الأبنية، فصدق المحتوى يمثل الخطوتين الأوليتين لإجراءات التحقق من صدق البناء وهما التعريف والعمليات الإجرائية للبناء .(King& King ,1990)

ويؤثر صدق المحتوى في بنية العوامل الكامنة في أدوات القياس ومفرداتها التي انتقاؤها لتمثل أوجه السمة المقاسة أو بنية العامل الكامن لأداة القياس حيث تفترض أن المفردة تقيس

الخاصية نفسها أو الوجه نفسه الذي حصل على القيمة ذات الدلالة، فان أداة القياس التي تعانى من عدم كفاية صدق محتواها سوف تفشل في تحقيق افتراض العوامل الكامنة لأداة القياس حيث أن المفردات لن تقع على أبعادها التي وضعت من اجل قياسها وبالتالي فالمقياس لن يحقق الأبعاد التي وضع لقياسها. (Haynes, Richard, Kubany,1995)

كما أن صدق المحتوى له علاقة بالصدق التنبؤى حيث أن الاختبار القادر على التنبؤ بالمستقبل للظاهره المقاسه من المفترض أن يكون بناءه صادق، ولا يتحقق ذلك إلا من خلال دعم الصدق التنبؤى بأدلة وافية مثل صدق المحتوى الذي يساعد على تقوية وتدعيم الصدق التنبؤى فإذا أشارت نتائج التحكيم أن الاختبار غير ممثل للمحتوى فإنه لا توجد علاقة بين الاختبار ومحكه فالضعف في تمثيل المحتوى يؤدى إلى الضعف في نواتج الاختبار كما أن ضعف تمثيل المفردات للمحتوى يؤدى إلى ضعف تمثيل الاختبار ككل .(Shepard,1996,Sireci,1998)

كما إن صدق المحتوى ذو طبيعة دينامية لذا تطوير تقييم أداة القياس لابد وان يتسق في السياق مع النظريات الحديثة حول البناء المستهدف وذلك لان التعريف الإجرائي والمجال والأوجه للأبنية المتعددة تتغير بتغير الزمن فان مناسبة وتمثيل عناصر أداة القياس للبناء المستهدف تكون غير مستقرة وبالتالي فان صدق المحتوى يختلف عبر الزمن فان البيانات الحديثة يجب أن تتطور وتتغير أيضا عبر الزمن. (Haynes & Waialae, 1994)

ومن ثم فان الطبيعية الدينامية لتعريف الأبنية أثرت في أربعة جوانب هي: إن مؤشرات صدق المحتوى لا يمكن الاحتفاظ بها ثابتة عبر الزمن، وان صدق المحتوى لأدوات التقييم السيكولوجي يجب أن تختبر بصفة دورية، وان أدوات القياس السيكولوجي تحتاج لمراجعة دورية لتعكس تعديلات البناء المستهدف ، وانه من الخطأ استخراج بيانات لبناء معدل من خلال أداة قياس غير معدلة . (Haynes, Richard, Kubany,1995)

ويتم التحقق من صدق المحتوى على مرحلتين متمايزتين هما : الجهود الأولية التي قام بها مطور المقياس لتحسين صدق المحتوى من خلال وضع التعريف وتحليل المجال أو المحتوى لتوليد المفردات والجهود البعدية لتقويم مناسبة محتوى المقياس من خلال تقييمات المحكمين والخبراء .

وبالنسبة للجهود الأولية فقد أشارت عدة دراسات إلى خطوات التحقق من صدق المحتوى مثل دراسة (Haynes, Richard, Kubany,1995) فقد أشارت إلى سبعة خطوات للتحقق من صدق المحتوى تتمثل في تعريف المجال والأوجه للبناء المستهدف والموضوع الذي يهدف للتحقق من صدق محتواه، إخضاع جميع عناصر أداة التقييم لصدق المحتوى، استخدام المجتمع والخبراء لتقييم توليد المفردات والعناصر، استخدام عدة محكمين لصدق المحتوى وتحديد الأحكام من خلال استخدام تدريج

للإجراءات، اختبار نسبة التمثيل للمفردات، استخدام التحليلات النفسية المتتابعة لتعزيز أداة القياس.

بينما أشارت دراسة (Allahyari, Rangi, Khosravi and Zayeri,2011) إلى خطوات التحقق من صدق المحتوى للاستبيانات وهى : تعريف مجال المحتوى للاستبيان مثل: الفشل المعرفي، تعريف التصنيفات الخاصة بمجال المحتوى مثل : الفشل المعرفي يعرف بأنه الفشل في الانتباه والذاكرة، وقياس الوظائف التي تؤدى لضعف الانتباه، تعديل المفردات وتطوير صور أولية للاستبيان حيث يتم عرض قائمة من المفردات على الخبراء للحكم عليها للتأكد من إن جميع عناصر الفشل المعرفي تم تمثيلها وأن كل المفردات التي تم وضعها تمثل جميع التصنيفات المقترحة في مجال المحتوى وأن المفردات تمثل مجال المحتوى للفشل المعرفي، تطوير مسودة للاستبيان ذات كفاءة عالية بعد إجراء التعديلات التي تم الإشارة إليها في الخطوة السابقة، تحديد نموذج صدق المحتوى حيث يتم تحديد الطريقة المناسبة للتحقق من صدق المحتوى لعرض الاستبيان على الخبراء ثم اختيار مؤشر صدق المحتوى، تعديل النموذج حيث يتم إجراء تعديلات على النموذج لتتماشي مع الفروق الفردية المراد إيضاحها فهنا تم التعديل على مؤشر صدق المحتوى من تدريج إلى تدريج خماسي، تعريف الخبراء الذين سوف يتم أخذهم كقائمة للتحقق من صدق المحتوى للحكم على مسودة الاستبيان وهنا يتم تحديد عدد الخبراء المناسبين وتعريف مجالات تخصصهم ومناسبتها لتحكيم مسودة الاستبيان، الحصول على البيانات المحصلة من تقييمات المحكمين وإجراء التحليلات الإحصائية المناسبة لها، الحساب الكمي للاتساق بين المحكمين باستخدام المعادلة المناسبة لحساب الاتساق بين المحكمين، حساب متوسطات تقديرات التحكيم حيث يتم حساب متوسط مؤشرات المحتوى لكل مفردة وللاستبيان ككل ويلي ذلك تحديد المفردات المقبولة والمفردات طبقا للمحك المناسب.

وقد صنفت طرق التحقق من صدق المحتوى إلى مدخلين هما : المدخل الكيفي ويتمثل في مجرد الحصول على أحكام وآراء الخبراء عن مدى تمثيل ومواءمة أداة القياس للمحتوى الذي تقيسه بينما المدخل الثاني وهو المدخل الكمي يتعدى مجرد الوقوف على الآراء إلى إجراء حسابات كمية لآراء المحكمين والخبراء باستخدام نماذج رياضية للتحقق من صدق المحتوى مثل: طريقة نسب الاتفاق لكاتز، وأسلوب الاختبار القبلي البعدى لبروف، والأساليب المعتمدة على مقاييس ليكارت، ومؤشر صدق المحتوى، ومؤشر كابا لكوهين. (Vogt,king,king,2004)

وقد وجهت العديد من الانتقادات للمدخل الكمي حيث يؤخذ على نسب الاتفاق الحساسية لعدد البنود وتجاهل الترجيح للبند، ويؤخذ على أسلوب الاختبار القبلي البعدى من خلال وجود محك خارجي أنها تفترض عدم وجود معرفة مسبقة لدى المتعلمين وأن الاختبار المحك والاختبار الجديد سوف يرتبطان ولكن قياس المجال السيكولوجي قد يكون خاطئ بمعنى أن الاختباران قد يرتبطان في الغاية ولكن قياس المجال ربما يكون في الأساس خاطئ بينما يؤخذ على طريقة كلاين وكوسيكوف

للعلاقة بين الأهداف والمفردات تأثرها بكل من عدد البنود وعدد الأهداف والحساسية لعدد المحكمين وصياغة التعليمات للمحكمين، ويؤخذ على مؤشر المناسبة لموريس وفتيس جيبونس أنه مؤشر محدود بالحسابات المستخرجة من الدرجات التي تقيس أبنية مختلفة مثل: الأهمية والمناسبة فضلا عن التباين بسبب الاختلافات في مقياس ليكارت الخماسي مقابل الثنائي وتعانى الطرق المعتمدة على ليكارت من درجة خطأ في تباين القياس في اتفاق الخبراء. (Miller,2003)

وقد وجهت اقتراحات زيادة حجم العينة لمواجهة التباين حيث يقترب من الاعتدالية ويقل خطأ القياس حيث تزداد الدقة مع زيادة حجم عينة المحكمين وأن زيادة عدد المحكمين أكثر من خمسة يسهل عملية الكشف واستبعاد القيم المتطرفة كما أن زيادة عدد المحكمين خمسة أو أكثر سوف يتحكم في عامل الصدفة في تقديرات الاتفاق ولكن القدرة على تحديد عدد كبير من المحكمين لتقليل الخطأ المعياري ليس دوما متاح. (Beckstead,2009)

وقد وضع (Lawshe,1975) طريقة لحساب صدق المحتوى تقوم على استخراج مؤشر لصدق المحتوى لكل مفردة من مفردات الاختبار على حده وللاختبار ككل وبالتالي تتضمن نتائج تلك الطريقة الكمية معاملين: معامل صدق المفردات ويمثل مدى تمثيل كل مفردة من مفردات الاختبار للمحتوى السلوكي الخاص بها والمحدد بالهدف السلوكي الذي يمثل من خلال تصنيف الأهداف, ومعامل صدق الاختبار : والذي يعبر عن تمثيل الاختبار للمحتوى السلوكي والأهداف المحددة من قبل التصنيف الذي يهدف المحتوى لتحقيقه .(Lai & Chang, 2007)

وقد وضع (Lawshe,1975) عده افتراضات أساسية تقوم عليها طريقته في التقدير الكمي لصدق المحتوى تتمثل في النقاط التالية :

1- يتراوح معامل صدق المحتوى بين −1 , +1
2- إذا كان أقل من نصف عدد المحكمين أشاروا أن المفردة غير ضرورية أي غير صادقة فإن قيمة معامل صدق المحتوى تكون بالسالب .
3- إذا أشار نصف عدد المحكمين أن المفردة ضرورية وصادقة والنصف الآخر أشار أنها غير ضرورية وغير صادقة فإن قيمة معامل صدق المحتوى تكون صفر .
4- إذا أشار كل المحكمين أن المفردة ضرورية وصادقة فإن قيمة صدق المحتوى تساوى الواحد الصحيح .

وقد عبر عن معادلة صدق المحتوى من خلال معادلة نسبة صدق المحتوى content validity ratio لقياس صدق المحتوى لكل مفردة من مفردات الاختبار حيث يتم عرض استمارة التحكيم على السادة المحكمين والخبراء في المجال للحكم على كل مفردة من مفردات الاختبار من حيث الانتماء للبعد أو السمة

المقاسه على سبيل المثال: عرض محكات التحكيم في شكل ثلاثة بدائل هي : ضرورية , مفيدة وليست ضرورية, غير ضرورية .والمعادلة صيغتها كالتالي :

$$CVR_i = \frac{n_e - \frac{N}{2}}{\frac{N}{2}}.$$

حيث تشير (CVRi) إلى نسبة صدق المحتوى للمفردة و (ne) إلى عدد المحكمين الذين أشاروا أن المفردة ضرورية, و (N) يشير إلى العدد الكلى للمحكمين

(Prasad & Reghunath , 2011)

أما إذا استخدم الباحث بدائل ثنائية البعد مثل (ضرورية / غير ضرورية), والبعد الثاني (مهمة/غير مهمة) فإن الباحث في تلك الحالة يلجأ لاستخدام صيغة المعادلة التالية :

$$CVR_{Relaxed} = \frac{N_e + N_i/2 - N/2}{N/2} \quad (1)$$

حيث يشير (CVR relaxed) إلى نسبة صدق المحتوى للمفردة, و(Ne) عدد الخبراء الذين أشاروا أن المفردة ضرورية , و (Ni) عدد المحكمين الذين أشاروا أن المفردة مهمة, (N) العدد الكلى للمحكمين أو الخبراء .

(Karaiskos, Kourouthanassis, Giaglis , 2009)

ومن خلال قيمة نسبة صدق المحتوى للمفردات يمكن حساب مؤشر صدق المحتوى للاختبار (content validity idex) من خلال إيجاد متوسط معاملات نسب صدق المحتوى للمفردات ككل كما يتضح من المعادلة التالية :

$$CVI = \frac{\sum_n^1 CVR}{Retained\ numbers}$$

بعد تطبيق المعادلتين (1 , 2) يكون الباحث قد حصل على نسبة صدق المحتوى لكل مفردة ومؤشر صدق المحتوى للاختبار, وبالتالي فهو في حاجة لمحك لكيفية الحكم على ذلك المعامل بأنه مقبول أو غير مقبول, ومن ثم فقد وضع (Lawshe,1975) جدول لمعامل صدق المحتوى يوضح العلاقة بين عدد المحكمين والقيمة الحرجة لمعامل صدق المحتوى التي تكون دالة عند مستوى(0.05) كما يتضح من الجدول (1)

جدول (1)

عدد المحكمين	أقل قيمة مقبولة لمعامل صدق المحتوى

0.99	5
0.99	6
0.99	7
0.75	8
0.78	9
0.62	10
0.59	11
0.56	12
0.54	13
0.51	14
0.49	15
0.42	20
0.37	25
0.33	30
0.31	35
0.29	40

حيث يتضح من الجدول (1) أن أقل عدد من المحكمين (5) محكمين وأعلى عدد بالجدول (40) يقابل العدد القيمة الحرجة لمعامل صدق المحتوى والتي إذا بلغها معامل صدق المحتوى للمفردات والاختبار أو زاد عنها يتم الإبقاء على المفردة والحكم على النتائج بأنها صادقة .

وقد عرض (Brinkman,2009) جدولا مطورا للحكم على معاملات

جدول (2)

حجم المحكمين	الحجم الحرج	قيمة المعامل الحرجة
5	5	1
6	6	1
7	7	1
8	7	0.75
9	8	0.78
10	9	0.80
11	9	0.64
12	10	0.67
13	10	0.54
14	11	0.57
15	12	0.60
20	15	0.50
25	18	0.44
30	20	0.33
35	23	0.31
40	26	0.30
45	29	0.29
50	32	0.28
60	37	0.23
70	43	0.23
80	48	0.20
90	54	0.20
100	59	0.18

يتضح من الجدول (2) أن عدد المحكمين تراوح من (5 محكمين إلى 100 محكم) كما أن الجدول يحدد العدد الحرج من المحكمين المقابل لعدد المحكمين والقيمة الحرجة لمعامل صدق المحتوى لكن هذا الجدول يعطى تقديرات صارمة حيث لاحظ ذلك (Brinkman,2009) لذلك قسم معاملات صدق المحتوى إلى معاملات صارمة وهى التي تستخدم في حالة الاختبارات الطويلة جدا حيث يحتاج الباحث أن يخرج منها بنسخة مختصرة لذلك يقتصر على حساب تكرارات المحكمين الذين أشاروا أن المفردات " ضرورية " فقط أما النمط الثاني فهو معامل صدق المحتوى غير الصارم والذي يستخدم في حال الاختبارات أو الاستبيانات القصيرة ومن ثم يقوم الباحث بتقدير تكرارات المحكمين الذين أشاروا أن المفردة " ضرورية " والمحكمين الذين أشاروا أن المفردة " مفيدة وليست ضرورية " في تصنيف واحد وبالتالي يقل عدد المفردات التي يجب حذفها ويرتفع معامل صدق

المحتوى , ويعتمد الباحث الجدول (1) لأنه الأكثر استخداما في الدراسات والبحوث العلمية والذي بني عليه (Lawshe,1975) طريقته في حساب معامل صدق المحتوى .

ومن ثم يمكن تفسير نتيجة معامل صدق المحتوى لكل من المفردات والاختبار في ضوء جدول (1) من خلال مقارنة قيمة معامل صدق المحتوى المحسوبة بقيمة معامل صدق المحتوى الجدولية فإذا بلغت القيمة المحسوبة القيمة الجدولية أو زادت عنها يتم الحكم على المفردات بأنها صادقة ويجب البقاء عليها كما أن قيم معاملات المحتوى التي بالسالب يتم استبعادها لأنها تدل على أن أقل من نصف عدد المحكمين أشاروا أن المفردة ضرورية أما قيمة معامل صدق المحتوى التي تساوى صفر فتشير إلى تساوى عدد المحكمين الذين أشاروا أن المفردة ضرورية مع عدد المحكمين الذين أشاروا أنها ليست ضرورية أما القيمة واحد صحيح فتؤكد أن كل المحكمين أشاروا أن المفردة ضرورية وفى تلك الحالة يجب التأكد من أن هؤلاء المحكمين خبراء في المجال بالفعل لأنه عندما يتفق الكل فهذا يؤدى إلى الشك في تمكن المحكمين لذا فإن كانوا خبراء بالفعل في المجال فيجب التأكيد على أن المفردة ضرورية أما إذا لم يتفق كل المحكمين على أن المفردة ضرورية فهذا يؤكد أن المفردة بالفعل غير ضرورية وإذا أشار نصف عدد المحكمين أو أكثر أن المفردة ضرورية فهي ضرورية وكلما زاد عدد المحكمين عن النصف زادت درجة الثقة في صدق النتائج (Allahyari et al., 2011).

وقد أشار (Lawshe,1975) إلى طريقة المتوسط الوزنى للمفردات حيث يقوم الباحث بتقدير البدائل الثلاث للتحكيم (ضرورية, مفيدة وليست ضرورية, غير ضرورية) بإعطائها درجات (2, 1, 0) على الترتيب ثم استخراج المتوسط الحسابي لتقديرات المحكمين لكل مفردة وبالتالي يكون المحك بالإضافة إلى قيمة معامل صدق المحتوى الجدولية قيمة المتوسط لتقديرات المحكمين والتي تقدر في هذا التدريج (1.5).

فالمفردات التي بلغت قيمتها قيمة معامل صدق المحتوى الجدولية ومتوسطها (1.5) فأكثر تكون مقبولة, والمفردة التي لم تبلغ قيمة معامل صدق المحتوى الجدولية لكن تراوحت قيمتها من (صفر – القيمة الجدولية) وقيمة المتوسط (1.5) فأكثر تكون مقبولة لأن أكثر من نصف عدد المحكمين أشاروا بأنها ضرورية, أما المفردة التي جاءت قيمة معامل صدق المحتوى لها أقل من الصفر أي بالسالب ومتوسطها أقل من (1.5) تكون مرفوضة .

ورغم أهمية طريقة معامل صدق المحتوى باعتبارها طريقة كمية لتقدير صدق المحتوى تتضمن حساب صدق المحتوى لكل من المفردات والاختبار وتعطى قيمة محسوبة لصدق المحتوى يمكن مقارنتها بقيمة معامل صدق المحتوى الجدولية التي تتمتع بثقة (95%) إلا أن العديد من

الأبحاث التربوية والنفسية تركز على نسب اتفاق المحكمين كطريقة لتقدير صدق المحتوى للمفردات حيث تتخذ نسبة الاتفاق (75%) كمحك للحكم على صدق المفردات (Yaghmal,2003)

لذا فإن البحث الحالي يركز على دراسة واقع ممارسة صدق المحتوى في البحوث التربوية والنفسية من خلال الاطلاع على الأبحاث التي تناولت صدق المحتوى والقيام بتحليل واقع استخدامها من حيث المفهوم والمصطلح المستخدم وطريقة التحقق من الصدق والتفسير له من ناحية ومن ناحية أخرى يتم دراسة ذلك الواقع كما يدركه الباحثون والخبراء من خلال طرح سؤال مفتوح عن واقع ممارسة صدق المحتوى ومشكلاته ثم ينتقل البحث لعرض مقارنة بين طريقة نسب الاتفاق وطريقة مؤشر صدق المحتوى ليعرض على البحث العربي والمصري نتائج تفيده في مجال القياس النفسي والتقويم التربوي وتعين المعلم والباحث على الاستخدام المقبول علميا لصدق المحتوى كما هو مطروح عالميا .

ثانيا : مشكلة البحث :

يعد صدق المحتوى أحد أنواع الصدق الثلاثة المعترف بها في نظرية الصدق التقليدية حيث يتوقف عليه تفسير الدرجات والنتائج المستخرجة وكثير من الأمور المتعلقة باتخاذ القرارات وخاصة في المجال النفسي والتربوي وهذا ما دفع العلماء نحو دراسته فقد دارت كثير من المناقشات العلمية والجدال والتضارب في الآراء ما بين مؤيد ومعارض ورافض رفضا كليا لفكرة صدق المحتوى وتعددت طرق التحقق منه وقياسه .

وقد توصلت نقاشات المتخصصين حول صدق المحتوى وممارساته إلى نقطتين أساسيتين هما : صدق المحتوى ضروري في بناء وتطوير أدوات القياس الجديدة فهو خطوة ضرورية في تطوير قياس تجربة جديدة لأنه يمثل آلية مبدئية لربط المفاهيم المجردة بمؤشرات قابلة للملاحظة والقياس ، أن مؤشر صدق المحتوى الكمي الذي يستخرج من نسب الاتفاق يعانى من التحيز لأنه لا يستطيع ضبط عامل صدفة الاتفاق لذا أوصوا باستخدام معامل كابا لكوهين لأنه يواجه تلك المشكلة . (Wynd et al., 2003)

فصدق المحتوى هو مخرجات التحكيم حول تمثيل وكفاية محتوى الاختبار بشكل مناسب للتعريف الإجرائي للمحتوى العام وبذلك يصلح استخدام صدق المحتوى مع اختبار التحصيل لأن لها محتوى محدد يمكن مقارنة المقياس في ضوئه كذلك مقياس الاتجاهات والاستبيانات يمكن حساب صدق المحتوى لها من خلال المجال السلوكي والذي يعتمد على التعريف الإجرائي والمعضلة هنا أن التعريف الإجرائي ليس بالضرورة أن يؤدى إلى تمثيل صادق للبناء المستهدف قياسه . (Beckstead,2009)

وقد وجه لصدق المحتوى عدة انتقادات تتمثل في أن الصدق يتعلق بالدرجات التي يتم الحصول عليها من تطبيق الاختبار وتفسيرها بينما صدق المحتوى يعتمد على شكل الاختبار قبل تطبيقه بدون أن يكون هناك أية درجات وبالتالي فهو ليس صدقا فالاختبارات في حد ذاتها ليس لها تحقق من الصدق والثبات إنها خصائص لاستجابات الاختبار وليس للاختبار نفسه . وأن تغطية المحتوى مهم أخذه في الاعتبار عند التحقق من بنية الاختبار وتفسيره ليكون موثوق فيه لكنه في حد ذاته ليس صدقا. (Sireci,1998)

وقد نشر (Beckstead, 2009) ورقة بحثية هدفت إلى إقناع العلماء المهتمين بنشر صدق المحتوى أنهم يضللون طاقاتهم وأن الممارسات الكمية لنسب الاتفاق تعد ضعيفة في الأساس .حيث أشار البحث إلى ما أشار إليه (Messick,1975) حيث اعتبر صدق المحتوى مجرد تمهيد للصدق فقط فصدق المحتوى يتعامل مع استنتاجات حول درجات الاختبار من خلال التعريف فإن كل الصدق هو الدقة في الاستنتاجات حول درجات الاختبار أما ما يطلق عليه صدق المحتوى فهو ليس صدقا على الإطلاق .

بينما أشار (Haynes & Waialae, 1994) إلى أن صدق المحتوى يقدم معلومات حول البيانات المحصلة من أداة القياس والاستدلالات التي تم الحصول عليها من تلك البيانات فالاستبيانات التي تعانى من خلل في صدق المحتوى سوف تكون الاستنتاجات الناتجة عن تطبيقها متأثرة بذلك الخلل، فالدرجة الحقيقية يمكن الحصول عليها من أداة قياس (سواء صادقة أو غير صادقة) ولكن البيانات سوف تكون مثيرة للشبهة عندما تكون عناصر أداة القياس غير مناسبة للبناء المستهدف والغرض والتقييم .

ويتضح للباحث من خلال الإطلاع على الأدبيات التي تناولت صدق المحتوى وجود انتقادات كثيرة لاستخدام نسب الاتفاق في تقدير صدق المحتوى تمثلت تلك الانتقادات فى أن الطريقة تقيس مدى مناسبة وعدم مناسبة المفردة ، ولا تهتم إذا كان هناك قائمة تشمل مفردات تمثل المجال ككل حيث أشار كوهين (1960) أنها طريقة أكثر بدائية تفتقد إلى مؤشر عدم الاتفاق علاوة على أنها تؤدى إلى تضخم الاتفاق الذي يرجع لتأثير عامل الصدفة .(Wynd et al., 2003)

ورغم تلك الانتقادات التي وجهت إلى طريقة نسب الاتفاق في تقدير صدق المحتوى فإن الباحث لاحظ اعتماد البحث التربوي العربي والمصري على نسب الاتفاق في الحكم على صدق المحتوى والذي يسميه أغلب الباحثين بمسمى " صدق المحكمين " .

ويرى العديد من علماء السلوك أن التقييمات المستخدمة يجب أن تتحقق فيها مؤشرات لصدق المحتوى،

ومع ذلك فإن مفهوم صدق المحتوى مثيرا للجدال منذ إنشاؤه حيث ظهرت مجادلات كثيرة ضد استخدام صدق المحتوى، حيث يرى أن صدق المحتوى فنية غير صحيحة وغير مقبولة سيكومتريا وهذا رأى غير مقبول على المستوى العلمى الواسع. لاحظ الباحث من خلال الاحتكاك بالباحثين والمتخصصين التربويين وجود خلط عند الكثير منهم بين مفهوم صدق المحتوى ومفهوم صدق المحكمين حيث يعتقد الكثير من الباحثين أن صدق المحكمين هو صدق المحتوى بينما يعتقد البعض الآخر أن صدق المحكمين للاستبيانات والتقارير الذاتية، كما يعتقد البعض أن صدق المحتوى ليس له معامل ومن ثم يقتصر فقط على الطرح الكيفي للبيانات أو النسبة العامة للاتفاق . (Sireci,1998)

كل تلك المعتقدات أثارت الباحث نحو التوسع في الأدبيات الخاصة بذلك حيث استنتج الباحث من تلك الأدبيات أن صدق المحتوى هو أحد ثلاثة أنواع معتمدة عالميا وعلميا لحساب الصدق وليس في الأدبيات ما يسمى بصدق المحكمين وأن التحكيم هو خطوة من خطوات التحقق من صدق المحتوى .

ثانيا : إن صدق المحتوى يشمل جميع أدوات القياس أي اختبارات القدرات العقلية واختبارات التحصيل ذات المحتوى المعروف وكذلك الاستبيانات وبطاقات الملاحظة التي تعتمد على تعريفات إجرائية واضحة ومفاهيم سلوكية مشتقة من التعريفات الإجرائية تساعد على بناء الأداة .

ثالثا : لاحظ الباحث أن صدق المحتوى له معامل وله معايير واضحة في المجتمع الغربي في حين أن مجتمعنا العربي لم يحدد أية معايير أو قيم أو حتى تطوير لفكرة صدق المحتوى .

كما لاحظ الباحث من خلال الأدبيات أن هناك عدة نقاط جدال موسعة لدى المختصين منذ ظهور صدق المحتوى تمثلت في تحقق مفهوم الصدق في صدق المحتوى حيث أن الهدف من الاختبار هو الحصول على درجات تساعد على الوصول إلى استنتاجات مناسبة تؤدى إلى اتخاذ قرارات صائبة ومن ثم فإن الصدق يتحقق بالدرجات وتفسيرها في حين أن صدق المحتوى أخذ شكلا مغايرا عن جميع أنواع الصدق فهو يقدر قبل تطبيق الاختبار لأنه حكم على شكل الاختبار ولغته ومناسبته للأهداف والتعريف الإجرائي والمرحلة العمرية والتعليمية والجنس..الخ وكل ذلك سابق للحصول على الدرجات مما أدى بالمتخصصين إلى اعتباره مجرد تمهيد للصدق فهو خطوة تمهيدية للإعداد وليس صدقا ومن ثم ظهرت اعتقادات تنادى بأن صدق المحتوى ليس صدقا ولكن تلك الاعتقادات لم تلاق رواجا عالميا فقد وجه لها العديد من الانتقادات. (Beckstead, 2009)

كما ظهرت أفكار جدلية حول المحكمين وعددهم المناسب والنسبة بين عدد المحكمين وعدد المفردات ومن ثم من هو المحكم؟ وكيفية تحديده ؟ فإذا تم تحديد المحكم بدقة فهذا يؤدى إلى نتائج

صادقة حول الحكم على أداة القياس .

وظهر جدال آخر موسع حول كيفية التحقق من صدق المحتوى فهناك من اتخذ توجها كيفيا فنادي بالتحقق الكيفي وآخرون اتخذوا نهجا كميا أدى إلى ظهور العديد من الطرق الكمية للتحقق من صدق المحتوى وظهرت نقاشات وشكوك حول تقديرات المحكمين وعلاقتها بذاتية المحكم، وتزييف الباحثون للنتائج ومدى إلمام المحكم بالموضوع فنتائج التحكيم تعانى من التضخم والتركيز على الاتفاق وإغفال جوانب عدم الاتفاق علاوة على قيام الباحثين باستخدام إجراء تقليص البدائل من تدريج خماسي إلى تدريج ثنائي على سبيل المثال مما أدى إلى إخفاء كثير من النواتج بالإضافة لعدم القدرة على تجنب عامل صدفة الاتفاق . (Beckstead, 2009)

ورغم تلك النقاشات والخلافات التي دارت في المجتمع الغربي إلا أن المجتمع العربي بعيد جدا عن هذه النقاشات مما أدى إلى قصور واضح لدى الباحثين في مفهوم صدق المحتوى وطرق التحقق منه وتفسيره وكيفيه تجنب الخطأ في التطبيق .

كما دارت نقاشات واسعة حول صدق المحتوى هل هو سمه أو حالة للاختبار بمعنى إذا كان صدق المحتوى سمة ثابتة للاختبار فهو لا يحتاج لإعادة التحقق منه مرة أخرى عبر الزمن أما إذا كان حالة وقتية فهو يحتاج لإعادة تحكيم عبر الزمن من أجل مواكبة التطورات وقد أشار المتخصصين أن صدق المحتوى حالة وقتية فهو مفهوم دينامي متطور لذا يحتاج لإعادة تطبيق عبر الزمن والثقافة . (Haynes, Richard, Kubany,1995)

وهذا المعنى يفتقد في البحث التربوي العربي حيث يقتصر التحقق من صدق المحتوى على التطبيق على المقياس المعد لأول مرة فقط وإذا أراد باحث إعادة تطبيق الاختبار بعد فترة زمنية لا يعيد النظر في صدق المحتوى بل يعتبره مسلمة أو سمة ثابتة للاختبار .

من خلال ذلك العرض تمثلت مشكلة البحث الحالي في دراسة واقع ممارسات صدق المحتوى في البحوث التربوية للإجابة عن التساؤلات التالية :

- ما واقع ممارسات التحقق من صدق المحتوى في البحوث التربوية ؟
- هل توجد علاقة بين عدد المحكمين ومعيار الحكم على صدق المحتوى ؟
- ما واقع تطبيق صدق المحتوى في البحوث التربوية كما يدركه الباحثون ؟
- هل تختلف عدد المفردات المقبولة باختلاف طريقة حساب صدق المحتوى (طريقة معامل صدق المحتوى – طريقة نسب الاتفاق) ؟
- هل يمكن تطوير جداول معامل صدق المحتوى لإيجاد النسبة الحرجة لطريقة نسب الاتفاق من

خلالها ؟

- هل يتأثر معامل صدق المحتوى بحجم عينة المحكمين (10 محكمين, 20 محكم) ؟

أهداف البحث :

يهدف البحث الحالي إلى :

-دراسة واقع ممارسات التحقق من صدق المحتوى في البحوث التربوية.

- دراسة العلاقة بين عدد المحكمين ومعيار الحكم على صدق المحتوى.

-دراسة واقع تطبيق صدق المحتوى في البحوث التربوية كما يدركه الباحثون.

- المقارنة بين طريقة معامل صدق المحتوى وطريقة نسب الاتفاق في قبول أو رفض المفردات .

- تطوير جداول معامل صدق المحتوى لتتضمن النسب المئوية لاتفاق المحكمين المقابلة لكل معامل من معاملات صدق المحتوى .

- التعرف على أثر حجم عينة المحكمين في قبول أو رفض المفردات طبقا لمعامل صدق المحتوى .

- تقديم مفهوم معامل صدق المحتوى ليساعد الباحثين في تطوير أساليب قياس صدق المحتوى كميا .

- تقديم مقارنة بين طريقة معامل صدق المحتوى وطريقة نسب الاتفاق مما يوضح للباحث العربي أهمية استخدام طريقة معامل صدق المحتوى كبديل لطريقة نسب الاتفاق .

- تطوير جداول معامل صدق المحتوى لإنتاج نسب مئوية حرجة مقابلة لعدد المحكمين مما يسهل على الباحث المستخدم لنسب الاتفاق الدمج بين الطريقتين .

- الوقوف على أثر حجم عينة المحكمين على صدق المحتوى مما يفيد الباحث في انتقاء الحجم المناسب .

أهمية البحث :

تتضح أهمية البحث من خلال النقاط التالية :

- تقديم مفهوم معامل صدق المحتوى ليساعد الباحثين في تطوير أساليب قياس صدق المحتوى كميا .

- يقدم البحث تقويما لواقع ممارسات التحقق من صدق المحتوى في البحوث التربوية مما يساعد على تنوير الباحث العربي بنقاط القوة والضعف في التحقق من صدق المحتوى .

- دراسة واقع ممارسات ومشكلات التحقق من صدق المحتوى كما يدركه الخبراء والباحثين مما يفيد في الوقوف على مفاهيم الباحثين عن صدق المحتوى وممارساته الفعلية وتصورات العالم لتلك

الفنيات العلمية .

- تقديم مقارنة بين طريقة معامل صدق المحتوى وطريقة نسب الاتفاق مما يوضح للباحث العربي أهمية استخدام طريقة معامل صدق المحتوى كبديل لطريقة نسب الاتفاق .
- تطوير جداول معامل صدق المحتوى لإنتاج نسب مئوية حرجة مقابلة لعدد المحكمين مما يسهل على الباحث المستخدم لنسب الاتفاق الدمج بين الطريقتين .
- الوقوف على أثر حجم عينة المحكمين على صدق المحتوى مما يفيد الباحث في انتقاء الحجم المناسب .

التعريفات الإجرائية :

صدق المحتوى : يقصد به تمثيل الاختبار من خلال مفرداته للمحتوى ويعرفه الباحث إجرائيا ما تقسية معادلة نسب اتفاق المحكمين ومعادلاتي نسبة صدق المحتوى للمفردات ومؤشر صدق المحتوى للاختبار حول تمثيل المفردات والاختبار للمحتوى السلوكي .(الباحث)

نسبة صدق المحتوى للمفردات: ما تقيسه معادلة نسبة صدق المحتوى للمفردات ونسب اتفاق المحكمين حول تمثيل المفردات للمحتوى السلوكي .(الباحث)

مؤشر صدق المحتوى للاختبار: ما تقيسه معادلة مؤشر صدق المحتوى والذي يمثل متوسط معاملات صدق المحتوى للاختبار ككل .(الباحث)

إجراءات البحث :

منهج الدراسة **:** استند البحث على إتباع المنهج الوصفى التحليلى البعدى من خلال التحليل البعدى لرسائل الدكتوراه من ناحية والمسحى من خلال جمع أراء الخبراء والمتخصصين حول واقع ممارسة صدق المحتوى والعلاقاتى من خلال ايجاد العلاقة الارتباطية بين معايير الحكم وعدد المحكمين والفارق من خلال البحث عن الفروق بين طريقتى نسب الاتفاق ومعامل صدق المحتوى .

عينة البحث : تكونت عينة البحث من عدد (30) من الخبراء المتخصصين فى مجال التربية وعلم النفس والمعلمين بالتربية والتعليم ، عدد (70) رسالة دكتوراه تم اجازتها بكلية التربية جامعة المنيا منذ عام (2000) حتى (2014)

أدوات البحث :

اعتمد الباحث في جمع البيانات على ثلاث أدوات هي :

- استبانة مهارات المعلم داخل حجرة الدراسة : قام الباحث بإعداد استبانة مكونة من (10) مفردات تمثل (10) مهارات للمعلم .
- استمارة تحليل صدق المحتوى من إعداد الباحث: حيث أعد الباحث استمارة لتحليل واقع ممارسات صدق المحتوى في البحوث التربوية وشملت عدة عناصر: القسم العلمى، المصطلح المستخدم، عدد

المحكمين، طريقة التحقق، معيار الحكم .

– سؤال مفتوح تم عرضه على الخبراء للإجابة عن واقع ومشكلات ممارسة التحقق من صدق المحتوى كما يدركونه. حيث طرح الباحث سؤال مفتوح على الخبراء " ما واقع ممارسات صدق المحتوى ومشكلاته من وجهة نظركم ؟"

المعادلات الإحصائية المستخدمة :

استخدم الباحث للتحقق من أسئلة البحث المعادلات التالية :

1–معادلة النسبة المئوية لتحديد نسب اتفاق المحكمين .

2– معادلة نسبة صدق المحتوى للمفردات.

3– معادلة مؤشر صدق المحتوى للاختبار .

4- معادلة المتوسط الحسابي .

5- معادلة مان ويتنى .

خطوات التطبيق :

1- تحليل رسائل الدكتوراه التي تم إجازتها بكلية التربية جامعة المنيا منذ عام (2000 وحتى 2014) ثم حساب الإحصائيات الوصفية لواقع استخدام صدق المحتوى من حيث المسمى المستخدم والطريقة وعدد المحكمين والمعيار المستخدم .

2- عرض التساؤلات المفتوحة عن واقع ممارسة صدق المحتوى ومشكلاته ثم استخلص الباحث نتائج التحليل في نقاط محددة .

3- إعداد استمارة التحكيم للسادة المحكمين (30محكم) تضمنت الاستمارة المفردات وثلاثة بدائل أمام كل مفردة هي :

- ضرورية
- مفيدة وليست ضرورية
- غير ضرورية

4–عرض الاستمارة على السادة المحكمين لإبداء رأيهم حول المفردات العشر .

5–حساب نسب الاتفاق ومؤشر صدق المحتوى ومتوسط المفردات .

6–اقتراح معادلة لتسهيل حساب مؤشر صدق المحتوى ومقارنتها بالمعادلة الأصلية

7–القيام بتطوير طريقة نسب الاتفاق لتكون دالة عند مستوى (0.05).

8–استخدام معادلة مان ويتنى لايجاد الفروق بين نسب الاتفاق ومؤشر صدق المحتوى ومتوسط المفردات فى حال التطبيق (10) محكمين مقابل التطبيق (20) محكم

نتائج البحث :

للإجابة السؤال الأول والذي ينص على " ما واقع ممارسات التحقق من صدق المحتوى في البحوث التربوية ؟" قام الباحث بالاطلاع على رسائل الدكتوراه التي منحت بكلية التربية جامعة المنيا منذ عام (2000 وحتى 2014) وقد وجد الباحث حوالي (70) رسالة استخدمت صدق المحتوى موزعة على الأقسام العلمية : قسم علم النفس التربوي، قسم الصحة النفسية، قسم أصول التربية، قسم تربية الطفل، قسم المناهج وطرق التدريس والجدول التالي يوضح نتائج تحليل الرسائل التي استخدمت صدق المحتوى للتحقق من صدق أدواتها .

جدول (3)تحليل لواقع استخدام صدق المحتوى

القسم	المصطلح المستخدم		الطريقة		المعيار	عدد المحكمين
	صدق المحتوى	صدق المحكمين	نسب الاتفاق	الكيفية		
علم النفس التربوي	%38	%62	%87	%13	%90-75	21-7
الصحة النفسية		%100	%78	%22	%100-70	16-7
أصول التربية		%100	%16	%84	%84-70	46-11
تربية الطفل		%100	%44	%56	%88-66	21-7
المناهج وطرق التدريس	%15	%85	%25	%75	%90-80	25-9

يتضح من الجدول (3) ما يلي :

1- انتشار استخدام الباحثين لمصطلح صدق المحكمين للتعبير عن صدق المحتوى حيث بلغت النسبة الكلية لاستخدام مصطلح صدق المحكمين حوالي (99%) وهى تدل على انتشار مصطلح صدق المحكمين بشكل كبير رغم أن المصطلح العلمى المعترف به علميا وعالميا هو مصطلح صدق المحتوى وكانت أكبر نسبة لذكر صدق المحتوى في قسم علم النفس التربوي وقد لاحظ الباحث استخدام الباحثين لصدق المحتوى مع جميع أنواع أدوات البحث فلم يقتصر على الاختبارات التحصيلية فقط بل شمل التقارير الذاتية والاستبيانات واستطلاعات الرأي وبطاقات الملاحظة ...الخ من الأدوات .

2- أن طرق التحقق من صدق المحتوى تنوعت بين الكيفي والكمي ولكنها اقتصرت في الكمي على طريقة نسب الاتفاق فقط حيث تشير النتائج أن نسبة (54%) من الرسائل استخدمت نسب الاتفاق بين المحكمين بينما (46%) من الرسائل استخدمت المدخل الكيفي ويلاحظ في كلا الحالتين اقتصار التحقق على العرض على السادة الخبراء والمحكمين موضحين أن نتائج التحكيم هي صدق المحتوى وهذه مغالطة لأن التحكيم نفسه خطوة من خطوات التحقق من صدق المحتوى سواء كان كمي أو كيفي

3- أن معيار الحكم المقبول غير ثابت وليس له قواعد محددة لاختياره في البحث التربوي حيث تراوح بين (66-100%) لكن بدون تحديد سبب لاختيار معيار بدل من معيار آخر وليس هناك مبرر علمي يستخدمه الباحث لتحديد نسبة الاتفاق المقبولة كمعيار

4- أن انتقاء وتحديد عدد الخبراء والمحكمين ليس له مبرر في رسائل الدكتوراه المحللة في البحث

الحالي فليس هناك مبرر يطرحه الباحث لتحديد الكاف من المحكمين وتراوح عدد المحكمين من (7-46) محكم

5-كما يلاحظ من الجدول انتشار المدخل الكيفي للتحقق من صدق المحتوى في البحوث التربوية في أقسام الصحة النفسية، أصول التربية ، تربية الطفل . بينما كان أكثر انتشار للمدخل الكيفي في أقسام أصول التربية، وكانت أعلى نسبة حوالي (84%) لتخصص علم النفس التربوي يليه قسم الصحة النفسية ثم تربية الطفل ثم المناهج وطرق التدريس بينما جاءت أعلى نسب لاستخدام المدخل الكيفي في قسم أصول التربية يليه قسم المناهج وطرق التدريس ثم قسم تربية الطفل .

6- ومن خلال دراسة الباحث للنسبة الكلية اتضح أن حوالة (1%) فقط من البحوث استخدمت مصطلح صدق المحتوى وحوالى (99%) من البحوث استخدمت مصطلح صدق المحكمين .

وللإجابة عن السؤال الثاني والذي ينص على " هل توجد علاقة بين عدد المحكمين ومعيار الحكم على صدق المحتوى ؟" قام الباحث بدراسة العلاقة بين عدد المحكمين ومعيار الحكم المقبول فدلت نتائج معامل الارتباط عن عدم وجود علاقة ارتباطية دالة بين عدد المحكمين ومعيار الحكم المقبول وهذا يؤكد أن كلا من عدد المحكمين ومعيار الحكم يتم اختيارهم اعتباطيا أو بالصدفة أو كما هو متاح لدى الباحث وهذا يؤدى لظهور الذاتية لدى الباحث في انتقاء المحكمين من ناحية واختيار المحك المناسب من ناحية أخرى والجدول (4) يوضح معامل ارتباط الرتب لسبيرمان بين عدد المحكمين والمحك المستخدم .

جدول(4)عدد المحكمين ومعايير الحكم والعلاقة بينهما وتقييم معايير الحكم فى ضوء معامل صدق المحتوى

مسلسل	عدد المحكمين	معيار الحكم %	تقييم المعيار فى ضوء معامل صدق المحتوى
1	10	90	مقبول
2	10	80	مرفوض
3	13	84	مقبول
4	9	80	مرفوض
5	8	75	مرفوض
6	7	85	مرفوض
7	12	75	مرفوض
8	11	85	مرفوض
9	10	90	مقبول
10	16	95	مقبول
11	15	98	مقبول
12	13	85	مقبول
13	14	93	مقبول

مسلسل	عدد المحكمين	معيار الحكم %	تقييم المعيار فى ضوء معامل صدق المحتوى
14	8	93	مرفوض
15	11	90	مقبول
16	7	85	مرفوض
17	11	100	مقبول
18	7	85	مرفوض
19	13	100	مقبول
20	17	70	مقبول
21	19	84	مقبول
22	9	66	مرفوض
23	12	85	مقبول
24	9	88	مقبول
25	14	80	مقبول
26	13	72	مرفوض
27	7	80	مقبول
28	15	80	مقبول
29	14	85	مقبول
30	17	80	مقبول
31	14	85	مقبول
32	23	90	مقبول
33	9	85	مرفوض
34	17	85	مقبول
معامل الارتباط	0.044		

يتضح من الجدول (4) أن قيمة معامل الارتباط بين مدى عدد المحكمين ومعيار الحكم جاءت غير دالة مما يشير لعدم وجود تلازم لدى الباحث بين عدد المحكمين ومعيار الحكم وأن معيار الحكم على الأدوات يكون عشوائيا خاضع لعامل الصدفة .

وللإجابة عن السؤال الثالث والذي ينص على " ما واقع تطبيق صدق المحتوى فى البحوث التربوية كما يدركه الباحثون ؟" تم عرض سؤال مفتوح على عدد من الخبراء والمتخصصين في مجال التربية وعلم النفس عن واقع ممارسات صدق المحتوى ومشكلاته ثم قام الباحث بتصنيف وترتيب استجابات الخبراء والتي أشارت إلى أن واقع ممارسات صدق المحتوى ومشكلاته تتمثل فيما يلي :

- عدم إلمام الباحثين بمفهوم صدق المحتوى .
- عدم إلمام الباحثين بإعداد جداول المواصفات .
- عدم تحديد بعض الباحثين للتعريفات الإجرائية .
- عدم وجود معيار لتمييز المحكم عن غيره .
- عدم وجود قواعد لاختيار المحكمين .

- عدم وجود معايير واضحة للتحكيم .
- تأثر المحكمين بالذاتية .

للإجابة عن السؤال الرابع والذي يتعلق بالمقارنة بين طريقتي معامل صدق المحتوى وطريقة نسب الاتفاق " قام الباحث بتحليل نسب الاتفاق التى تم استخدامها كمحك للحكم على صدق الادوات من خلال تحليلها طبقا للمحك المقبول لمعامل صدق المحتوى حيث استخدم الباحث مقارنة بين المحك الذى استخدمه الباحثون فى دراساتهم فعليا والمحك المفترض استخدامه طبقا لمعامل صدق المحتوى والجدول (4) يوضح ذلك حيث اتضح أن عدد الدراسات التى اشارت لكل من المحك وعدد المحكمين بلغت (34) دراسة بلغت عدد المحكات المقبولة (22) بينما بلغت المحكات المرفوضة (12) أي أن حوالى (35%) من الدراسات محكاتها غير موثوق بها حيث قيمها غير دالة عند مستوى (0.05) وبالتالى فإن (35%) من رسائل الدكتوراه أشارت إلى تحقق صدق المحتوى فى مقاييسها كانت غير صادقة وهذا أمر يلفت النظر لضرورة الانتقال من التركيز على استخدام نسب الاتفاق إلى استخدام معامل صدق المحتوى .

ثم قام الباحث بتطبيق استبيان مكون من (10) مفردات على السادة المحكمين وتحليلها باستخدا طريقة نسب الاتفاق من ناحية و استخدام معادلة نسبة صدق المحتوى لتحديد معامل صدق المحتوى للمفردات ثم معادلة مؤشر صدق المحتوى لتحديد معامل صدق المحتوى للاختبار ثم استخراج متوسط تقديرات المحكمين لكل مفردة من مفردات الاختبار وبالتالي تم الحكم على المفردات في طريقة معامل صدق المحتوى من خلال قيمة معامل صدق المحتوى الجدولية والتي بلغت (0.62) عند مستوى (0.05) وقيمة المتوسط الحسابي (1.5) .

والطريقة الثانية نسب الاتفاق تم فيها استخراج النسبة المئوية لاتفاق المحكمين حول أن المفردات ضرورية والجدول (5) يوضح ذلك.

جدول (5) المقارنة بين طريقتى معامل صدق المحتوى ونسب الاتفاق

المفردات	البدائل			طريقة معامل صدق المحتوى			طريقة نسب الاتفاق		المعامل نهائي
	ضرورية	مفيدة	غيرضرورية	لمعامل	لمتوسط	الحكم	نسب الاتفاق	الحكم	
1	9	1	0	0.80*	1.9	مقبول	0.90	مقبول	0.8
2	7	3	0	0.40	1.7	مقبول	0.70	مرفوض	0.4
3	9	1	0	*0.80	1.9	مقبول	0.90	مقبول	0.8
4	7	2	1	0.40	1.6	مقبول	0.70	مرفوض	0.4
5	5	2	3	صفر	1.2	مرفوض	0.50	مرفوض	-
6	7	3	1	0.40	1.7	مقبول	0.70	مرفوض	0.4

7	2	7	0	-0.60	1.1	مقبول	0.20	مرفوض	-
8	8	2	0	0.60	1.8	مقبول	0.80	مقبول	0.6
9	8	2	0	0.60	1.8	مقبول	0.80	مقبول	0.6
10	9	1	0	*0.80	1.9	مرفوض	0.90	مقبول	0.8
معامل صدق المحتوى للاختبار				0.42					0.60

- *دال عند مستوى (0.05)

يتضح من جدول (5) ما يلي :

- بالنسبة لطريقة معامل صدق المحتوى:

1- جاءت ثلاثة مفردات فقط معاملاتها أعلى من القيمة الجدولية (0.62) وهى المفردات (1، 3، 10) حيث بلغت (0.80).

2- ومع استخدام معامل صدق المحتوى أكبر من أو يساوى (0.62) لأن عدد المحكمين (10) محكمين ومتوسط المفردة (1.5) كمحك لقبول المفردات فإنه ترتب على ذلك

- قبول المفردات غير المحققة لشروط معامل صدق المحتوى والتي تتراوح قيمة المعامل لها من صفر إلى (0.62) وقيمة متوسط المفردة لها أكبر من أو يساوى (1.5) وبالتالي تم قبول المفردات (2، 4، 5، 6، 8، 9) حيث أن القيمة صفر فما فوق تدل على أن نصف عدد المحكمين أو أكثر أشاروا أن المفردة ضرورية .
- قبول المفردات التي بلغت قيمة معامل الصدق لها (0.62) ومتوسطها أكبر من أو يساوى (1.5) وبالتالي تم قبول المفردات (1، 3، 10)
- رفض المفردات إلى بلغت قيمة معامل الصدق لها أقل من الصفر أي بالسالب ومتوسطها أقل من(1.5) وبالتالي تم رفض المفردة (7) لأن قيمتها بلغت (-.60) ومتوسطها (1.1)
- بلغت قيمة المعامل للاختبار ككل بعد استبعاد المفردة المرفوضة (0.60) وهى قيمة مقبولة.
- ومن ثم تم قبول (8) مفردات واستبعاد (2) مفردة .
- بالنسبة لطريقة نسب الاتفاق :

استخدمت القيمة (75%) كمحك لقبول المفردات التي بلغت ذات القيمة أو زادت عنها وبالتالي تم قبول جميع المفردات عدا المفردات رقم (2، 4، 5 ، 6 ، 7) اى تم قبول (5) مفردات واستبعاد (5) مفردات .

- الفرق بين الطريقتين :
- استخدمت طريقة مؤشر صدق المحتوى قيمة جدولية تتمتع بثقة تقدر (95%) لأنها دالة عند مستوى (0.05) ولم تقتصر على ذلك بل استخدمت متوسط المفردات لتوضح تقديرات المحكمين كدرجات

وبالتالي استطاعت أن تقدم معلومات ذات معنى من حيث المعامل ومن حيث المتوسط ولم يتم رفض سوى مفردة واحدة.

- استخدمت طريقة نسب الاتفاق محك واحد وهو نسبة (75%) وهذه النسبة مجرد قيمة ليس لها مدلول سوى أن أكثر من نصف المحكمين أشاروا أن المفردة ضرورية لكن هل هذا المحك يكون ثابت دائما مع اختلاف حجم المحكمين ؟ ورغم أنها ليس لديها فترة ثقة وليس لديها تفسير لاختلاف حجم عينة المحكمين فإنه تم رفض خمس مفردات قبلت منها الطريقة الأولى اربعة مفردات كما لا توجد قيمة لصدق محتوى الاختبار يمكن التنبؤ بها من خلال هذه الطريقة .
- لذا ينصح الباحث باستخدام طريقة مؤشر صدق المحتوى لأنها تعطى معامل واضح له فترة ثقة يختلف باختلاف العينة بالإضافة لمحك متوسط المفردات
- كما ينصح الباحث بتقنين طريقة نسب الاتفاق لترتبط بطريقة مؤشر صدق المحتوى لتكون لها فترة ثقة ومحك آخر يتمثل في متوسط المفردة وللاختبار ككل قيمة تمثل متوسط قيم الصدق للمفردات .

بالنسبة للإجابة عن السؤال الخامس والمتعلق بإمكانية تطوير مؤشر صدق المحتوى ليتضمن نسب الاتفاق المقابلة لحجم عينة المحكمين والتي تكون دالة عند مستوى(0.05) فقد قام الباحث بإعادة النظر في معادلة نسبة صدق المحتوى فوجد أنه يمكن صياغتها بطريقة أسهل تعطى نفس النتائج وهى

نسبة صدق المحتوى =(م −0.50)/0.50

حيث أن م= متوسط اتفاق السادة المحكمين حول المفردة بأنها ضرورية

0.50= قيمة ثابتة تمثل 50% من عدد المحكمين

وللتحقق من صدق المعادلة قام الباحث بتقدير المعامل باستخدام المعادلة الأصلية وتقدير المعامل باستخدام المعادلة المعدلة والجدول (6) يوضح ذلك

جدول (6) المقارنة بين المعادلة الأصلية والمعدلة

المفردات	البدائل			طريقة معامل صدق المحتوى		
	ضرورية	مفيدة	غير ضرورية	المعامل المعادلة الأصلية	المعامل المعادلة المقترحة	الحكم
1	9	1	0	0.80*	0.80*	مقبول
2	7	3	0	0.40	0.40	مقبول
3	9	1	0	0.80*	0.80*	مقبول

المفردات	البدائل			طريقة معامل صدق المحتوى		
	ضرورية	مفيدة	غير ضرورية	المعامل المعادلة الأصلية	المعامل المعادلة المقترحة	الحكم
4	7	2	1	0.40	0.40	مقبول
5	5	2	3	صفر	صفر	مرفوض
6	7	3	1	0.40	0.40	مقبول
7	2	7	0	0.60-	0.60-	مقبول
8	8	2	0	0.60	0.60	مقبول
9	8	2	0	0.60	0.60	مقبول
10	9	1	0	*0.80	*0.80	مرفوض
معامل صدق المحتوى للاختبار				0.42	0.42	

يتضح من الجدول (6) عدم وجود فروق بين تقديرات معامل صدق المحتوى بالمعادلة الأصلية وتقديرات معامل صدق المحتوى بالمعادلة المعدلة حيث أنها أعطت نفس القيم وبالتالي تصلح المعادلة المعدلة لحساب نسبة صدق المحتوى للمفردات .

ثم قام الباحث بالاطلاع على جدول قيم مؤشرات صدق المحتوى الجدولية واستخراج العدد الحرج من المحكمين المقابل لتلك القيم ثم قسمة العدد الحرج على العدد الكلى للمحكمين المقابل له في الجدول وضرب الناتج في مائة ليتم التوصل إلى النسب المئوية الحرجة الدالة عند مستوى (0.05) والمقابلة لحجم المحكمين والجدول (7) يوضح ذلك

جدول (7) عدد المحكمين والنسبة الحرجة

عدد المحكمين	أقل قيمة مقبولة لمعامل صدق المحتوى	النسبة الحرجة
5	0.99	99.8
6	0.99	99.8
7	0.99	99.9
8	0.75	87.5
9	0.78	88.9
10	0.62	81.0
11	0.59	79.7
12	0.56	78.3
13	0.54	76.9
14	0.51	75.0
15	0.49	74.7
20	0.42	71.0
25	0.37	68.8
30	0.33	66.7
35	0.31	65.7
40	0.29	64.8

يتضح من الجدول (7) عدد المحكمين وأقل قيمة لمعامل صدق المحتوى يقابلها النسبة المئوية الحرجة وبالتالي تم تقنين طريقة النسبة المئوية من خلال إعداد محكات دالة تناسب أحجام العينة المختلفة كما أنه من خلال ذلك يمكن استخراج صدق المحتوى للاختبار والذي يعبر عنه بأنه متوسط النسب المئوية لاتفاق المحكمين على المفردات والتي يجب ألا تقل عن النسبة الحرجة المقابلة لها بالجدول والمعادلة التالية تعبر عن ذلك نسبة صدق المحتوى للاختبار = مجموع نسب الاتفاق للمفردات ÷ عدد المفردات

وللإجابة عن السؤال السادس المتعلق بدراسة اختلاف صدق المحتوى بطريقة معامل صدق المحتوى باختلاف حجم العينة قام الباحث بتثبيت النسب المئوية ثم استخراج عدد المحكمين المتفقين على أنها ضرورية بافتراض أن البديل ثنائي (ضروري / غير ضروري) حيث تم استخراج العدد المقابل للنسبة المئوية في حالة (عدد المحكمين 10 محكمين , وعدد المحكمين 20 محكما) ثم حساب معامل صدق المحتوى للمفردات ومتوسط المفردات لكل مجموعة والجدول (8) يوضح ذلك

جدول (8) حجم عينة المحكمين ونسب الاتفاق ومعامل صدق المحتوى

النسبة	عدد10محكمين	عدد20محكما	معامل صدق المحتوى	م10	م20
0.10	1	2	0.80-	1	1.1
0.20	2	4	0.60-	1.2	1.2

النسبة	عدد10محكمين	عدد20محكما	معامل صدق المحتوى	م10	م20
0.30	3	6	0.40-	1.3	1.3
0.40	4	8	0.20-	1.4	1.4
0.50	5	10	0.00	1.5	1.5
0.60	6	12	0.20	1.6	1.6
0.70	7	14	0.40	1.7	1.7
0.80	8	16	0.60	1.8	1.8
0.90	9	18	0.80	1.9	1.9
1.00	10	20	1.00	2	2

يتضح من الجدول (8) ما يلي :

- أن قيم المعامل جاءت بالسالب عندما كان عدد المحكمين الذين أشاروا أن المفردة ضرورية أقل من نصف عدد المحكمين .
- أن القيمة المقابلة للنسبة(50%) جاءت تساوى صفر
- أن النسب فوق (50%) جاءت قيم المعاملات لها بالموجب
- يلاحظ تأثر قبول المفردات بالنسبة لمحك معاملات صدق المحتوى بحجم العينة فمثلا لعدد المحكمين (10) يكون محك معامل صدق المحتوى (0.62) وبالتالي تم رفض جميع القيم السالبة وقبول النسبتين (0.90, 100)
- في حين أنه في حالة حجم العينة (20) محكما فإن محك معامل صدق المحتوى يساوى (0.42) ومن ثم تم رفض جميع المفردات السالبة وقبول النسب (0.80 , 0.90, 100)
- وبالتالي اختلف قبول أو رفض المفردات طبقا لمحك المعامل باختلاف حجم العينة
- ولكن باستخراج متوسط المفردات لكل مجموعة يلاحظ عدم وجود اختلاف في متوسط المفردات يرجع لحجم العينة
- كما أنه بتطبيق المحكين معا لكل مجموعة وجد عدم اختلاف قبول أو رفض المفردات باختلاف حجم العينة وبالتالى ينصح الباحث باستخدام المحكين معا وعدم الاقتصار على محك معامل صدق المحتوى للمفردات لأنه يتأثر بحجم العينة فكلما زاد حجم العينة قلت القيمة الحرجة وبالتالي زادت عدد المفردات المقبولة

وللوقوف على نتائج واقعية غير مفترضة لجأ الباحث الى تطبيق الاستبانة على عدد (30) محكم صنفهم الباحث لمجموعتين (مجموعة 10 محكمين ومجموعة 20 محكم) ثم تم إجراء التحليلات على المجموعتين والجدول التالى يوضح ذلك

جدول (9) نسب الاتفاق ومؤشر صدق المحتوى ومتوسط المفردات

(تطبيق 10 محكمين مقابل 20 محكم)

المفردات	عدد (20) محكم			عدد(10) محكمين		
	نسب الاتفاق	المعامل	المتوسط	نسب الاتفاق	المعامل	المتوسط
1	**0.95**	0.9	1.95	**0.9**	**0.80***	**1.9**
2	**0.5**	0.5	1.7	**0.7**	**0.4**	**1.7**
3	**0.9**	0.8	1.9	**0.9**	***0.80**	**1.9**
4	**0.6**	0.2	1.6	**0.7**	**0.4**	**1.6**
5	**0.35**	0.3-	1.2	**0.5**	**صفر**	**1.2**
6	**0.9**	0.8	1.9	**0.7**	**0.4**	**1.7**
7	**0.2**	0.6-	1.1	**0.2**	**0.6-**	**1.1**
8	**0.9**	0.8	1.9	**0.8**	**0.6**	**1.8**
9	**0.7**	0.4	1.7	**0.8**	**0.6**	**1.8**
10	**0.85**	0.70	1.85	**0.9**	***0.80**	**1.9**
معامل صدق لمحتوى للاختبار		0.42			0.42	

يتضح من الجدول (9) ما يلى :

عدم وجود تأثر منتظم بشكل ملحوظ فى كل من نسب الاتفاق ومؤشر صدق المحتوى للمفردات و متوسط المفردات فى حال التطبيق على (20) محكم عن التطبيق على (10) محكمين وللتأكد استخدم الباحث اختبار مان ويتنى لايجاد الفروق بين التطبيقين لعينتين مستقلتين للمتغيرات الثلاثة نسب الاتفاق ومؤشر صدق المحتوى للمفردات ومتوسط المفردات كما يتضح من الجدول(10)

جدول (10) الفروق بين نسب الاتفاق ومؤشر صدق المحتوى ومتوسط المفردات فى

(تطبيق 10 محكمين مقابل 20 محكم)

	عدد المحكمين	عدد المفردات	متوسط الرتب	مجموع الرتب	Mann-Whitney U	الدلالة
نسب الاتفاق	20	10	10.6	106		
	10	10	10.4	104	49	غير دال
مؤشر صدق المحتوى	20	10	10.95	109.5	45.5	
	10	10	10.05	100.5		غير دال
متوسط المفردات	20	10	11	110	45	غير دال
	10	10	10	100		

يتضح من الجدول عدم وجود فروق دالة إحصائيا بين نسب الاتفاق فى التطبيقين (20 محكم و 10 محكمين) كذلك اتضح عدم وجود فرق بين مؤشرات صدق المحتوى فى التطبيقين (20 محكم و 10 محكمين) واظهرت النتائج عدم وجود فروق دالة بين متوسط المفردات فى التطبيقين (20 محكم و 10 محكمين) .

الاستنتاجات :

يتضح من نتائج البحث وجود مشاكل في ممارسات صدق المحتوى في البحث التربوي

تمثلت في ضعف إلمام الباحثين بمفهوم صدق المحتوى وخطوات تطبيقه فالباحث يركز على التحكيم باعتبار أنه هو الصدق ذاته مع أن التحكيم هو خطوة من الخطوات وتبدو ملامح عدم الإلمام بصدق المحتوى من خلال استخدام الباحثين لمصطلح صدق المحكمين وكأن المحكم هو الصادق أو غير ذلك حتى أن المعنى اللغوي الحرفي لمصطلح صدق المحكمين يشير إلى التحقق من صدق المحكمين أنفسهم مثله في ذلك مثل ثبات التصحيح مثلا ومن ثم فالاستخدام اللغوي والاصطلاحي غير صحيح والتقنية غير مستندة لوثائق علمية عند تطبيقها فهي ترجع لما يعرف بالعرف أو ما تعارف عليه الباحثين أو عادات وتقاليد البحث في التحقق من صدق المحتوى ومعاييره فعندما نسأل باحث لماذا اخترت طريقة نسب الاتفاق أو معيار (80%) مثلا فلا نجد إجابة سوى أن العرف لدى قسمهم العلمى نهج ذلك المنهج دون أن يقدم مستند علمي يدعم كلامه ولو أن الباحث رجع للمراجع العلمية لوجد أن ذلك العرف ليس صحيحا وأن العالم تخطاه منذ حقب من الزمن .

ولم يقتصر الأمر على المفهوم بل تعدى ذلك إلى سيطرة طريقة نسب الاتفاق على التحقق من صدق المحتوى رغم ما أشير إليه من عيوب سابقة الذكر إلا أن الباحث يجد سهولة في إجراء نسب الاتفاق كما أنها تعطى قيما متضخمة يظن الباحث أنها تعبر عن صدق المحتوى في حين أنها لا تدل على ذلك مطلقا، واستخدم الباحثون معايير كمية للحكم على نسب الاتفاق تلك المعايير ليس لها أساس علمي مذكور في البحوث التربوية أي أن الباحث لم يستند لشئ عن ذكر المعيار حيث لم يجد الباحث علاقة ارتباطية بين عدد المحكمين ومعيار الحكم .

وقد أشارت نتائج استطلاع رأى الخبراء لضعف الإلمام بصدق المحتوى وعدم وجود مفهوم واضح لمن هو المحكم ؟ وعدم دراية الباحث بالتعريف الإجرائي وكيفية إعداده وضعف تدريب الباحث على إعداد جداول المواصفات وعدم وجود معيار محدد للحكم كل ذلك جعل صدق المحتوى يجد قبول أقل من صدق البناء والصدق المرتبط بالمحك رغم أن صدق المحتوى كما يرى الباحث يعد أكثر أنواع الصدق واقعيا لأنه يحتك بالخبراء والمختصين الذين تعايشوا مع القضية وعاصروها وهو عامل بشرى لا يتوافر في باقي أنواع الصدق التي تعتمد على الأسلوب الرياضي البحت كمحدد للصدق مهملة الدور البشرى الذي يعايش المشكلة علاوة على أن الإحصاء قد تعطى نتائج تتعارض مع الواقع ونضطر لقبولها رغم عدم واقعيتها .

كما يتضح من نتائج البحث أهمية طريقة معامل صدق المحتوى كطريقة كمية تعبر عن تمثيل الاختبار للمحتوى من حيث المفردات والاختبار ككل وبالتالي تقدم نسبة صدق المحتوى للمفردات حيث توضح نسبة صدق المحتوى للمفردة بالنسبة للمفردات الأخرى وهذه الدرجة يتوافر لها محك ذو ثقة معترف بها علميا هذا المحك هو القيمة الجدولية المقابلة لحجم عينة المحكمين والدالة

عند مستوى (0.05) كما أنها تعطى قيمة كلية للاختبار توضح صدق محتوى الاختبار ككل تلك القيمة أيضا لها محك هو القيمة الجدولية ذات مستوى الدلالة (0.05) . Lai & Chang, (2007)

تفيد طريقة معامل صدق المحتوى في تقنين صدق المحتوى فبعد أن كان صدق المحتوى لا معامل له أصبح له معامل قابل للتفسير , علاوة على تقديم الطريقة لنمطين من التعامل مع البيانات النمط الصارم والذي يستخدم في حالة الاختبارات الطويلة والنمط غير الصارم في حالة الاختبارات القصيرة .لا تكتفي الطريقة باستخدام القيمة الجدولية عند مستوى(0.05) كمحك حيث اتضح من نتائج البحث أنها تتأثر بحجم عينة المحكمين حيث كلما زاد حجم عينة المحكمين قلت القيمة الحرجة وبالتالي زادت عدد المفردات المقبولة , لذا لجأت لمحك آخر وهو متوسط المفردة ويتمثل في متوسط تقديرات المحكمين للمفردة مما يؤدى لتلاشى تأثير حجم العينة وكما اتضح من نتائج البحث عدم تأثر قبول أو رفض المفردات بحجم العينة عندما تم استخدام المحكين معا . (Brinkman,2009)

وبالتالي أفادت النتائج في إيضاح ايجابيات طريقة معامل صدق المحتوى التي تميزها عن طريقة نسب الاتفاق فاتخاذ محك(75%) لقبول المفردات في طريقة نسب الاتفاق ليس له فترة ثقة ولا يعطى قيمة للاختبار ككل يمكن الحكم في ضوئها ولا يراعى خطأ تحيز المحكمين أو خطأ الصدفة ولذا لجأ الباحث لتقنين طريقة نسب الاتفاق وتطوير طريقة معامل صدق المحتوى بعد أن توصلت نتائج تحليل محكات نسب الاتفاق إلى أن حوالى (35%) من رسائل الدكتوراه استخدمت محكات غير موثوق فيها وغير صادقة وبالتالى فالنتائج التى توصلت لها تلك الرسائل أيضا غير موثوق فيها وتحتاج لإعادة نظر .

حيث توصل الباحث لمعادلة مبسطة لحساب معامل صدق المحتوى أسهل من المعادلة الأصلية في استخراج النتائج ولها ثابت محدد وتعطى نفس النتائج , كما توصل الباحث لاستخراج نسب الاتفاق الحرجة المقابلة لحجم العينة المناسب مما يعد توفيق بين الطريقتين ومساهمة في تقنين طريقة نسب الاتفاق حيث لا يحتاج الباحث لاستخراج معامل صدق المحتوى كل ما عليه هو استخراج نسب الاتفاق للمفردات ثم مقارنتها بنسب الاتفاق الجدولية عند حجم عينة المحكمين المستخدم هذه النسب الحرجة دالة عند مستوى (0.05) .

كما أفاد البحث في تقديم إمكانية استخراج نسبة الاتفاق على الاختبار ككل من خلال استخراج متوسط نسب اتفاق المحكمين بعد استبعاد المفردات غير الصادقة وهذا يتيح للباحثين مقارنة تلك النسبة بالقيمة الجدولية عند مستوى(0.05) والحكم على صدق الاختبار إذا بلغت القيمة المحسوبة القيمة الجدولية أو زادت عنها .

توصيات البحث :

ومن النتائج السابقة يوصى الباحث بإجراء مزيد من الدراسات حول طرق التحقق من صدق المحتوى من ناحية وإنتاج لبرامج لتنمية مفهوم صدق المحتوى وطرق تطبيقه لدى الباحث التربوى مما يساعد على تفادى الأخطاء التى يقع فيها الباحثون .فمن الضرورى طرح برامج لتوعية الباحثين وطلاب العلم والبحث العلمى بالإجراءات الحديثة المستخدمة لتفادى الأخطاء التى تنتج عن الاستخدام الخاطئ لصدق المحتوى وذلك لتنمية الناحية التطبيقية لدى هؤلاء الطلاب والباحثين فى التطبيق الصحيح لصدق المحتوى

كما توصى الدراسة بإجراء مزيد من البحوث المقترحة حول " تحليل صدق المحتوى باستخدام تقنيات حديثة " وإجراء مقارنات حول أفضلية تلك الطرق مع اقتراح حلول تفيد الباحثين فى الوقوف على آخر التطورات العلمية لتطبيق صدق المحتوى .

المراجع :

Allahyari, T., Rangi, N. H., Khosravi, Y., & Zayeri, F. (2011). Development and Evaluation of a New Questionnaire for Rating of Cognitive Failures at Work. International Journal of Occupational Hygiene, 3(1), 6–11.

Anastasi, A. (1988). *Psychological testing* (6th ed.).(New York: Macmillan)

Beckstead, J.W.(2009). Content validity is naught. Int. J. Nurs. Stud.;46:1274–1283. DOI:

http://dx.doi.org/10.1016/j.ijnurstu.2009.04.014.

Brinkman, W.-P.(2009). Design of a Questionnaire Instrument, Handbook of Mobile Technology Research Methods, ISBN 978-1-60692-767-0, pp. 31-57, Nova Publisher

Cronbach, L. J. (1971). Test validation.(In R. L. Thorndike (Ed.), *Educational measurement.* (2nd ed. pp. 443—507). Washington, DC: American Council on Education.)

Haynes, S. N. ; Richard,D.C.S; Kubany,E.S. (1995). Content Validity in Psychological Assessment: A Functional Approach to Concepts and Methods. *Psychological Assessment* ,Vol. 7, No. 3, 238-247

Haynes, S. N. & Waialae, K. (1994). Psychometric foundations of behavioral assessment.(In R. Fernández-Ballestros (Ed.), *Evaluacion Conductual Hoy* . Madrid: Ediciones Piramide.)

Karaiskos, D.; Kourouthanassis, P.E.; and Giaglis, G.M. (2009) Towards a

Validated Construct for Information Systems Pervasiveness: An Exploratory Assessment. In the *Proceedings of BLED 2009* pp.468-483

Lai, S. K. & Chang, C. M. (2007). The Application of Aiken Content Validity Coefficient and Homogeneity Reliability Coefficient in Enterprise Business Risks Scale for Hot-Spring Hotel Industries. *Journal of Management Practices and Principles*, 1(2), 64-75.

Lawshe, C. H. (1975). The quantitative approach to content validity.(*Personnel Psychology, 28*563—575.)

Messick, S. (1993). Validity.(In R. L. Linn (Ed), *Educational measurement* (2nd ed. pp. 13—104). Phoenix: American Council on Education and Oryx Press.)

Miller,J.,M.(2003). *Test Validation: A Literature Review.* On *http://plaza.ufl.edu/millerjm/validity_litreview_072203.pdf*

Prasad and Reghunath, (2011). "Evaluation of Safety Performance in a Construction Organization in India: A Study," ISRN Civil Engineering, vol. 2011, Article ID 276545, 6 pages, 2011. doi:10.5402/2011/276545

Shepard, L. A.: 1996, 'The centrality of test use and consequences for test validity', *Educational Measurement: Issues and Practice* ,16, pp. 5–24.

Sireci,(1998).The Construct of Content Validity. *Social Indicators Research*, 45: 83–117

Wynd, C.A., Schmidt, B., & Schaefer, M.A. (2003). Two quantitative approaches for estimating content validity. *Western Journal of Nursing Research*, 25, 508–518.

Vogt DS1, King DW, King LA.(2004). Focus groups in psychological assessment: enhancing content validity by consulting members of the target population. *Psychological Assessment* ,Sep;16(3):231-43.

Yaghmal,F. (2003). Content validity and its estimation, *Journal of Medical Education*,vol.3,N.1:25-27

The content validity in educational research: reality and development

By

Mohamed Ibrahim Mohamed

The current research aims to Study the reality of true practices for content validity in educational research throw analysis educational research

according to experts opinions and also throw the comparison between the content validity index and the proportion of agreement , in addition to develop tables index of the content validity index and recognize on the effect of size of judges on accept or reject items by using two ways which are the content validity index and the proportion of agreement and to make sure of that the researcher has analyzed number of (70) PhD thesis which have accepted in faculty of Education Minia university and he has applied search tool on number of (30) judges, the results has shown that there are problems in practicing of the content validity as a result of researchers weakness in knowing of content validity concept beside lack of the clear criteria to determine judges and criteria way on tools and satisfaction on the way of the proportion of agreement although it`s effected by the item of coincidence agreement and magnified estimates , the results have shown the importance of using the item of content validity as a quantitative way represents the acting of test for content which represents the content validity index for items and the content validity index as general in the level of significance (0.05) in addition to it`s using the item to decrease size simple of judgment also helps the researcher to reach simple equation to calculate the content validity index and develop tables for to the way of the proportion of agreement on the level of Significance .