فعالية أدوات بديلة لقياس مهارات ما وراء المعرفة في ضوء التصميم متعدد الطرق

الاستـــلام: 2/ سبتمبر /2023
التحكيم: 10/ سبتمبر /2023
القبـــــول: 20 / سبتمبر /2023

أ.د/محمود فتحي عكاشة [1،*]
أ. م. د/إيمان صلاح ضحا [2]

© 2023 University of Science and Technology, Aden, Yemen. This article can be distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

© 2023 جامعة العلوم والتكنولوجيا، المركز الرئيس عدن، اليمن. يمكن إعادة استخدام المادة المنشورة حسب رخصة مؤسسة المشاع الإبداعي شريطة الاستشهاد بالمؤلف والمجلة.

[1] أستاذ علم النفس التربوي- العميد الأسبق لكلية التربية - جامعة دمنهور. البريد الالكتروني: m.okasha@edu.dmu.edu.eg
[2] أستاذ علم النفس التربوي المساعد - كلية التربية- جامعة دمنهور. البريد الالكتروني: eman_daha@edu.dmu.edu.eg
* عنوان المراسلة: okasha_mahmod@edu.dmu.edu.eg

فعالية أدوات بديلة لقياس مهارات ما وراء المعرفة في ضوء التصميم متعدد الطرق

الملخص:

هدفت الدراسة للوصول إلى منهجية صالحة في قياس ما وراء المعرفة عبر أدوات بديلة جرى إعدادها والتحقق من صلاحيتها في ضوء التصميم متعدد الطرق الذي يجمع بين الطرق الكمية والكيفية في القياس.

أجريت دراسة استطلاعية، طبِّق فيها اختبار سلوك حل المشكلة المبرمج على عينة من طلبة الجامعة، وجرى جمع البيانات الكمية والكيفية كافة المتعلقة بطرق ومسارات حلهم، كما جرى جمع بروتوكولاتهم اللفظية أثناء قيامهم بالحل، وقد استفدنا من تلك البيانات في إعداد أدوات الدراسة التي قنِّنَت على عينة مكونة من (110) طالبًا وطالبة من طلبة كلية التربية في الفصل الدراسي الثاني للعام الجامعي(2013/2014) .

وقد توصلت الدراسة إلى وجود ارتباطٍ دال إحصائيًا بين درجات الطلبة على البروتوكولات اللفظية المتزامنة (TAP) ودرجاتهم على كل من الاختبار متعدد الاستجابات (MRT) والاستبيان المحدد بالمهمة (TSQ)، وقد كانت المهارات المقاسة بالاختبار متعدد الاستجابات الأكثر مساهمة في تفسير الأداء والتنبؤ به. كما كشفت الدراسة أيضًا أن درجات المعايرة البعدية كانت مؤشرًا جيدًا للمراقبة ما وراء المعرفية، واتسمت بالدقة لدى فئة الطلبة مرتفعي الأداء مقارنة بالمعايرة القبلية، كما كشفت البيانات التي جُمِعت من الآثار عن معلومات ذات قيمة حول مهارات الطلبة ونشاطهم في كل خطوة من خطوات الحل.

الكلمات المفتاحية: ***قياس مهارات ما وراء المعرفة، التصميم متعدد الطرق، البروتوكولات اللفظية، القياس المحدد بالمهمة.***

https://doi.org/10.20428/ijtd.v14i1.2309

The Effectiveness of Alternative Tools for Measuring Metacognitive Skills in Terms of the Multi Methods Design

Mahmod Okasha [1,*]
Eman Salah [*]

Abstract

The study aimed at reaching a valid methodology to measure metacognition through alternative tools constructed and validated in terms of multi-method design that combines between qualitative and quantitative methods in measurement.

A pilot study was conducted in which the computerized problem solving test was applied on sample of undergraduate students, the qualitative and quantitative data concerning their methods and paths they used to solve the problem in addition their thinking aloud protocols were collected, These data are used in preparing the study tools that applied on a sample consisted of (110) students from the faculty of education in(2013/2014).

The study revealed that there was a statistically significant correlation between students' scores in the concurrent methods (TAP) and the offline methods "MRT", "TSQ", The skills measured by the multiple response test were the most contributors to interpreting and predicting performance.

The study revealed also that students' post calibration scores was a good indicator of metacognitive monitoring, and it was more accurate than pre calibration for high performing students. Traces data presented valuable information about students' skills and activity at each step of the solution.

Keywords: *Metacognitive Skills Measurement, Multi Method Design , Verbal Protocol ,Task Specific Measurement.*

[1] Professor of Educational Psychology - Former Dean of the Faculty of Education - Damanhour University.
[2] Assistant Professor of Educational Psychology - Faculty of Education - Damanhour University.
(*) Corresponding Email: okasha_mahmod@edu.dmu.edu.eg

المقدمة والإطار النظري:

تعدُّ ما وراء المعرفة أحد الموضوعات الرئيسة في المنظور المعرفي للتعلم وحل المشكلة، حيث يعد وعي الفرد بعمليات تفكيره التي يقوم بها أثناء قيامه بفهم المشكلة أو المفهوم وقدرته على التحكم فيها وتوجيهها من أهم الجوانب في عملية التعلم وحل المشكلة، ولذلك فقد أولى كثير من الباحثين مهارات ما وراء المعرفة اهتماماً ملحوظًا، ووضعوها في أولويات أبحاثهم منذ فترة طويلة؛ وبالرغم من تلك الجهود المتزايدة والتي تؤكد على أهمية ما وراء المعرفة، فإن العمل على تطوير أدوات تقييمها لم يحظَ باهتمامٍ موازٍ وقد كان ذلك لكونها وعيًا أو عمليات داخلية ليست سلوكيات ظاهرة قابلة للملاحظة والقياس (Santiago,2008).

تواجه عملية قياس وتقييم مهارات ما وراء المعرفة صعوبات عديدة، فعند التصدي إلى مستوى الوعي المرتبط بالمهارة التي يتم تقييمها يفترض أن يكون لدى الأفراد القدرة على الوصول إلى العملية الحقيقية التي يجري تقييمها، وأنهم يمكنهم على الأقل القيام ببعض الارتباطات المسببة لسلوكهم أو مخرجات تعلمهم، ولكن في مجال التعلم الضمني يذكر كاسكالور (Cascallar, et al., 2006) أنه نتيجة لتنوع آليات نظامنا المعرفي توجد قدرة محدودة للوصول للمجال الكامل لنشاط المعالجة أو الروابط السببية الضمنية الكامنة وراء السلوكيات الظاهرة في أي وقت.

ورغم تلك الصعوبات فقد تنوعت طرق قياس ماوراء المعرفة، وقد صنّفها فينمان (Veenman,2005) في ضوء علاقتها الزمنية بالأداء إلى: الطرق المحتملة الاستباقية التي تكون قبل المهمة (Prospective Methods) والطرق الاسترجاعية التي تكون بعد المهمة (Retrospective Methods) ويطلق على تلك الطرق **(Offline Methods)** حيث إنها لا تستخدم أثناء الانهماك والانشغال في المهمة، ومن أمثلة تلك الطرق الاستبيانات والقوائم التي تطبّق قبل أو بعد إكمال المهمة أو حل المشكلة؛ بالإضافة إلى الطرق المتزامنة (Concurrent Methods) والتي تطبق أثناء الانهماك في المهمة ولذلك يطلق عليها **(Online Methods)** والتي من أمثلتها الملاحظات النظامية وبروتوكولات التفكير بصوت مسموع.

ومع تنوع طرق قياس مهارات ما وراء المعرفة فقد تعرضت الكثير منها إلى النقد الشديد وبخاصة النقد الموجه إلى مسألة صدق هذه الأدوات؛ فقد أشارت هادوين (Hadwin,et al.,2001) إلى أنه قد يقرر الطلبة في الاستبيانات المهارات والإستراتيجيات التي يعتقدون بفاعليتها ليس لأنهم استخدموها بالفعل، حيث إنهم يعدّلون معالجتهم الإستراتيجية لتلائم أغراضهم المختلفة، وبذلك تلعب المرغوبية الاجتماعية دورًا كبيرًا في تقارير الأفراد واختيارهم لبدائل الاستجابة، كما أشارت ألان (Allan,1997) إلى أن الفرد يختار الاستجابة التي تصف مدى استخدامه للمهارات والإستراتيجيات التي تصفها المفردات، مع أنه قد يكون استخدم فعليًا إستراتيجيات أخرى لم تتعرض لها تلك المفردات.

فضلاً عن أن تلك الاستبيانات تقيس المعالجة الإستراتيجية كاستعدادٍ أو كسمةٍ دائمة نسبيًا للفرد تتعلق بطريقة تعامله بأسلوب محدد في المواقف والمهام، وهناك أسباب عديدة لتثبيط الثقة في قدرة الطلبة على القيام بتعميمات دقيقة فيما يتعلق بمدى استخدامهم لإستراتيجيات محددة. (Winne, et al., 2002)

ويضيف ويني (Winne, et al., 2002) أن ما يتم تقريره في الاستبيان هو المهارات والعمليات التي طبقت بوعي، أما بالنسبة للمهارات غير المتاحة في الوعي فلا يتم تقريرها، علاوة على أنه قد تحدث بعض عمليات الذاكرة المتحيزة أو التعديلية إذا طبقت الاستبيانات بشكل متأخر عن المهام.

https://doi.org/10.20428/ijtd.v14i1.2309

كما أنه توجد مشكلة خطيرة تواجه مقاييس التقرير الذاتي ذات النمط المتكرر في الاستجابة أشار إليها فينمان (Veenman,et al.,2003) وهي أن التقارير الذاتية تعكس مفاهيم الطلبة للأنشطة التي يؤدونها، وبينما يجري تقرير هذه المفاهيم قد يختار الطلبة نقاطًا مرجعية متنوعة لمقارنة مفاهيمهم عن أدائهم (مثل مستواهم الفردي، وجهة نظر مزعومة لمدرسهم، أو المستوى القياسي المفترض للطالب المثالي، أو على العكس المستوى القبلي للطلبة الضعفاء)، لذلك فقد يتواجد لدى الطالب نمط إجابة ثابت لاختياره إحدى النقاط المرجعية أثناء التقرير.

بالإضافة إلى ذلك ففي الاستبيانات عادة ما يُكلف الفرد بإعطاء تقديرات تراكمية واسترجاعية لكيفية أدائه على المهام، وبذلك فعلى الفرد أن يتناول المعلومات من الذاكرة طويلة المدى، ومن غير المحتمل أن يكون لدى الفرد سجل محفوظ في الذاكرة طويلة المدى للأنشطة العقلية، وتبعًا لذلك انتهى سانتياجو (Santiago.,2008) إلى أن الاستبيانات تعاني من نقطة ضعف شديدة، وهي أنها لا تقيس المهارة التي تم استخدمَت بقدر ما تقيس تقرير عن المهارة التي تصفها المفردة، وبذلك فهي تقيس"إما قدرة الفرد على استرجاع المهارة في حالة المقاييس الاسترجاعية البعدية، أو قدرة الفرد علي التنبؤ باستخدام المهارة في حالة المقاييس القبلية".

وفي هذا الصدد قدّم عددٌ من الباحثين بيانات تشكك في صدق التقارير الذاتية في قياسها لمهارات ما وراء المعرفة، وأكّدوا أنه من أهم مهددات استخدام تلك الاستبيانات هو انخفاض معاملات صدقها مثل: (Baker & Cerro, 2000; Cromley & Azavedo, 2007; Hadwin,et al,2001; Richardson, 2004; Veenman, 2005). ؛ وهذا بدوره يشكك في النتائج التي تجمع بتلك الطرق والتي غالبًا ما يلجأ إليها الباحثون نتيجةً لانخفاض تكلفتها، وسهولة استخدامها، وسرعة الحصول على النتائج عبرها مقارنةً بالطرق الأخرى.

وقد كان البديل لدى الباحثين لتلافي العيوب التي تواجه الاستبيانات عند استخدامها في قياس مهارات ماوراء المعرفة استخدام الطرق المتزامنة التي تطبّق أثناء قيام الفرد بمهمة التعلم أو حل المشكلة، ومن أشهر تلك الطرق استخدام بروتوكولات التفكير بصوتٍ مسموع (Thinking aloud protocol) ؛ وقد اعتبرت تلك الطرق أدوات فعّالة للحصول على المعالجة الإستراتيجية المتزامنة، فقد كشفت دراسة كلٌ من (Ericsson&Simon.,1980; Desoete., 2007; Desoete., 2008; Santiago., 2008) صلاحية المقاييس المتزامنة في قياس ما وراء المعرفة لدى الأفراد وقدرتها على التنبؤ الجيد بمجموعات الأداء.

وعزز ماجليانو (Magliano,et al.,1999) صلاحية البروتوكولات اللفظية كنوافذ على المعالجة المتزامنة عن طريق شرح التقارب بين بيانات التفكير بصوتٍ مرتفع وبيانات الأداء، وأضاف فينمان (Veenman,2005) أن الطلبة لا يقررون بدقة الإستراتيجيات التي يستخدمونها استخداماً دائماً أثناء المهام أو التي انتهوا من استخدامها، ولكنهم يقررون بدقة الإستراتيجيات التي يستخدمونها حاليًا، وبذلك فبروتوكولات التفكير بصوتٍ مرتفع أو الاستبيانات التي تستخدم في الوقت الحقيقي والتي يجري فيها التقرير الذاتي للإستراتيجيات والمهارات التي تستخدم في ذلك الوقت لها درجة مقبولة من الصدق.

فضلاً عن ذلك، فقد كشفت دراسة (Bannert & Mengelkamp., 2008) أن طريقة التفكير بصوتٍ مسموعٍ تعدُّ طريقة تقييم متزامنة حساسة لما وراء المعرفة، حيث لم تتداخل مع عمليات الفرد ما وراء المعرفية، ولم تغير مستوياتها، كما أنها لم تحوّل الأداء على المهمة، الأمر الذي في حالة حدوثه يُعدُّ مهدداً أساسيًا يحول دون الاعتماد على نتائج القياس بتلك الطريقة، وأكّدت أن هذا يمكن أن يحدث فقط في حالة تضمن تعليمات القياس

قيام الأفراد بالتأمل وإعطاء التفسيرات الإضافية التي قد تتطلب وقتًا أطول للمعالجة، مما يغيِّر من شكل عمليات الفرد عن مثيلتها في حالة التفكير الصامت.

كما أكد فينمان (Veenman.,2007) أن المقاييس المتزامنة تتنبأ بشكل مرتفع بمخرجات التعلم، بينما لا تقوم بذلك المقاييس غير المتزامنة. كما توصل فينمان أيضًا (Veenman.,2011) إلى أن نقطة القوة الرئيسة في الطرق المتزامنة هي أن سلوك المتعلم الفعلي يكون وفقًا لمعيار محدد خارجيًا يجري إرساء قواعده خارج تباين الخطأ الناتج عن إدراكات المتعلم الذاتية.

ولم تخلُ تلك الطرق المتزامنة عند استخدامها في قياس مهارات ما وراء المعرفة من النقد فيذكر كل من إريكسون وسيمون (Ericsson & Simon., 1980) أنها تكشف فقط عن المعلومات المتاحة في الذاكرة العاملة، وتكون في متناول الوعي التي يمكن التعبير عنها، كما أنها من المحتمل أن تكون متأثرة بعوامل التواصل (Long & Bourg., 1996)؛ فضلاً عن أن تلك الطريقة تتطلب إجراءات معملية حساسة، كما تستغرق الكثير من الوقت في ضوء التعليمات التي تقدم للطلبة، وتسجيل الملاحظات، وتحليل درجات البروتوكولات، مما يجعلها غير مناسبة للعينات الكبيرة.

ونظرًا للمهددات المختلفة التي تواجه كل من الطرق المتزامنة وغير المتزامنة في قياسها لمهارات ما وراء المعرفة، كما أنه لم يُتح إلى الآن استخدام المقاييس السلوكية لرصد الاستجابات باستخدام تكنولوجيا المعلومات الحديثة باتِّباع وسائل آلية دقيقة، بما يمكِّن الباحث من المقارنة بين ما يقوله الشخص وما يفعله عند نقطةٍ معينةٍ من المهمة، إضافةً إلى مقارنة كل ذلك بحركة العين وأداء البدن ونشاط المخ أثناء التعامل مع المهمة، فقد أكّد فينمان (Veenman, et al., 2006) على أنه من الأفضل عدم الاعتماد على نوعٍ واحدٍ من المقاييس في قياس تلك العمليات العقلية العليا؛ وذلك لتجنب مصدر الخطأ في حالة استخدام أداةٍ واحدة، وأوصى بضرورة استخدام أداةٍ أخرى ثم المقارنة بين النتائج للحصول على المعلومات المطلوبة بدقة.

وبذلك فقد اقترح طرقًا منهجية تستخدم أكثر من طريقة في القياس بما يسمى "التصميم متعدد الطرق (Multi Method Design) يجري فيه استخدام أكثر من أداة في أوقاتٍ مختلفةٍ طبقًا للأداء على المهمة (Across-method and time design) والذي يجمع بين القياس القبلي والمتزامن أو المتزامن والبعدي أو الثلاث القياسات.

وقد تبنّى كثير من الباحثين أمثال (Desoete, 2007; Desoete, 2008; Santiago, 2008; Bannert & Mengelkamp, 2008; Van Haut – Wolters, 2009) هذه المنهجية في القياس باستخدام التصميمات متعددة الطرق التي من عبرها يمكن تلافي مصدر الخطأ في كل أداةٍ على حدة، ولكن في تلك الدراسات كافة كانت هناك ملاحظة مشتركة، حيث فوجئ الباحثون باختلاف نتائج القياسات المتزامنة وغير المتزامنة، حيث تراوحت معاملات الارتباط بين نتائج الاستبيانات والبروتوكولات في تلك الدراسات من (-0.7) إلى (0.42) وكانت جميعها غير دالة إحصائيًا، مما أدى إلى وقوعهم في شكٍ وحيرة، وغالبًا ما كانت تفسر تلك النتيجة بأن: الاستبيانات أقل كفاءة في الكشف عن الأنشطة والمهارات التي يقوم بها الفرد بالفعل أثناء تنفيذ مهمة أو حل مشكلة، وخاصة أن تلك الدراسات كشفت عن ارتباطات مرتفعة بين مقاييس التفكير بصوتٍ مرتفع والأداء على المهام بصوره المختلفة مقارنة بالاستبيانات بكثير. وبذلك فقد استنتج الكثير من الباحثين أن نتائج الاستبيانات مشكوك في صحّتها وصدقها فيما يتعلق بقياس مهارات ما وراء المعرفة حتى عندما كانت داخل تصميمات متعددة الطرق.

https://doi.org/10.20428/ijtd.v14i1.2309

ولكن تذكر تشيلنجس (Schellings,2011) أنهم بهذه الاستنتاجات يهملون أسبابًا أخرى قد تكمن وراء تلك الارتباطات المنخفضة بين مقاييس التفكير بصوتٍ مرتفع والتقارير الذاتية التي حددتها في سببين:

السبب الأول يتعلق بالإستراتيجيات والعمليات المقاسة: حيث إن طرق القياس المختلفة قد تهدف إلى قياس إستراتيجيات وعمليات مختلفة في محتواها، فقد تستهدف الاستبيانات قياس عمليات معينة، بينما تحلل بروتوكولات التفكير بصوتٍ مسموعٍ في ضوء عمليات مختلفة؛ فضلاً عن أنه في أغلب الدراسات التي استخدمت تصميم الطرق المتعددة كان يجري تقرير الارتباطات بين الدرجات الكلية في الوقت نفسه، فمعظم الأدوات تتكون من مقاييس فرعية متعددة.

السبب الثاني يتعلق بالمهمة أو المشكلة التي عليها يجري القياس: فغالبًا ما تستخدم الدراسات التي تلتزم بمنهجية الطرق المتعددة كثير من الاستبيانات والمهمات العامة التي تجعل الفرد يقرر أساليبه واستراتيجياته المعتادة أثناء المواقف، في حين أن بروتوكولات التفكير بصوتٍ مسموعٍ تستهدف مهمة بعينها، وغالبًا ما يسفر عنها تسجيل دقيق لكل المهارات والعمليات التي استخدمت آنذاك. وبذلك فمن المنطقي أن تكون نتائج تلك القياسات مختلفة؛ حيث إنه يوجد مهددان أساسيان لصدق بناء الاستبيان أشار إليهما ميسك (Messick, 1995) أحدهما: يتعلق بالتباين غير المرتبط بالبناء (The construct- irrelevant variance)، فالتقييم عن طريق هذا الاستبيان واسع جداً ويحتوي على معلومات غير مناسبة لتلك المهمة، ولذلك يجب أن يفهم المفحوصون معنى المفردات المتضمنة في الاستبيان قبل الاستجابة عليها، وأن يحدد المهام التي تتعلق بها هذه المفردات، وفي أغلب الأحوال لا يهتم الباحثون بذلك. والآخر: يتعلق بالتمثيل الناقص للبناء (Construct- under representation) وفيه يكون التقييم عن طريق هذا الاستبيان ضيق جداً قد لا يحتوي على كل العمليات والمهارات التي قد يقوم بها المفحوص أثناء المهمة؛ وبالتالي يفشل في بناء الجوانب الحرجة للوصول إلى الهدف المراد تقييمه، مثل تلك الحالة قد تحدث في البروتوكولات فقط عندما لا يتلفظ الفرد بكل الأفكار التي طرأت عليه أثناء أدائه للمهمة أو حله للمشكلة.

ومن ذلك يبدو أن بناء الاستبيان على مهمة محددة قد يقلل من التباين غير المناسب للبناء، كما أن وضع كل خصائص وعمليات المهمة في الاعتبار عند بناء الاستبيان قد يقلل التمثيل الناقص للبناء وقد يجعل كلاً من مؤشرات الاستبيان والبروتوكولات متقاربة في قياسها لتلك المهارات الخفية صعبة القياس.

وفي هذا الصدد أكدت هادوين (Hadwin, et al., 2001) فعالية تقارير الطلبة الذاتية محددة السياق عن التقارير الذاتية التي لا تعكس آثار هذا السياق، وأشارت إلى أنه إذا كان على الباحث استخدام منهجية التقرير الذاتي، فإن استخدام استبيانات مفصّلة على مهامٍ أو سياقاتٍ محددةٍ يكون الحل الأفضل، كما تشير جارفيلا (Jarvela,et al.,2008) أنّ تلك الاستبيانات التي تعتمد على مبادئ نمذجة الخبرة (Experience sampling method) والتي تهدف إلى جمع بيانات في وقتٍ مناسبٍ وسياقٍ حقيقي تعدُّ بمثابة مدخلٍ مبشّرٍ لدراسة الحالات والعمليات الداخلية التي يقوم بها الفرد، وذلك في مواقف حقيقية مما يزيد من صدق الأداة.

وقد قامت تشيلنجس (Schellings, 2011) بإجراء مقارنة عادلة بين الاستبيانات والبروتوكولات عن طريق بناء استبيان محدد بمهمة قراءة محددة (Task specific questionnaire) ومقارنته ببروتوكولات التفكير بصوتٍ مرتفع، وقد توصلت إلى ارتباطات مقبولة بينهما. وأكدت على أن الاستبيانات محددة المهمة يمكن أن تكون بدائل كافية ومعقولة لطرق التفكير بصوتٍ مرتفع المعملية شديدة الحساسية.

كما أكّد كل من براتن وساميلستين (Bråten & Samuelstuen, 2007) على أن الاستبيانات المحددة بالمهمة أكثر فائدة عن بروتوكولات التفكير بصوتٍ مرتفع، عندما تكون العينة كبيرة نسبيًا، كما أنها تتنبأ جيدًا بالأداء. وقد قاما بمقارنتها بالآثار المادية التي يتركها الطالب وهو بصدد قيامه بالمهمة "Traces"، وقد كشفت دراستهما أن تقارير الطلبة الذاتية المحددة بالمهمة قد تكون جديرة بالثقة تمامًا بشرط أن يكون للمفردات نظائر مادية ناشئة عن الأنشطة العلنية، وأحد أسباب ذلك قد يكون أن مفردات التقرير الذاتي في هذه الحالات تتصرف كإيماءات استرجاعية فعّالة تعزز ذاكرة المتعلمين لما فعلوه بالفعل. كما كشفت دراستهما أيضًا أن كلاً من التقارير الذاتية محددة المهمة والآثار الخاصة بها الموجودة في المهمة تتنبأ بالأداء على هذه المهمة والأداء وراء هذا السياق.

وقد أشار ويني (Winne,2006) في هذا الصدد إلى نقطةٍ مهمة، وهي أنه قد لا يعبر عن المعالجة الإستراتيجية بالضرورة علنيًا بطريقة تخلف وراءها آثارًا، لذلك فقد تشكل الآثار فقط عينة من عمليات المتعلمين، كما أنها قد تكون أقل دقة إذا لم يتم إنتاج نظائر مادية لمفرداتها ولذلك فقد نصح الباحثون بضرورة استخدام التقارير الذاتية المحددة المهمة مع البيانات المتاحة من خلال الآثار كلما أمكن ومقارنة ذلك بالإنجاز الفعلي، مع عدم إغفال طريقة التفكير بصوتٍ مرتفع التي يمكن تضمينها في التصميمات متعددة الطرق In:(Bråten & Samuelstuen, 2007) .

ومن ناحية أخرى فقد اقترح كل من كروملي وأزيفيدو (Cromley & Azevedo, 2011) مدخلاً آخر لقياس مستوى المهارات بشكل محدد بالمهمة؛ وذلك عبر الاستدلال على الإستراتيجية التي استخدمها الفرد في إتمام المهمة عن طريق طرح أسئلة اختيار من متعدد (Multiple Responses Test) حيث تشير كل استجابة من الاستجابات إلى استخدام إستراتيجية معينة تعكس مستوى محددًا من المهارة، وبذلك يُمكن تصنيف الأفراد على أساس مستوياتهم في المهارات المستخدمة بناءً على اختياراتهم لبدائل معينة دون غيرها، وقد اتبع سانتياجو (Santiago, 2008) طريقة مماثلة لذلك، حيث صنف المفحوصين إلى (مرتفعين – متوسطين – منخفضين) في استخدامهم لما وراء المعرفة طبقًا لإستراتيجيتهم المستخدمة في الحل، والتي جرى تحديدها قبليًا، مما يجعل قياس ما وراء المعرفة متزامنًا مع الأداء على المشكلة ولا يخضع لتحيزات الباحثين، كما أنه لا يدرك المفحوصون هذه العملية مما يجعل ذلك مريحًا وموضوعيًا؛ وربما تتلافى تلك الطريقة التي تتضمن المفردات متعددة الاستجابات المشكلة الخطيرة التي تواجهها مقاييس التقرير الذاتي ذات النمط المتكرر في الاستجابة والتي أشار إليها فينمان (Veenman,et al.,2003).

وإذا كانت الدراسة الحالية بصدد هذه القضية البحثية المهمة التي تتعلق بقياس مجموعة من أهم المهارات التي تسهم في أداء الفرد، وفي ظل قصور الدراسات العربية في تلك القضية، فقد هدفت الدراسة الحالية إلى الوصول إلى منهجيةٍ صادقةٍ تتوافر فيها المعايير السيكومترية المعتمدة في المجال لقياس مهارات ما وراء المعرفة بتطوير المقاييس المتاحة، وإعداد أدوات بديلة يجري تقليل مهددات صدقها إلى الحد الأدنى، وذلك عبر بناء كل من: الاستبيانات محددة المهام التي تتعلق بمهام محددة بدلاً من الاستبيانات العامة التي تتطلب وضع خصائص تلك المهام في الاعتبار وتحليل كل العمليات والمهارات التي يمكن استخدامها لإنجاز تلك المهام، والاختبار متعدد الاستجابات الذي يُبنى بعد تحديد الاحتمالات والمسارات كافة التي يمكن أن يتخذها الأفراد أثناء قيامهم

https://doi.org/10.20428/ijtd.v14i1.2309

بالمهمة، والتي يمكن منها تصنيف الأفراد وفقًا لمستوى المهارة المستخدمة، وبذلك يمكن معرفة مستوى المهارة ما وراء المعرفية لدى الفرد باختياره لإحدى الاستجابات دون الأخرى.

ويعتمد إعداد هذه الأدوات في الأساس على بيانات كيفية متنوعة متعلقة بخصائص المهام وخصائص الأفراد وطرقهم في الحل وحكمهم عليها بمجرد رؤيتهم لها، وكيفية تصرفهم عند الوقوع في الأخطاء أو مواجهة الصعوبات... وغيرها من البيانات التي يمكن جمعها بطرق كيفية تشتمل على البروتوكولات اللفظية والملاحظات والمقابلات، فضلاً عن البيانات التي يمكن جمعها عن طريق الآثار (Traces) والتي يمكن منها الاستدلال على المهارة التي استخدمها الفرد ومستواها.

وللتحقق من صدق تلك الأدوات البديلة في قياس مهارات ما وراء المعرفة يلزم مقارنتها بالمقاييس المتزامنة المتمثلة في البروتوكولات اللفظية بالاعتماد على التصميم متعدد الطرق.

وفي ضوء الاهتمام بتأثير الاختلاف بين المهام وبعضها، حيث تحدد طبيعة المهمة نوعية المهارات وانتشارها بين فئات الأفراد مختلفي الأداء، فقد هدفت الدراسة أيضًا إلى التحقق من قدرة كل أداة من الأدوات التي جرى إعدادها على التمييز بين مهارات ما وراء المعرفة النوعية المستخدمة في كل مشكلة على حدة بين الأفراد مرتفعي ومنخفضي الأداء على حل المشكلة.

وبذلك تتحدد مشكلة الدراسة في السؤالين الآتيين:

1- ما مدى صدق الأدوات التي جرى إعدادها في الدراسة الحالية في قياس مهارات ما وراء المعرفة؟
2- ما مدى قدرة أدوات الدراسة على تمييز مهارات ما وراء المعرفة النوعية المستخدمة في كل مشكلة على حدة بين الأفراد مرتفعي ومنخفضي الأداء على حل المشكلة؟

أهمية الدراسة:

1- الاستفادة من مقاييس التقييم البديلة ومن مميزات التقارير الذاتية ذات الفائدة العملية واسعة النطاق.
2- يمكن للباحثين والتربويين استخدام الأدوات التي جرى بناؤها في حال التحقق من صلاحيتها في قياس ما وراء المعرفة بشكل سليم وبطريقة تتلافى عيوب الأدوات الشائعة حاليًا.
3- استكمال القصور في الدراسات العربية حيث تندر الدراسات التي اهتمت بالقياسات المتعددة الطرق لمهارات ما وراء المعرفة أو بابتكار أدوات لقياسها بشكل سليم وموضوعي خصوصًا في مجال حل المشكلة.
4- فهم طبيعة العمليات والمسارات التي يقوم بها الأفراد مختلفي الأداء أثناء حل المشكلة.
5- إثراء النظرية الخاصة بإسهام ما وراء المعرفة في حل المشكلة.

مصطلحات الدراسة:

مهارات ما وراء المعرفة: Metacognitive skills "هي عمليات تحكم وظيفتها التخطيط والمراقبة والتقويم لأداء الفرد في حل المشكلة، وتعدُّ مهارات تنفيذية مهمتها توجيه وإدارة مهارات التفكير المختلفة العاملة في حل المشكلة" (Sternberg, 1992).

حل المشكلة Problem solving: " نشاط عقلي يهدف إلى التغلب على عائقٍ معينٍ لتحقيق هدف، أو الإجابة عن سؤالٍ بشرط ألا يستطيع الفرد استدعاء إجابةٍ فوريةٍ من الذاكرة"(Sternberg,1996).

التصميم متعدد الطرق Multi methods design: "أحد التصميمات البحثية يتم فيه استخدام أكثر من طريقة في القياس طبقًا للأداء على المهمة، بحيث تتضمن إحداهما القياس المتزامن المتمثل في (البروتوكولات اللفظية) وتتضمن الأخرى القياس غير المتزامن المتمثل في (الاستبيانات)، ثم المقارنة بين نتائج القياسات المختلفة للحصول على المعلومات المطلوبة بدقة (Veenman, 2005) .

الاستبيان المحدد بالمهمة Task specific questionnaire: "أداة تقرير ذاتي يتم بناء مفرداتها لقياس مهارات الأفراد ما وراء المعرفية التي قاموا بها أثناء إنجازهم لمهمةٍ محددة، أو حلهم لمشكلةٍ بعينها(Shellings, 2011) .

الطريقة:

تضمّن إعداد أدوات الدراسة خطوات عديدة:

1) إجراء دراسة استطلاعية (Pilot study) بهدف إعداد أدوات الدراسة، وتطلب إجراء تلك الدراسة خطوات عديدة:

1. إعداد (اختبار سلوك حل المشكلة) المكوَّن من (14) مشكلة عامة غير متعلقة بمحتوى دراسي معين بالرجوع إلى الأدبيات والدراسات السابقة ذات الصلة مثل: (الأعسر ،1998 ؛ البنا ،1996 ؛ جروان ،1999 ؛ زكي ،2007؛ عامر ،2002؛ عبد القوي 1993 ؛ عبد القوي ،2006؛ محفوظ ،1985؛ هيبة ،2007).
2. عرض الاختبار في صورته المبدئية على مجموعة من المحكمين المتخصصين في المجال لإبداء آرائهم فيه، ومجموعة من الخبراء المتخصصين في الحاسبات والبرمجة لبرمجة تلك المشكلات وتحويلها إلى مشكلات تفاعلية تحل عبر حركات الماوس ولوحة المفاتيح. وفي الوقت نفسه تجمع بروتوكولات الطلبة اللفظية أثناء الحل، علاوةً على ذلك تسجَّل كل حركات الماوس التي يقوم بها الطالب، بحيث يتيح لكل طالب بعد تطبيقه للأداة ملف باسمه يشتمل على ملفين، الأول عبارة عن ملف فيديو مسجل فيه كل المحاولات والحركات التي قام بها الطالب منذ لحظة فتحه البرنامج وكل تلفظاته التي تفوّه بها وقت حدوثها متزامنة مع محاولاته، والثاني ملف خاص بالنتائج التي حصل عليها في كل مشكلة متضمنة الدرجة والزمن المستغرق، وعدد الحركات التي قام بها ووسائل المساعدة التي استعان بها.
3. تطبيق (اختبار سلوك حل المشكلة) على عينة مكوَّنة من (60) طالبًا وطالبة من طلبة كلية التربية المقيدين بالفصل الدراسي الثاني للعام الدراسي (2013/2012م) وقد كان التطبيق فرديًا على بعض الطلبة، وجماعيًا على بعضهم الآخر، وقد قامت الباحثة بإجراء الملاحظات النظامية والمقابلات المتعمقة لجمع البيانات الكيفية المتعلقة بحل الطلبة وطرق تفكيرهم.
4. فرز ملفات الطلبة الخاصة بالنتائج لتصنيفهم إلى مستويات في سلوك حل المشكلة تمهيدًا للبحث عن مهارات ما وراء المعرفة ذات المستويات المختلفة في ملفات كل فئة وذلك في كل مشكلة على حدة؛ ولتحديد الطلبة ذوي المستويات المختلفة في سلوك حل المشكلة قامت الباحثة بالآتي:

- ترتيب جميع الطلبة (60) طالبًا تنازليًا على حسب الدرجة، ثم الزمن، ثم الحركات، بالترتيب في كل مشكلة على حدة، وقد كان هذا الترتيب لأن الطالب مكلف بحل المشكلة والتوصل

https://doi.org/10.20428/ijtd.v14i1.2309

إلى الحل الأخير الذي يجري بناءً عليه تقدير الدرجة وهو الذي يقرر إنهاءه لمحاولاته والذهاب إلى المشكلة التالية بالضغط على مفتاح (Next) وبذلك فهو غير مقيد في ذلك بزمن ولا بحركات، ولذلك فالمعيار الأول هو الدرجة، ثم الزمن، حيث إن الأهم من الحركات أن يتحكَّم الفرد في سرعته، ويدير وقته ليصل إلى الحل الصحيح، وينتقل إلى المشكلة التالية.

- تحديد مجموعة الطلبة المرتفعين في سلوك حل المشكلة، وهم أعلى (27٪) من الطلبة، وهم من حصلوا على أعلى درجات في أقل زمن وأقل عدد من الحركات، والطلبة المنخفضين في سلوك حل المشكلة، وهم أقل(27٪) من الطلبة، وهم من حصلوا على أقل درجات في أكبر زمن وأكبر عدد من الحركات. وما بين الفئتين هم فئة الطلبة المتوسطين في سلوك حل المشكلة.

5. حساب متوسط زمن ومعامل صعوبة كل مشكلة، واستبعاد المشكلات غير المناسبة (أربعة مشكلات).
6. تحليل ملفات الفيديو الخاصة بطلبة كل فئة على حدة؛ لفحص وتحديد المهارات النوعية التي استخدمها طلبة كل فئة في كل مرحلة من مراحل الحل، وتحديد مسارات تفكيرهم وتحديد الأخطاء الأكثر انتشارًا، فضلاً عن تحديد الطرق التي يكثر اتباعها لتصويب الأخطاء في كل فئة على حدة. وقد أسفر عن فحص ملفات الطلبة في العشر مشكلات التي حُدِّدت استبعاد مشكلتين لم يتضح فيهما فروق بين المداخل والطرق التي اتبعها الطلبة في المجموعات المختلفة لحل المشكلة، كما لم تكشف تلفظاتهم عن فروق نوعية في طرق تفكيرهم ومعالجتهم للمشكلة، *وبذلك فقد انتهت الباحثة إلى تحديد ثمان مشكلات تراوح معامل صعوبتها من (60٪) إلى(80٪) جرى إعداد أدوات الدراسة الحالية في ضوئها.*

2) إعداد أدوات الدراسة الأساسية (بطارية قياس سلوك حل المشكلة، ومهارات ما وراء المعرفة) في ضوء فحص ملفات الطلبة مختلفي الأداء الخاصة باختبار سلوك حل المشكلة وفي ضوء البيانات التي جُمِعت من الملاحظات والمقابلات" وتضمنت تلك الخطوة ما يلي:

1. إعداد مخطط ترميز تفصيلي(Coding-schema) لبروتوكولات التفكير بصوت مسموع الخاصة بالطلبة في كل مشكلة على حدة، يتضمن كل الأنشطة الموجبة والسالبة (التي تعبر عن مهارات التخطيط والمراقبة والتقويم) التي قام بها طلبة كل مجموعة أثناء حلهم للمشكلة والذي تضمن بناؤه خطوات عديدة، هي:

1. تحديد كل الأنشطة والعمليات التي تعكس مهارات ماوراء المعرفة المراد البحث عنها في كل مشكلة.
2. تحليل كل مشكلة على حدة وفق متطلبات أدائها ووصف أفضل الطرق لحلها، وتحديد تتابعات عمليات تجهيز المعلومات الممكنة للأداء الناجح عليها، مع تحديد البدائل الممكنة كافة.
3. تحديد خطوات وعمليات حل المشكلة التي يمكن القيام بها عبر مهارات ماوراء المعرفة المحددة.
4. الوصف المفصل لنوع العمليات والأفعال والأنشطة المتوقعة التي يمكن أن يقوم بها الطالب أثناء حل المشكلة والتي قد يقرها في البروتوكول، والتي توضع تحت مكوِّن معيَّن دون غيره، وفي هذه الخطوة يجري التفسير الموضوعي لعبارات البروتوكول إلى مكوِّنات بحيث يُستدل من العبارة أو النشاط الذي قام به الطالب على وجود المهارة.

2. تحليل بروتوكولات الطلبة في ضوء مخطط التشفير والتحقق من صدق وثبات التحليل:

حيث جرى فتح ملف الفيديو في الفقرة الخاصة بالمشكلة وتفريغ البروتوكول اللفظي (البروتوكول الخام) ثم تجزئة بياناته إلى عبارات ومقاطع صغيرة لوضعها تحت كل مكوِّن من مكوِّنات التحليل طبقًا لمخطط التشفير، بحيث تخصص درجة واحدة صحيحة لكل عبارة تنتمي إلى مكوِّن وبذلك جرى تحويل البروتوكول الخام (يحتوي على البيانات الخام) إلى بروتوكول مشفر يحتوي على معلومات يمكن التعامل معها إحصائيًا.

وللتحقق من صدق البروتوكولات اللفظية جرى تحييد مجموعة من العوامل والمقومات التي تؤثر على صدق البروتوكولات، والتي حددها روسو (Russo, et al., 1989) في شكلين: ***الأول هو التفاعلية (Reactivity)*** حيث يكون البروتوكول اللفظي تفاعليًا إذا غيرت التلفظات العملية الأساسية والذي قد ينتج عنه تغيير في مخرج هذه العملية أو كإطالة بسيطة في زمن الاستجابة (Response time)، وهناك عوامل عديدة تؤدي إلى التفاعلية عند الحصول على بروتوكولات لفظية جرى تحييدها، وهي:

- تأثير المجرب: فهو قادر على تغيير وربما تحويل تدفق أفكار الطلبة، ولذلك فقد حرصت الباحثة عند إعداد الأداة وبرمجتها أن تكون بعيدة تمامًا عن الطلبة أثناء قيامهم بالحل، كما أن الأداة مزودة بخاصية تنبيهية عند صمت الطالب أكثر من عشرين ثانية دون تدخل من قبل الباحثة.
- سعة الذاكرة العاملة المحددة التي تقيد عمليات التجهيز وعمليات التلفظ، حيث قد يؤدي تركيز المشارك على عمليات التلفظ إلى الحمل الزائد على الذاكرة العاملة، مما يعيق الأداء في المهام، الأمر الذي قد لا يحدث في حالة التفكير الصامت، ولذلك فقد أكدت الباحثة في التعليمات قبل بداية الاختبار أن يحرص الطالب على التفكير بصوتٍ مسموعٍ من دون أي مجهود أو عمليات وسيطة للتعبير، فليس من المطلوب أن يخطط الطالب لما سوف يقوله أو يحاول شرح أو تفسير ما يقوله وما يفعله، أو يحاول فرز ما يقوله وما لا يقوله، هذا النوع من التفكير بصوتٍ مسموعٍ ينتمي إلى المستوى الثاني وفقًا لأريكسون وسيمون (Ericsson & Simon, 1980) والذي لا يؤثر في بناء العمليات المعرفية لأداء المهمة.
- الظروف غير العادية مثل الضوضاء أو التوتر الناتج عند الالتزام بمسؤوليات أخرى، ولذلك فقد روعي قبل تطبيق الأداة توفير المكان والتوقيت المناسب للطلبة، ولم يكن تطبيق الأداة تعسفيًا على كل الطلبة في الشعبة في نفس التوقيت، بل مناسبًا لمواعيدهم ولظروفهم الشخصية.

أما الشكل الثاني لعدم صدق البروتوكول هو عدم المصداقية (Non veridicality) حيث يكون البروتوكول غير صادق إذا لم يعكس بدقة العملية الأولية الكامنة، ويتضمن ذلك أخطاء السهو والنسيان (أي عدم تقرير أحداث عقلية حدثت) وأخطاء التزييف (أي تقرير أحداث عقلية لم تحدث).

ومن أهم العوامل التي تؤدي إلى عدم مصداقية البروتوكول الفاصل الزمني بين انهماك الفرد في المشكلة والحصول على البروتوكول، حيث إن البروتوكول الاسترجاعي يكون أكثر عرضة لعدم المصداقية، ونظرًا لأن البروتوكول المستخدم في الدراسة الحالية هو بروتوكول متزامن فهو أكثر مصداقية؛ لأن عمليات التجهيز وعمليات التلفظ تحدثان في نفس الوقت، فعمليات التلفظ تعكس بشكل مباشر ما يقوم به الفرد من عمليات تجهيز.

https://doi.org/10.20428/ijtd.v14i1.2309

وللتأكد من صدق تحليل البروتوكولات اللفظية جرى حساب معامل الارتباط بين درجات مهارات ماوراء المعرفة المقدرة عبر عملية تحليل البروتوكولات ودرجات الطلبة على سلوك حل المشكلة (صدق المحك الخارجي) حيث ينبغي أن تتصرف أداة التقييم كما توقعت النظرية طبقًا لعلاقتها بالمتغيرات الأخرى وتنبؤها بمخرجات التعلم (Messick, 1995; Veenman, 2007; Veenman, 2011) وقد بلغ معامل الارتباط (0.76)، مما يدل على وجود ارتباطٍ قوي بين درجات الطلبة على مهارات ماوراء المعرفة التي أسفرت عنها عملية تحليل البروتوكولات وبين درجاتهم على سلوك حل المشكلة مما يدل على تحقق الافتراضات النظرية التي انطلقت منها الباحثة، مما يؤكد وجود صدق نظري في عملية تحليل البروتوكولات.

بالإضافة إلى ذلك جرى التحقق من ثبات تحليل البروتوكولات اللفظية في ضوء مخطط التشفير بحساب (ثبات المصححين)، حيث جرى تفريغ ملف الفيديو الخاص بأحد الطلبة وكتابته ونسخه ثلاث صور، وإعطائه لثلاثة مصححين آخرين من ذوي الخبرة لكي يقوموا بتصحيحها وفقًا للإجراءات والأسس المتبعة من قبل الباحثة لكي يقوموا بتقدير الدرجة التي يستحقها الطالب على كل مهارة ماوراء معرفية وفقًا للمعايير والأسس المحددة سلفًا. ثم حساب ثبات التحليل في كل بروتوكول على كل مشكلة من الثمان المشكلات، كل على حدة، وذلك بواسطة نسب الاتفاق ومعامل الثبات المركب بين المحللين، وقد امتدت قيم معاملات الثبات المركب بين المحللين ما بين(0.838) إلى (0.924) وهي قيم مرتفعة، مما يدل على أن تحليل البروتوكولات يجري في ضوء قواعد وأسس ثابتة.

3. إعداد الاستبيان المحدد بالمهمة: Task specific questionnaire

جرى إعداد استبانة تقرير ذاتي بحيث تتضمن مفرداتها مهارات ماوراء المعرفة المستخدمة في مهام بعينها؛ وجرى صياغة مفرداتها المتعلقة بكل مشكلة بالاستعانة بالتصنيف الموجود في ترميزات مهارات ماوراء المعرفة التي شُفِّرَت في مخطط التشفير، ولأن مخطط التشفير يعد مخططًا تفصيليًا يتضمن كل الأنشطة الموجبة والسالبة فقد روعي أن تجمع مفردات الاستبيان بين الموجب والسالب وذلك في كل مشكلة. وقد وصل عدد مفردات الاستبيان إلى (48) مفردة، بواقع ست مفردات لكل مشكلة مفصّلة على الثلاث المهارات المراد قياسها، وهي التخطيط، والمراقبة، والتقويم، بحيث يكون لكل مهارة مفردتين أحدهما موجبة والأخرى سالبة، وكان مقياس الإجابة مكونًا من بديلين: نعم أو لا، تتراوح الدرجة على الاستبيان من (صفر إلى 48).

4. إعداد اختبار متعدد الاستجابات Multiple Responses Test (MRT) لقياس مستوى المهارات ماوراء المعرفية الكامنة، وذلك عبر الخطوات الآتية:

1. تحليل كل مشكلة وفق متطلبات أدائها، ووصف أفضل الطرق لحلها، وتحديد تتابعات عمليات تجهيز المعلومات الممكنة للأداء الناجح عليها بفحص ملفات كل فئة من فئات الطلبة ذوي المستويات المختلفة.
2. صياغة مفردات الاختبار وبدائل الاستجابات، بحيث تعكس مسارات الأفراد في مراحل الحل كافة، تتكون كل مفردة من جذع يُسأل فيه الطالب عن إجراء أو خطوة قام بها وهو بصدد حله للمشكلة، سواء أكان هذا الإجراء في المرحلة الأولى للحل منذ بداية عرض المشكلة أو أثناء محاولاته أو في المرحلة الأخيرة للحل قبل انتقاله إلى الشاشة التالية، *وبدائل الاستجابات* التي جرى صياغتها بحيث تعكس مسارات الأفراد بعد هذا

الإجراء، والتي تعكس مستويات مختلفة من المهارة الكامنة المراد الاستدلال عليها عبر تلك الإجراءات. جرى صياغة أربعة بدائل على كل مفردة جميعها صحيحة تمثل درجات مختلفة من المهارة المراد قياسها، بحيث يدل أحد البدائل على المستوى المرتفع من المهارة، والآخر على المستوى المنخفض، والبديلين الباقيين على المستوى المتوسط منها، بحيث يحصل الطالب على ثلاث درجات إذا اختار البديل الذي يعبر عن المستوى المرتفع، ودرجتين إذا اختار أحد البديلين اللذين يعبران عن المستوى المتوسط، ودرجة واحدة إذا اختار البديل الذي يعبر عن المستوى المنخفض فيما عدا المفردة رقم (23) فقد كان عليها بديلان على المستوى المرتفع، وبديل واحد على المستوى المتوسط، وبديل واحد على المستوى المنخفض، مع مراعاة أن تعرض المفردات وفقًا لاختيارات الفرد حيث تتغير المفردات التالية على حسب استجابات الفرد على المفردات السابقة، ولكن عددها ثابت مع الجميع.

كما أنه في بعض المفردات جرى تخيير الفرد بالإجابة عليها بحيث تكون الدرجة على السؤال متوقفة على الدرجة الفعلية التي حصل عليها الفرد في المشكلة، ومن أمثلة هذه المفردات المفردة رقم (17) الخاصة بمهارة المراقبة[*] في مشكلة (السودوكو)، وهي:

أجب عن هذا السؤال فقط إذا وقعت في خطأ أثناء حلك ... عندما وقعت في خطأ فإنك ...

أ- عرفت أنه يوجد خطأ وبحثت عنه قبل أن أنتقل للخطوة التالية.

ب- عرفت أنه يوجد خطأ واستكملت خطواتي إلى أن يظهر لي.

ج- عرفت أنه يوجد خطأ ولذلك أعدت المشكلة من البداية خوفًا من ضياع الوقت.

د- لم أعرف أنني وقعت في خطأ إلا في الخطوات الأخيرة.

إذا اختار الفرد البديل (أ) يحصل على الدرجة (3) لأنه يعبر عن المستوى المرتفع، وإذا اختار (ب) أو (ج) يأخذ الدرجة (2) لأنه يعبر عن المستوى المتوسط، وإذا اختار البديل (د) يأخذ الدرجة (1) لأنه يعبر عن المستوى المنخفض. أما في حالة عدم إجابة الفرد تتوقف درجته على هذا السؤال وفقًا للدرجة الفعلية التي أحرزها على المشكلة، بحيث إذا لم يجيب وأخذ الدرجة العظمى على المشكلة فيحصل على الدرجة (3) التي تدل على المستوى المرتفع من المهارة، أما إذا لم يأخذ الدرجة العظمى فيحصل على الدرجة (1) والتي تدل على المستوى المنخفض من المهارة.

3. بلغ عدد مفردات الاختبار (59) مفردة، تتراوح الدرجة عليه من(59 -177) روعي فيها الوزن النسبي للمفردات الخاصة بكل مهارة مع ملاحظة أن الإجابة عن مفردات الاختبار غير مقيدة بزمن، حيث لن تغلق الشاشة على الطالب إلا بعد اختيار البديل الذي يناسبه.

5. برمجة الأدوات التي أعِدَّت، والتي اشتملت على خمسة عناصر، هي:

برمجة الأدوات التي أعِدَّت على شكل بطارية أدوات لقياس "سلوك حل المشكلة ومهارات ما وراء المعرفة" بحيث تعرض على الطالب المشكلات المكلف بحلها المتضمنة في اختبار سلوك حل المشكلة، كل مشكلة في نافذة منفصلة محددة بالزمن، وبعد انتهائه منها يعرض عليه مفردات الاختبار متعدد الاستجابات الخاصة بتلك

[*] أصبحت هذه المفردة تحت بُعد التقويم بعد التحكيم.

https://doi.org/10.20428/ijtd.v14i1.2309

المشكلة، ثم مفردات الاستبيان المحدد بالمهمة الخاصة بتلك المشكلة. وقد أضِيفَ بند آخر يتعلق بأحكام الفرد الخاصة بقدرته على حل المشكلة بمجرد رؤيتها، وأحكام الفرد الخاصة بثقته في الحل الذي توصل إليه بعد الانتهاء منها كمؤشر للمراقبة ما وراء المعرفة، وقد كان ذلك بناءً على كثير من البيانات التي جُمِعت عند فحص ملفات الطلبة. حيث كثرت تعليقات الطلبة الخاصة بقدرتهم على الحل وتقديراتهم لما توصلوا إليه، لذلك أضِيفَ هذا التقييم بشكل منفصل في كل مشكلة لقياس ثقتهم في قدرتهم على الحل، وثقتهم في ما توصلوا إليه من حل، وهذا التقييم يسمى المعايرة (Calibration) بصورتيها القبلية والبعدية.

ويقصد بمقياس المعايرة هنا: مدى التطابق بين تقدير الطلبة لقدراتهم أو لأدائهم (أحكام ما وراء المعرفة) وأدائهم الفعلي (المهمة المعيارية). وقد جرى حساب المعايرة القبلية والبعدية بواسطة مقياس التحيز (Bias) والذي يمثل متوسط الفرق بين الدرجة المتوقعة في التنبؤات القبلية أو الدرجة المتوقعة في التنبؤات البعدية والدرجة الفعلية للأداء (Mengel Kamp & Bannert, 2010). ويتيح مقياس التحيز للانحراف بين معدل الثقة ودرجة الأداء أن يكون موجباً أو سالباً، وبالتالي فهو يقدم معلومات عن اتجاه ومقدار نقص المطابقة بين الحكم والأداء، واتجاه الفرق يعطي معلومات عن الثقة العالية مقابل الثقة المنخفضة، في حين يقدم مقدار الفرق الذي يبدأ من الصفر معلومات عن مقدار الخطأ في الحكم. وكلما كانت أحكام الفرد على أدائه تتطابق مع أدائه الفعلي كان مؤشر التحيز صفرًا دل ذلك على أن الفرد لديه معايرة أفضل، أما إذا كانت أحكامه على أدائه أكبر من الدرجة الفعلية التي حصل عليها كانت درجة التحيز موجبة، ويشير ذلك إلى الإفراط في الثقة (Over Confidence)، بينما إذا كانت أحكام الفرد على أدائه أقل من الدرجة الفعلية التي يحصل عليها كانت درجة التحيز سالبة، ويشير ذلك إلى النقص في الثقة (Under Confidence) (Baker,2010).

وقد أضِيفَ مقياس المعايرة بالاستعانة بالخبراء المبرمجين، حيث قاموا بإضافة بند بعد عرض المشكلة مباشرة يشتمل على مسطرة متدرجة حدها الأدنى صفر، وحدها الأعلى الدرجة العظمى على المشكلة (مع مراعاة أن المشكلات تختلف في درجاتها العظمى) وعلى هذه المسطرة مؤشر يقبل الحركة، حيث يكلف الفرد قبل أن يبدأ محاولاته في حل المشكلة أن يحدد لنفسه الدرجة التي يتوقع أن يحرزها على هذه المشكلة، وهذه هي الدرجة الأولى. ثم يبدأ الفرد بالحل ويستمر في محاولاته إلى أن ينتهي، وينتج عن ذلك الدرجة التي قدّرها البرنامج لحله، وهذه هي الدرجة الثانية، ثم بعد ذلك تعرض عليه نفس المسطرة المتدرجة، حيث يكلف الفرد بأن يحدد الدرجة التي يقدرها لحله الذي توصل إليه، وهذه هي الدرجة الثالثة. ومن هذه الدرجات يقوم البرنامج بحساب درجات المعايرة بشكل تلقائي، بحيث تحسب درجة المعايرة القبلية لكل مشكلة (Prediction) من: الدرجة التي توقعها الفرد لنفسه قبل محاولاته (الدرجة الأولى) – الدرجة الفعلية التي حصل عليها الفرد بتقدير البرنامج (الدرجة الثانية). وتحسب درجة المعايرة البعدية لكل مشكلة (Postdiction) من الدرجة التي قدرها الفرد لحله، أي درجة معايرة الأداء (الدرجة الثالثة) – الدرجة الفعلية التي حصل عليها الفرد (الدرجة الثانية).

رصد الآثار الدالة على ارتفاع وانخفاض مهارات الفرد عبر تحركاته منذ بدايتها إلى نهايتها:

ولأن كل مشكلة من الثمان المشكلات المكوّنة للاختبار يتطلب حلها قيام الفرد بعددٍ معيَّنٍ من الحركات أو الخطوات الصحيحة والتي يختلف حدها الأدنى في مشكلة عن الأخرى، ولكي يقوم الفرد بحل

المشكلة فإنه يقوم بعدد من الحركات والخطوات التي قد تتجاوز في معظم الأحيان العدد المطلوب (الحد الأدنى) للحركات، فالمهارات الثلاث يمكن أن تأخذ الشكل الموجب أو السالب أثناء محاولات الفرد وحركاته.

ولأن الحركات الصحيحة المطلوبة لحل المشكلة سُجِّلت في البرنامج خطوةً بخطوة، فكانت هناك إمكانية لتسجيل كل حركة فعلية يقوم بها الفرد في اتجاهه للحل ومقارنتها بالحركات الصحيحة المسجلة في كل مشكلة، بحيث توضع الإشارات الخاصة بالمهارات الثلاث في كل تقاطع بين حركة المفحوص وأقرب حركة صحيحة قام بها.

وبذلك فكل حركة يقوم بها الفرد يمكن للبرنامج أن يقوم بتحديد موقعه على متصل الخطوات الصحيحة، ثم توضع الإشارات على هذا التقاطع بين حركته وأقرب حركة صحيحة، بحيث يمكن من ذلك التعرف على مهارات الفرد وتغيراتها في مراحل الحل كافة، وتحديد المراحل والخطوات التي كانت فيها هذه المهارات نشطة وفعالة، والمراحل التي كانت فيها غير نشطة.

ثم جرى توضيح المطلوب للخبراء المبرمجين مع تفصيل لكل الحركات التي توضع فيها الإشارات الموجبة والسالبة في كل مهارة بحيث ينتج عن هذا الإجراء منحنى لمسار الحل في كل مشكلة موضح به الإشارات الموجبة والسالبة لكل مهارة.

6. الحصول على النسخة المبرمجة من الأدوات التي أعِدَّت:

المكونة من ثمان مشكلات جرى حوسبتها بحيث ينتج لكل طالب ملف باسمه يشتمل على ملفين: *أحدهما* خاص بالنتائج، متضمنة درجات الطالب على كل مشكلة على حدة، ثم درجاته على الاختبار متعدد الاستجابات، ثم درجاته على الاستبيان محدد بالمهمة، ودرجات المعايرة القبلية والبعدية، ثم منحنى يوضح حالات المهارات التي قام بها الفرد سواء أكانت موجبة أو سالبة في كل حركة قام بها الفرد لحل المشكلة مقارنة بالحركات الصحيحة التي ينبغي القيام بها لحل المشكلة المسجلة في البرنامج، *والآخر* ملف فيديو مسجل عليه كل محاولات وتلفظات الفرد أثناء الحل، حيث كان موضح في التعليمات أن التفكير بصوتٍ مسموعٍ شرط أساسي لإتمام هذه الأدوات.

الدراسة الأساسية:

تطبيق أدوات الدراسة "البطارية" على العينة الأساسية:

قامت الباحثة بتطبيق أدوات الدراسة في صورتها النهائية على عينة الدراسة الأساسية التي تكوَّنت من (110) طالبًا وطالبة من طلبة كلية التربية المقيدين بالفرقة الثانية والثالثة في الفصل الدراسي الثاني لعام (2014/2013)؛ وقد كان التطبيق جماعيًا على معظم الطلبة (70) طالبًا وطالبة، وفرديًا على بعضهم الآخر (40) طالبًا وطالبة.

لوحظ أثناء جلسات التطبيق الأولى أن أغلب الطلبة يشتكون من طول الأدوات وكثرة عدد المفردات التي تتطلب المزيد من التركيز الذي يستغرق الكثير من الوقت، لذلك فقد قامت الباحثة بالتنبيه على الطلبة بحل المشكلات

https://doi.org/10.20428/ijtd.v14i1.2309

الست الأولى باتباع التعليمات المطلوبة في كل مشكلة، وتركت لهم الحرية في حل المشكلتين الأخيرتين على حسب تركيزهم واستعدادهم؛ حتى لا يكونوا مضطرين لاختيار أية بدائل دون تركيز.

جرى جمع ملفات الطلبة واستخراج البيانات الكمية والكيفية لكل طالب، وحصر الطلبة الذين أتموا الثمان المشكلات بمتطلباتها كافة، وقد كان عددهم عشرين طالبًا وطالبة من إجمالي (110) طالبًا وطالبة أتموا جميعًا الست المشكلات الأولى، لذلك جرى استبعادهم من الاختبار ببنودهم.

جرى تحليل بروتوكولات الطلبة المتزامنة على كل مشكلة وفقًا لمخطط التشفير، وذلك بفحص ملفات الفيديو، ورصد كل البيانات الخاصة بكل مشكلة (الدرجة، الزمن،الحركات، بيانات خاصة بالاختبار، بيانات خاصة بالاستبيان، بيانات خاصة بالمعايرة، بيانات خاصة بشكل المنحنى الممثل من حركات الطالب أثناء الحل). ثم رصد كل البيانات على البرنامج الإحصائي (SPSS) لاستخدامه في معالجة البيانات وتحليل النتائج.

التحقق من صلاحية أدوات الدراسة "البطارية"للاستخدام:

(1) صلاحية اختبار سلوك حل المشكلة للاستخدام: جرى التحقق من صلاحية المشكلات الست للاستخدام بحساب معامل الثبات بطريقة (ألفاكرونباخ) وقد امتدت قيم معاملات الثبات للمشكلات الست من (0.935) إلى (0.955) وهي قيم مرتفعة، وكانت قيمة معامل الثبات الكلي (0.957) مما يدل على ثبات الاختبار ككل. وحساب الاتساق الداخلي للاختبار: حيث كانت معاملات الارتباط بين درجة كل مشكلة والدرجة الكلية مرتفعة ودالة إحصائيًا عند مستوى (0.01) مما يدل على وجود اتساق داخلي للاختبار مما يعكس صدق بنائه. بالإضافة إلى حساب القدرة التمييزية للمشكلات الست بين الأفراد الحاصلين على درجات مرتفعة والحاصلين على درجات منخفضة، وكانت قيم جميع معاملات التمييز مرتفعة وموجبة تتراوح بين (0.888) و (0.949) مما يدل على ارتفاع القدرة التمييزية للمشكلات الست مما يعكس صدقها.

(2) صلاحية الاختبار متعدد الاستجابات (MRT) للاستخدام:

بحذف المفردات الخاصة بالمشكلتين الأخيرتين أصبح عدد مفردات الاختبار (45) خمسًا وأربعين مفردة، موزعة على المشكلات الست، جرى التحقق من صلاحية الاختبار متعدد الاستجابات (MRT) للاستخدام بحساب:

[أ] معامل ألفا كرونباخ (Cronbach's Alpha) لمفردات كل مهارة على حدة (التخطيط، المراقبة، التقويم) بعدد مفردات كل مهارة وفي كل مرة تحذف درجات إحدى المفردات من الدرجة الكلية للمهارة التي تنتمي إليها ومقارنة القيم الناتجة بمعامل ألفا الكلي للمهارة بدون حذف أي مفردة (Cronbach's Alpha if item deleted)

واتضح من هذه الخطوة وجود إحدى عشرة مفردة يستلزم حذفها، حيث كان معامل ألفا في "حالة حذف المفردة" أكبر من معامل ألفا الكلي للمحور مما يدلّ ذلك على عدم ثبات المفردة، حيث إن ذلك يعني أن حذفها أدى إلى رفع معامل الثبات للمحور، وأن تدخلها يؤدي إلى خفض معامل الثبات للمحور". ثم جرى حساب معاملات الثبات للمفردات المتبقية وعددها (34) مفردة بنفس الطريقة، واتضح أن قيم معاملات الثبات لكل مفردة أقل من معامل الثبات للمحور الذي تنتمي إليه المفردة، أي أن جميع المفردات ثابتة، كما كانت معاملات الثبات للمحاور الثلاث (التخطيط،

المراقبة، التقويم) على التوالي (0.880، 0.888، 0.870) وهي قيم مرتفعة، مما يؤكد ثبات الاختبار بعد حذف المفردات والاطمئنان لنتائجه.

[ب] حساب القدرة التمييزية لكل مفردة من مفردات الاختبار.

بعد حذف المفردات غير المناسبة في الخطوة السابقة جرى حساب معاملات التمييز للمفردات المتبقية وعددها (34) مفردة، وقد امتدت قيم معاملات التمييز للمفردات التي تنتمي لمحور التخطيط بين (0.364، 0.654) وامتدت بين (0.589، 0.744) للمفردات التي تنتمي لمحور المراقبة، وبين (0.534، 0.984) للمفردات التي تنتمي لمحور التقويم، وكلها قيم موجبة ومقبولة، ولذلك تكون كل المفردات (34 مفردة) صادقة بمحك قدرتها التمييزية بين الأفراد.

[ج] التحقق من الاتساق الداخلي للاختبار:

جرى حساب معاملات الارتباط بعد حذف المفردات غير المناسبة، وقد اتضح أن قيم معاملات الارتباط بين درجات المفردات ودرجة المحور الذي تنتمي إليه جميعها قيم مقبولة ودالة إحصائيًا عند مستوى (0.01). كما كانت قيم معاملات الارتباط بين درجات المحاور الثلاثة (التخطيط، والمراقبة، والتقويم)، والدرجة الكلية (0.806، 0.578، 0.732) على التوالي، وجميعها قيم مرتفعة ودالة إحصائيًا، مما يدل على وجود ارتباطٍ قوي بين درجة كل محور والدرجة الكلية، وهذا دليل على وجود اتساق داخلي للاختبار ككل مما يعكس صدق بنائه.

[ء] المكوّن الكيفي لعملية صدق المفردات:

ولضرورة توضيح البعد الاجتماعي النفسي لتطور مفردة الاختبار وتفسير الدرجة، وضرورة إدراج الأدلة الكيفية وتأكيد فكرة تطبيق نظرية صدق الاختبار التي تتضمن مكوِّنًا كيفيًا كما أشار سيكولسكي (Secolsky,et al,2011) .

ولأنه يمكن الدراسة العميقة لكل مفردة من مفردات الاختبار؛ لأن برمجة الأدوات أتاحت تسجيل تلفظات الطلبة وردود أفعالهم اللفظية على كل مفردة من المفردات، فقد جرى الرجوع إلى المفردات التي حُذفت لتوثيق العناصر السلوكية والسياقية لعملية الصدق للوصول إلى فهمٍ جيدٍ وكافٍ للتفاعل الممكن لهذه الجوانب الكيفية للعملية مع الأدلة السيكومترية.

وجرى هذا الإجراء استجابةً لما أوصت به تشيلنجس (Schelling, 2011) بضرورة استخدام بروتوكولات التفكير المسموع، وجمع البيانات الكيفية من المقابلات والملاحظات حول كيفية استجابة المفحوصين على المفردات المتضمنة في الاختبارات والاستبيانات المحددة بالمهمة وذلك أثناء وبعد استجابتهم عليها، حيث يوفر ذلك معلوماتٍ ذات قيمة حول كيفية تفسير الطلبة للمفردات وكيفية إدراكهم لها، كما يساعد على شرح وتفسير النتائج التي يتم الحصول عليها، وربما قد يساعد ذلك على تحديد إستراتيجيات أخرى قد أغفلها الباحث، مما يساعد على تطوير مقاييس أكثر دقة.

https://doi.org/10.20428/ijtd.v14i1.2309

ومن هذه الخطوة جرى التوصل إلى أدلةٍ كيفيةٍ تؤكد الأدلة السيكومترية التي تم الحصول عليها فيما يتعلق بالمفردات التي حققت أدلة سيكومترية مقبولة وعددها (34) مفردة، والمفردات التي لم تحقق أدلة سيكومترية مقبولة، والتي تم حذفها وعددها (11) مفردة.

(3) صلاحية الاستبيان المحدد بالمهمة (TSQ) للاستخدام:

بحذف المفردات الخاصة بالمشكلتين الأخيرتين المحذوفتين وعددها (12) مفردة، أصبح عدد مفردات الاستبيان المحدد بالمهمة (36) مفردة، موزعة على الست المشكلات، وللحصول على الأدلة السيكومترية للتحقق من صدق المفردات جرى حساب:

[أ] **حساب معامل (ألفاكرونباخ) لمفردات كل مهارة على حدة بعدد مفردات كل مهارة وفي كل مرة تحذف درجات إحدى المفردات من الدرجة الكلية للمهارة التي تنتمي إليها ومقارنة القيم الناتجة بمعامل ألفا الكلي للمهارة بدون حذف أي مفردة. وقد جاءت معاملات ثبات المفردات أقل من معاملات ثبات المحاور التي تنتمي إليها، أي أن جميع المفردات ثابتة، حيث إن تدخل المفردة لا يؤدي إلى خفض معامل الثبات الكلي للمحور الذي تنتمي إليه العبارة. كما أن معامل الثبات الكلي للاستبيان بطريقة ألفا يساوي (0.906) وهي قيمة مرتفعة.**

[ب] **حساب معامل الثبات بطريقة التجزئة النصفية بمعادلة جتمان، وقد كانت معاملات الثبات للمحاور الثلاث التخطيط، والمراقبة، والتقويم، هي على التوالي (0.549، 0.688، 0.720) ومعامل الثبات الكلي يساوي (0.627) وهي معاملات ثبات مرتفعة، مما يدل على الثبات الكلي للاستبيان.**

[ج] **حساب القدرة التمييزية لكل مفردة من مفردات الاستبيان المحدد بالمهمة (TSQ) وقد امتدت قيم معاملات التمييز لمفردات محور التخطيط بين (0.331، 0.561) ولمفردات محور المراقبة بين (0.445، 0.616) ولمفردات محور التقويم بين (0.543، 0.729) وجميعها قيم موجبة ومرتفعة، مما يعطي ثقة في صدق المفردات وقدرتها على التمييز.**

[د] **التحقق من الاتساق الداخلي للاستبيان المحدد بالمهمة: وقد كانت قيم معاملات الارتباط بين درجة كل مفردة ودرجة المحور الذي تنتمي إليه مرتفعة ودالة إحصائيًا عند مستوى (0.01)، كما جاءت قيم معاملات الارتباط بين درجة كل محور والدرجة الكلية أكبر من (0.78) مما يدل على وجود ارتباط قوي بين درجة كل محور ودرجة الاستبيان ككل، وهذا يدل على وجود اتساق داخلي للاستبيان ككل.**

[هـ] **المكون الكيفي لعملية صدق مفردات الاستبيان المحدد بالمهمة:**

جرى استعراض ملفات الفيديو المسجلة للطلبة في المقاطع الخاصة بالاستجابة على الاستبيان المحدد بالمهمة لفحص كيفية استجابة الطلبة على كل مفردة، للتأكد من أن الأدلة السيكومترية المقبولة والمرتفعة للاستبيان المحدد بالمهمة لم تكن وليدة الصدفة أو ناتجة عن اختيارات عشوائية غير دقيقة، وإنما كانت ناتجة عن إجراءات وتقييمات صادقة.

وقد اتضح من الفحص أن مفرداته كانت واضحة ومباشرة وغير محيرة، ولاسيما أنه يتطلب منهم الاختيار بين بديلين إما نعم أو لا، وقد سهل ذلك عملية الاستجابة. كما أن ذلك قد يكون متوقعًا؛ لأنه من سمات الاستبيان المحدد بالمهمة أن مفرداته تشير إلى المهمة المحددة التي انتهى المفحوص من حلها على التو، حيث تساعد المفردات على إعادة السياقات النفسية التي وقعت فيها المعالجة الإستراتيجية، مما لا يسمح بعملية النسيان أو التشوش.

فكافة مفردات الاستبيان تقريبًا تسأل عن تفاصيل دقيقة في المشكلة التي قاموا بحلها على التو والمعروضة أمامهم أثناء الاستجابة على الاستبيان، فلا مجال يسمح بالفهم غير الصحيح أو الأحكام والتعميمات العامة التي تعد مهددًا لصدق المفردات.

(4) الصورة النهائية لأدوات الدراسة المقننة (بطارية أدوات لقياس سلوك حل المشكلة ومهارات ما وراء المعرفة).

مما سبق يتضح أن كل من (اختبار سلوك حل المشكلة، والاختبار المتعدد الاستجابات، والاستبيان المحدد بالمهمة) في صورهم النهائية بعد حذف المفردات غير المناسبة يتمتعون بدرجةٍ كبيرةٍ من الثبات والصدق وفقًا لما جرى الحصول عليه من أدلة سيكومترية وكيفية، وبذلك يمكن الاطمئنان للنتائج التي جرى الحصول عليها من تطبيق المقاييس الثلاثة على عينة التطبيق الأساسية وعددها (110) طالبًا وطالبة في الإجابة عن أسئلة الدراسة.

نتائج الدراسة ومناقشتها:

تحليل النتائج المتعلقة بالسؤال الأول ومناقشتها.

ينص السؤال الأول على :"ما مدى صدق الأدوات التي أعِدَّت في الدراسة الحالية في قياس مهارات ما وراء المعرفة؟"
للإجابة عن هذا السؤال اتَّبعت الباحثة الخطوات الآتية:

[1] التحقق من الصدق التلازمي الذي يتمثل في علاقة الأدوات بعضها ببعض، والذي يمكن تدعيمه بواسطة الصدق التقاربي (Convergent validity) حيث ينبغي أن تسير طرق القياس المختلفة في نفس الاتجاه طالما أنها تقيس نفس البناء، وللتحقق من ذلك قامت الباحثة بالآتي:

[1-1] حساب معامل الارتباط بين درجات الطلبة على القياسات المتزامنة لمهارات ما وراء المعرفة المتمثلة في البروتوكولات اللفظية المتزامنة (TAP) ودرجاتهم على القياسات غير المتزامنة المتمثلة في الاستبيان المحدد بالمهمة (TSQ) والاختبار متعدد الاستجابات (MRT).

يوضح جدول (1، 2) النتائج التي جرى التوصل إليها، حيث كانت جميع معاملات الارتباط على مستوى كل مشكلة وعلى المستوى الإجمالي دالة إحصائيًا عند مستوى دلالة (0.01) أي أنه يوجد ارتباط موجبٌ ودالٌ إحصائيًا بين درجات الطلبة على القياسات المتزامنة لمهارات ما وراء المعرفة المتمثلة في البروتوكولات اللفظية المتزامنة ودرجاتهم على القياسات غير المتزامنة المتمثلة في كل من الاختبار متعدد الاستجابات، والاستبيان المحدد بالمهمة. ومن المرجح أن يكون السبب في ذلك طبيعة بناء كل من (MRT ,TSQ) اللذان جرى بناؤهما في ضوء البيانات الكيفية التي جُمِعت من المقابلات والملاحظات والبروتوكولات اللفظية نفسها على مهام محددة، وتتفق تلك النتيجة مع ما توصلت إليه دراسة كل من:

https://doi.org/10.20428/ijtd.v14i1.2309

(Braten & Samuelstuen, 2007; Samuelstuen & Braten, 2007; Schellings, 2011) والتي توصلت إلى ارتباطات مقبولة ودالة إحصائيًا بين المقاييس المتزامنة والمقاييس غير المتزامنة المتمثلة في الاستبيانات المحدد بالمهمة.

جدول (1) : معاملات الارتباط بين درجات مهارات ما وراء المعرفة كما قيست بكل من (MRT، TAP)

المهارة	المشكلة الأولى	المشكلة الثانية	المشكلة الثالثة	المشكلة الرابعة	المشكلة الخامسة	المشكلة السادسة	الاجمالي
التخطيط	0.370**	0.579**	0.297**	0.374**	0.443**	0.365**	0.664**
المراقبة	0.464**	0.463**	0.464**	0.224*	0.349**	0.355**	0.462**
التقويم	0.500**	0.057	0.577**	0.306**	0.269**	0.320**	0.351**
الدرجة الكلية	0.556**	0.408**	0.490**	0.323**	0.542**	0.357**	0.650**

جدول (2) : معاملات الارتباط بين درجات مهارات ما وراء المعرفة كما قيست بكل من (TSQ، TAP)

المهارة	المشكلة الأولى	المشكلة الثانية	المشكلة الثالثة	المشكلة الرابعة	المشكلة الخامسة	المشكلة السادسة	الإجمالي
التخطيط	0.380**	0.255*	0.268**	0.502**	0.427**	0.574**	0.500**
المراقبة	0.399**	0.434**	0.281**	0.226*	0.437**	0.377**	0.562**
التقويم	0.420**	0.182	0.217*	0.407**	0.363**	0.358**	0.448**
الدرجة الكلية	0.568**	0.254*	0.313**	0.545**	0.493**	0.434**	0.676**

[2-1] حساب معامل الارتباط بين درجات الطلبة على المقاييس غير المتزامنة لمهارات ما وراء المعرفة المتمثلة في الاختبار متعدد الاستجابات (MRT) والاستبيان المحدد بالمهمة (TSQ).

يوضح جدول (3) النتائج حيث تراوحت معاملات الارتباط على مستوى المهارات الفرعية ما بين (0.403، 0.507)، وكان معامل الارتباط الكلي يساوي (0.694) وجميعها معاملات ارتباط دالة إحصائيًا عند مستوى دلالة (0.01).

جدول (3) : معاملات الارتباط بين درجات الطلبة على *TSQ، MRT*

المهارة	المشكلة الأولى	المشكلة الثانية	المشكلة الثالثة	المشكلة الرابعة	المشكلة الخامسة	المشكلة السادسة	الإجمالي
التخطيط	0.294**	0.369**	0.300**	0.394**	0.381**	0.449**	0.339**
المراقبة	0.447**	0.438**	0.285**	0.382**	0.473**	0.452**	0.518**
التقويم	0.381**	0.371**	0.214*	0.549**	0.273**	0.531**	0.582**
الدرجة الكلية	0.458**	0.403**	0.428**	0.429**	0.422**	0.507**	0.694**

وربما يرجع السبب في هذه القيم المرتفعة والدالة لمعاملات الارتباط سواءً على مستوى المهارة الفرعية أو على مستوى الدرجة الكلية للمهارات في كل مشكلة على حدة أو على المستوى الإجمالي إلى أنّ كلًا من (MRT & TSQ) يندرجان تحت نفس المنهجية في القياس وهي القياس غير المتزامن (Offline Measurement) كما أنهما يستهدفان قياس نفس المهارات، وقد جرى صياغة مفرداتهم من تحليلات كيفية تتضمن تتبع محاولات الطلبة ومساراتهم في الحل وبروتوكولاتهم اللفظية أثناء حل نفس المشكلات؛ فكل من مفردات الاختبار والاستبيان تحمل إيماءات واضحة خاصة بالمشكلة المعروضة، مما لا يسمح بأي مجال لأية تعميمات أو ممارسات قبلية من قبل الفرد على أية مشكلات أو خبرات سابقة.

[2] التحقق من الصدق التنبؤي الذي يتمثل في قدرة الأدوات الثلاثة على التنبؤ بالأداء على حل المشكلة:

تشير معظم الدراسات أن مهارات ما وراء المعرفة تؤدي إلى مخرجات تعلم جيدة، وبالتالي ينبغي أن تكون أداة التقييم التي تستهدف قياس مهارات ما وراء المعرفة مؤشرًا كافيًا وملائمًا لمخرجات التعلم (Messick,1995; Veenman,2007; Veenman,2011) .

وللتحقق من ذلك جرى استخدام نموذج تحليل الانحدار المتعدد بطريقة الانحدار التدريجي (Stepwise Regression) وحساب معامل التحديد أو نسبة التباين الكلي (مربع معامل الارتباط) لمعرفة مدى إسهام كل أداة من أدوات القياس في التنبؤ بالأداء المتمثل في سلوك حل المشكلة، وقد اتضح من التحليل وجود ثلاثة نماذج كما هو موضح بجدول (4):

جدول (4) :الانحدار التدريجي لاختبار قدرة مهارات ما وراء المعرفة المقاسة بكل من البروتوكولات والاختبار والاستبيان على التنبؤ بالأداء

النماذج	قيمة (ف)	مستوى الدلالة	معامل الارتباط (R)	معامل التحديد (R^2)	النسبة المئوية للمساهمة	ثابت المعادلة	معامل الانحدار	قيمة (ت)	مستوى الدلالة
النموذج الأول الاختبار (MRT)	134.255	0.01	0.744	0.554	٪55.4	25.610-	0.564	11.587	0.01
النموذج الثاني الاختبار (MRT) البروتوكولات (TAP)	93.020	0.01	0.797	0.635	٪8.1 (٪63.5)	28.529-	0.356 0.720	5.803 4.862	0.01 0.01
النموذج الثالث الاختبار (MRT) البروتوكولات (TAP) الاستبيان (TSQ)	65.949	0.01	0.807	0.651	٪1.6 (٪65.1)	32.369-	0.308 0.574 0.134	4.795 3.597 2.224	0.01 0.01 0.05

يتضح من جدول (4):

- ثبات صلاحية النماذج الثلاثة نظرًا لارتفاع قيمة "ف" المحسوبة، فقد بلغت (65.949) عند مستوى معنوية (0.01) مما يعني أن لكل أداة من الأدوات الثلاثة إسهامًا إيجابيًا ودالًا إحصائيًا في التنبؤ بسلوك حل المشكلة.

https://doi.org/10.20428/ijtd.v14i1.2309

- وكان أفضل النماذج هو النموذج الثالث، حيث إنه بإضافة (TSQ) إلى كل من (MRT) (TAP). أدى ذلك إلى ارتفاع معامل الارتباط إلى (0.807) وكذلك معامل التحديد ليصبح (0.651). وبذلك فإضافة (TSQ) أدى إلى زيادة نسبة إسهام الأدوات المستخدمة في تفسير الأداء والتنبؤ بمقدار(1.6٪) لتصل إلى (65.1٪).
- المهارات المقاسة بكل أداة من الأدوات الثلاث لها تأثير إيجابي وذو دلالة إحصائية على الأداء المتمثل في سلوك حل المشكلة، ويمكن التعبير عن معادلة الانحدار بين المهارات المقاسة بالأدوات الثلاث والأداء على سلوك حل المشكلة كالتالي:

الأداء (سلوك حل المشكلة) = 32.369 + (0.308 × المهارات المقاسة بالاختبار MRT) + (0.574 × المهارات المقاسة بالبروتوكولات المتزامنة TAP) + (0.134 × المهارات المقاسة بالاستبيان المحدد بالمهمة TSQ).

ومن هذه النتائج يتضّح أنه يمكن التنبؤ بسلوك حل المشكلة من المهارات المقاسة بالأدوات الثلاث ولكن كان الاختبار متعدد الاستجابات الذي يندرج تحت المقاييس غير المتزامنة أكثرها إسهامًا في تفسير الأداء والتنبؤ به، وتختلف هذه النتيجة مع عدد من الدراسات السابقة مثل: (Bannert& Mengel Kamp, 2008;Cooper, et al., 2008; Desoete, 2007; Desoete, 2008; Santiago, 2008). والتي استخدمت التصميم متعدد الطرق (Multi Method Design) في قياس مهارات ما وراء المعرفة والذي يجمع بين المقاييس المتزامنة المتمثلة في التفكير بصوتٍ مسموعٍ والمقاييس غير المتزامنة المتمثلة في الاستبيانات، حيث كشفت جميعها عن اختلاف نتائج القياسات المتزامنة وغير المتزامنة، كما كانت نتائج المقاييس المتزامنة هي الأكثر واقعية والأكثر ارتباطًا بالأداء على المهام بصوره المختلفة.

ويمكن أن يرجع سبب الاختلاف بين نتائج تلك الدراسات ونتائج الدراسة الحالية إلى نوعية المقاييس غير المتزامنة المستخدمة، ففي تلك الدراسات كانت الاستبيانات المستخدمة عامة غير متعلقة بمهام محددة، أما في الدراسة الحالية فقد جرى بناء المقاييس غير المتزامنة بشكل محدد بالمهمة حيث تتعلق جميع مفرداتها بالمهمة التي كان الفرد بصددها على التو.

وهذه النتائج تؤكد فكرة أنه عبر التدخل في تصميم المقاييس غير المتزامنة بحيث تكون على مهام محددة تقرَّب مؤشرات كل من المقاييس المتزامنة وغير المتزامنة بحيث تستهدف شيئًا واحدًا، وهذا مؤشرٌ للصدق (الصدق التقاربي). وتتفق تلك النتيجة مع ما توصلت إليه دراسة (Bråten & Samuelstuen, 2007;Samuelstuen & Bråten, 2007; Schellings, 2011) التي توصلت إلى ارتباطات مقبولة ودالة إحصائيًا بين المقاييس المتزامنة والمقاييس غير المتزامنة المتمثلة في الاستبيانات المحدد بالمهمة.

أسهمت المهارات المقاسة بالأدوات الثلاث في تفسير الأداء والتنبؤ به، وهذا دليل على صدق تلك الأدوات في قياسها للمهارات ما وراء المعرفية ذات العلاقة القوية مع الأداء. وقد كانت المهارات المقاسة بالاختبار متعدد الاستجابات الأكثر إسهامًا في تفسير الأداء والتنبؤ به عن البروتوكولات المتزامنة ويليها الاستبيان، وبذلك يمكن القول إنه عند إعداد الاختبارات بشكل محدد بالمهام يتم تقليل المهددات التي تهدد صدق المقاييس غير المتزامنة إلى درجة جعلتها أعلى من البروتوكولات المتزامنة شديدة الحساسية.

وربما كان الاختبار الأفضل، حيث يُكلف الفرد فيه باختيار أحد البدائل المتعلقة بطريقة تفكيره أو مساره في الحل أو تصرفه عند مواجهته لصعوبة محددة من بين أربعة بدائل تعبر عن مسارات مختلفة في التفكير. أي أن المفردة محددة والمسار الذي يختاره الفرد محدد بالفعل، لا مجال للتأويل أو الفهم الخاطئ، أو الاختيار التقريبي من قبل الطالب، أو التشفير الخاطئ من قبل الباحثة، فضلًا عن ذلك فإن الاختبار متعدد الاستجابات يتضمن مفردات تكيفية متوقفة على اختيار الفرد لبديل معين في أحد المفردات، فتأتي المفردات بعدها متعلقة باختياره، مما يسمح بتتبع مسار الفرد وطريقة تفكيره وقدرته على التكيف الإستراتيجي. وهو الأمر الذي قد يفسر كيف أصبحت المهارات المقاسة بالاختبار أكثر قدرة على التنبؤ بالأداء عن الاستبيان المحدد بالمهمة، والذي لا يتضمن مفردات تكيفية، كما أن مفرداته تتطلب نمطًا مختلفًا من الاستجابة بين بديل أحدهما نعم، والآخر لا.

وبهذا يمكن تأكيد ما أشارت إليه تشيلنجس (Schellings, 2011) في قولها:

"إذا كانت المقاييس غير المتزامنة مبنية بدقة وجرى اختبارها بطرق سليمة فقد تصبح بدائل كافية ومقنعة للمقاييس المتزامنة المكثفة المعملية في قياس إستراتيجيات ومهارات الأفراد".

[3] استخدام المعايرة (Calibaration) كمحك آخر لكل من الصدق التلازمي والصدق التنبؤي للأدوات.

جرى استخدام المعايرة لقياس دقة المراقبة ماوراء المعرفية (Metacognitive Monitoring Accuracy) للاستدلال على جودة مهارات ماوراء المعرفة عمومًا بوصفها في مركز العديد من النماذج الخاصة بمهارات ماوراء المعرفة وذلك على غرار عددٍ من الدراسات مثل (Desoete, 2007; Desoete, 2008; Sampaio & Brewer, 2009).

جرى استخدام درجات معايرة الطلبة (القبلية والبعدية) كمحك آخر للصدق التلازمي عن طريق دراسة الارتباطات بين 555555درجات الطالب على المراقبة ما وراء المعرفية المقاسة بالأدوات الثلاثة (الاختبار، والاستبيان، والبروتوكول) ودرجات معايرة الطلبة (القبلية والبعدية)، ثم دراسة الارتباط بين درجات معايرة الطلبة (القبلية، والبعدية) والأداء الفعلي بوصفها محكًا للصدق التنبؤي؛ وذلك انطلاقًا من النتيجة التي أشارت إليها ستون (Stone,2000) والتي تؤكد أن المستويات المختلفة من المعايرة ينتج عنها أطوار مختلفة في إكمال المهمة، وبالتالي مستويات مختلفة في الأداء، وذلك في كل مشكلة على حدة. وجرى ذلك عبر الخطوات الآتية:

1- استخراج درجات المعايرة القبلية والبعدية لكل مشكلة، والتي تتراوح في الثلاث المشكلات الأولى من (-6 إلى +6) وفي الثلاث المشكلات الأخيرة من (-3 إلى +3).
2- تحويل درجات معايرة الطلبة القبلية والبعدية إلى مقياس متدرج، بحيث يحصل الأفراد الذين كانت درجة معايرتهم صفرًا على أعلى درجة، ويحصل الأفراد الذين تطرفت درجة معايرتهم إلى الدرجة القصوى سواءً بالموجب أو بالسالب على أقل درجة، حيث كانت أحكامهم غير دقيقة، فهم مبالغون في أحكامهم إذا كانت درجات معايرتهم موجبة، أو مقللون في أحكامهم إذا كانت درجات معايرتهم سالبة.
3- حساب معامل الارتباط بين دقة أحكام الأفراد القبلية والبعدية التي جرى استخراجها من الخطوة السابقة ودرجات المراقبة ما وراء المعرفية المقاسة بكل من البروتوكولات (TAP) والاختبار (MRT) والاستبيان (TSQ) في كل مشكلة على حدة.

https://doi.org/10.20428/ijtd.v14i1.2309

4- حساب معامل الارتباط بين دقة أحكام الأفراد القبلية والبعدية التي جرى استخراجها من الخطوة (2) والأداء في كل مشكلة على حدة. ويوضح جدول (5)، (6) النتائج التي تم الحصول عليها.

جدول (5) : معاملات الارتباط بين دقة الأحكام القبلية وكل من درجات المراقبة ما وراء المعرفية والأداء

المشكلة مهارات المراقبة المقدرة بـ	المشكلة الأولى	المشكلة الثانية	المشكلة الثالثة	المشكلة الرابعة	المشكلة الخامسة	المشكلة السادسة	الإجمالي
البروتوكولات	0.113	0.171	0.284**	0.011	0.314**	0.173	0.465**
الاختبار	0.062	0.011	0.136	0.200*	0.351**	0.164	0.121
الاستبيان	0.119	0.075	0.177	0.041	0.132	0.164	0.327**
الأداء على المشكلة	0.357**	0.047	0.260**	0.240*	0.385**	0.638**	0.512**

جدول (6) :معاملات الارتباط بين دقة الأحكام البعدية وكل من درجات المراقبة ما وراء المعرفية والأداء

المشكلة مهارات المراقبة المقدرة بـ	المشكلة الأولى	المشكلة الثانية	المشكلة الثالثة	المشكلة الرابعة	المشكلة الخامسة	المشكلة السادسة	الإجمالي
البروتوكولات	0.264**	0.256**	0.250*	0.212*	0.3**	0.366**	0.448**
الاختبار	0.273**	0.264**	0.29**	0.209*	0.38**	0.208*	0.315**
الاستبيان	0.407**	0.259**	0.36**	0.39**	0.2*	0.253*	0.465**
الأداء على المشكلة	0.641**	0.218*	0.55**	0.604**	0.89**	0.304**	0.771**

وتوضح النتائج أنه: ارتبطت دقة أحكام الأفراد البعدية ارتباطًا موجبًا ودالاً مع مهارات المراقبة ما وراء المعرفية المقدرة عبر كل من البروتوكولات، والاختبار، والاستبيان، على مستوى كل مشكلة من المشكلات الست وعلى المستوى الإجمالي، كما ارتبطت أيضًا ارتباطًا موجبًا ودالاً مع الأداء على مستوى كل مشكلة على حدة وعلى المستوى الإجمالي، وكانت معاملات الارتباط تتراوح ما بين (0.218 إلى 0.89).

في حين كانت معاملات الارتباط بين دقة أحكام الأفراد القبلية غير دالة مع مهارات المراقبة ما وراء المعرفية المقدرة بالبروتوكولات في كل المشكلات ماعدا المشكلة الثالثة والخامسة والمقدرة بالاختبار في كل المشكلات ماعدا المشكلة الرابعة والخامسة والمقدرة بالاستبيان في المشكلات كافة. وهذا يعني أن المعايرة البعدية أكثر دقة من المعايرة القبلية.

ويتفق ذلك مع ما توصل إليه كل من (Cummings, 2006; Snyder, et al., 2011) وقد أشار (Cummings, 2006) إلى أنه ربما يرجع السبب في أن التنبؤات البعدية أكثر دقة من التنبؤات القبلية، إلى أن الأفراد يعتمدون في إصدارهم للتنبؤات القبلية والبعدية على معلومات مختلفة. فالتنبؤات القبلية تعد أحكام مراقبة مستقبلية (Prospective Monitoring Judgment) بمعنى أن يراقب الفرد مهاراته وخلفياته السابقة قبل القيام بالمهمة وهذا بدوره يتطلب مهارة أكثر، في حين أن الفرد عند إصداره الحكم البعدي يكون لديه معلومات أخرى تتعلق بطبيعة المهمة ودرجة صعوبتها، فضلاً عن معلوماته حول أدائه الذي انتهى منه على التو.

ومن ذلك يمكن القول إن درجات معايرة الطلبة البعدية تعد مؤشرًا جيداً ومرتبطًا مع مهارات المراقبة المقدرة من خلال الأدوات الثلاث التي تستهدف قياس مهارات ما وراء المعرفة، ويُعدُّ ذلك محكًا لصدق الأدوات وصدق مقياس

المعايرة (الصدق التقاربي)، كما أن ارتباط درجات معايرة الطلبة البعدية بالأداء يُعدُّ محكًا خارجيًا يدل على صدق الأدوات وفقًا للنظرية (الصدق المرتبط بالمعيار الخارجي).
(Messick, 1995; Stone, 2000; Veenman, 2007; Veenman, 2011)

تحليل النتائج المتعلقة بالسؤال الثاني ومناقشتها.

ينص السؤال الثاني على :"ما مدى قدرة أدوات الدراسة على تمييز مهارات ما وراء المعرفة النوعية المستخدمة في كل مشكلة على حدة بين الأفراد مرتفعي ومنخفضي الأداء على حل المشكلة؟" وللإجابة عن هذا السؤال قامت الباحثة بالخطوات الآتية:

1- تحديد مجموعة المرتفعين والمنخفضين في حل المشكلة، وذلك كما يلي:

أ- ترتيب عينة الدراسة (110) طالبًا وطالبة في كل مشكلة على حدة تنازليًا من الأعلى إلى الأدنى في الدرجة الفعلية على المشكلة، ثم إعادة ترتيبهم تصاعديًا على حسب الزمن من الأدنى إلى الأعلى في الزمن مع عدم الإخلال بالترتيب الأول، ثم إعادة ترتيبهم تصاعديًا على حسب عدد الحركات من الأدنى إلى الأعلى في عدد الحركات مع عدم الإخلال بالترتيبين الأول والثاني. وبذلك فقد ترتبت عينة الدراسة الحالية على أساس الدرجة، ثم الزمن، ثم عدد الحركات.

ب- تحديد مجموعة الطلبة المرتفعين في الأداء وهم أعلى (27٪) وعددهم (30) طالبًا وطالبة، وهم الطلبة الأعلى في الدرجة والأقل في الزمن وفي عدد الحركات وفئة الطلبة المنخفضين في الأداء وهم أقل (27٪) وعددهم (30) طالبًا وطالبة، وهم الطلبة الأقل في الدرجة والأعلى في الزمن وفي عدد الحركات، وذلك في كل مشكلة على حدة، ثم على المستوى الإجمالي.

2- دراسة الفروق بين الأفراد مرتفعي ومنخفضي الأداء على سلوك حل المشكلة في مهارات ما وراء المعرفة المقاسة بكل أداة على حدة {البروتوكولات اللفظية المتزامنة (TAP) و بالاختبار متعدد الاستجابات (MRT) وبالاستبيان المحدد بالمهمة (TSQ) } بواسطة اختبار (ت) لعينتين مستقلتين. والنتائج موضحة بجدول (7، 8، 9) .

جدول (7) :الفروق في مهارات ما وراء المعرفة كما قيست بواسطة تحليل البروتوكولات بين مجموعتي الطلبة المرتفعين والمنخفضين في حل المشكلة

المهارة	المشكلة	المرتفعون		المنخفضون		قيمة (ت)
		المتوسط	الانحراف	المتوسط	الانحراف	
التخطيط	الأولى	6.30	1.66	3.83	1.95	3.13**
المراقبة		5.63	2.11	3.77	1.52	3.93**
التقويم		5.10	1.65	3.63	1.77	3.32**
المجموع		17.03	3.61	12.23	3.33	5.35**
التخطيط	الثانية	4.43	1.43	3.77	1.28	1.90
المراقبة		7.83	1.91	6.63	2.20	2.25*
التقويم		7.80	2.59	6.07	2.60	2.58*
المجموع		20.07	4.08	16.47	4.73	3.15**

https://doi.org/10.20428/ijtd.v14i1.2309

التخطيط	الثالثة	4.47	5.94	4.20	0.71	1.24
المراقبة		4.97	1.10	4.10	0.71	3.63**
التقويم		4.60	1.25	3.37	1.41	3.58**
المجموع		14.03	1.75	11.67	0.71	6.86**
التخطيط	الرابعة	6.23	1.45	4.97	1.77	3.03**
المراقبة		4.93	1.76	4.00	1.44	2.25*
التقويم		6.87	1.68	5.77	2.27	2.14*
المجموع		18.03	3.36	14.73	4.18	3.37**
التخطيط	الخامسة	3.63	1.22	2.90	1.16	2.39*
المراقبة		2.57	0.97	2.00	1.20	2.01*
التقويم		3.37	1.38	2.67	1.27	2.05*
المجموع		9.57	2.10	7.57	2.65	3.24**
التخطيط	السادسة	5.20	1.32	4.53	1.22	2.03*
المراقبة		4.20	1.73	3.00	1.51	2.86**
التقويم		4.07	1.74	2.83	1.49	2.95**
المجموع		13.47	3.01	10.37	2.33	4.46**
التخطيط	الإجمالي للمشكلات الست	30.27	3.44	25.20	3.59	5.58**
المراقبة		30.13	5.05	23.50	3.37	5.98**
التقويم		31.80	4.21	24.33	5.20	6.11**
المجموع		92.20	8.88	73.03	7.92	8.83**

قيمة (ت) الجدولية حيث درجة الحرية (58) عند مستوى 0.05 = 2 وعند مستوى0.01 = 2.66

جدول (8) :الفروق في مهارات ما وراء المعرفة كما قيست بالاختبار متعدد الاستجابات بين مجموعتي الطلبة المرتفعين والمنخفضين في حل المشكلة

المهارة	المشكلة	المرتفعون		المنخفضون		قيمة (ت)
		المتوسط	الانحراف	المتوسط	الانحراف	
التخطيط	الأولى	4.30	1.34	2.53	0.82	6.15**
المراقبة		2.67	0.55	2.20	0.89	2.45*
التقويم		2.40	0.56	1.97	0.67	2.71**
المجموع		9.37	1.71	6.70	1.66	6.12**
التخطيط	الثانية	4.07	1.05	1.67	1.12	8.55**
المراقبة		5.27	0.87	4.50	0.73	3.69**
التقويم		6.43	0.82	4.03	1.54	7.53**
المجموع		15.77	1.59	10.20	1.73	12.97**
التخطيط	الثالثة	4.90	0.99	4.23	2.12	1.55
المراقبة		5.07	1.08	3.47	2.83	2.89**
التقويم		2.53	0.51	2.10	1.41	1.58

المجموع		12.50	1.83	9.80	3.54	3.71**
التخطيط	الرابعة	5.07	0.94	3.67	1.63	4.08**
المراقبة		4.30	1.37	3.50	1.38	2.25*
التقويم		4.70	0.88	3.63	0.93	4.58**
المجموع		14.07	2.03	10.80	2.01	6.26**
التخطيط	الخامسة	5.27	0.94	4.53	1.46	2.31*
المراقبة		2.67	0.55	2.17	0.83	2.75**
التقويم		4.20	1.16	3.57	1.17	2.11*
المجموع		12.13	1.89	10.27	2.59	3.19**
التخطيط	السادسة	6.03	1.33	5.13	1.80	2.21*
المراقبة		2.10	0.76	1.53	0.63	3.15**
التقويم		5.70	0.60	4.27	1.17	5.97**
المجموع		13.83	1.80	10.93	2.35	5.37**
التخطيط	الإجمالي للمشكلات الست	29.63	2.89	21.77	3.48	9.52**
المراقبة		22.07	2.36	17.37	2.48	7.51**
التقويم		25.97	2.06	19.57	2.53	10.75**
المجموع		77.67	3.73	58.70	5.32	15.98**

جدول (9) :الفروق في مهارات ما وراء المعرفة كما قيست بالاستبيان المحدد بالمهمة بين مجموعتي الطلبة المرتفعين والمنخفضين في حل المشكلة

المهارة	المشكلة	المرتفعون		المنخفضون		قيمة (ت)
		المتوسط	الانحراف	المتوسط	الانحراف	
التخطيط	الأولى	1.67	0.66	1.23	0.63	2.61*
المراقبة		1.60	0.56	1.03	0.61	3.72**
التقويم		1.57	0.68	1.13	0.73	2.38*
المجموع		4.83	1.32	3.40	1.13	4.52**
التخطيط	الثانية	1.37	0.67	1.00	0.69	2.08*
المراقبة		1.80	0.41	1.47	0.57	2.60*
التقويم		1.53	0.57	1.13	0.68	2.46*
المجموع		4.07	1.29	3.60	1.40	3.15**
التخطيط	الثالثة	1.47	0.57	1.10	0.00	3.51**
المراقبة		1.80	0.41	1.20	0.71	4.02**
التقويم		1.73	0.45	0.97	0.00	9.33**
المجموع		5.00	0.95	3.27	0.71	8.03**

https://doi.org/10.20428/ijtd.v14i1.2309

التخطيط	الرابعة	1.63	0.61	1.23	0.57	2.62*
المراقبة		1.50	0.63	1.13	0.51	2.48*
التقويم		1.77	0.43	1.07	0.64	4.97**
المجموع		4.90	1.35	3.43	0.97	4.83**
التخطيط	الخامسة	1.83	0.38	1.53	0.63	2.24*
المراقبة		1.57	0.57	1.23	0.63	2.16*
التقويم		1.67	0.66	1.30	0.65	2.16*
المجموع		5.07	0.98	4.07	1.14	3.64**
التخطيط	السادسة	1.70	0.47	1.27	0.69	2.85**
المراقبة		1.43	0.67	1.00	0.69	2.44*
التقويم		1.37	0.760	0.83	0.59	3.02**
المجموع		4.50	1.20	3.10	1.15	4.61**
التخطيط	الإجمالي للمشكلات الست	9.67	0.99	7.37	1.71	6.36**
المراقبة		9.70	1.21	7.07	1.44	7.68**
التقويم		9.63	1.25	6.43	2.05	7.32**
المجموع		29.00	2.08	20.87	3.78	10.33**

توضح النتائج وجود فروق ذات دلالة إحصائية عند مستوى (0.01) بين مجموعة الطلبة مرتفعي ومنخفضي الأداء على كل مشكلة من المشكلات الست، وعلى المستوى الإجمالي في متوسط مجموع مهارات ما وراء المعرفة (التخطيط، والمراقبة، والتقويم) المقدرة بواسطة كل من (تحليل البروتوكولات، وبالاختبار متعدد الاستجابات، وبالاستبيان المحدد بالمهمة) .

ولكن هناك تمايزات وفروق بين مجموعتي الطلبة مرتفعي ومنخفضي الأداء في نوع المهارات المستخدمة، وفي نسبة استخدام كل مهارة في كل مشكلة من المشكلات، كما قيست بواسطة كل من تحليل البروتوكولات، وبالاختبار متعدد الاستجابات، وبالاستبيان المحدد بالمهمة؛ حيث كانت الفروق دالة عند مستوى (0.01) لبعض المهارات، ودالة عند مستوى (0.05) للبعض الآخر، ولم تكن دالة للتخطيط في المشكلة الثانية والثالثة كما قيست بتحليل البروتوكولات. وللتخطيط والتقويم في المشكلة الثالثة كما قيست بالاختبار.

وبذلك فقد اتفقت الثلاث أدوات في تمييزها للمهارات الأكثر إسهاماً في ارتفاع الأداء في بعض المشكلات وفي تمييزها لنسبة استخدام المهارات بين المشكلات في بعضها الآخر.

وربما يدعم هذا صدق الأدوات مجتمعة واستهدافها لنفس العمليات وصدقها في التمييز بين فئتي الطلبة وبين المشكلات في المهارات المستخدمة، كما أن تلك الاختلافات الناتجة ربما يمكن إعزاؤها لطبيعة الأدوات وطريقة

الاستجابة عليها وطريقة تقدير الدرجة؛ ولذلك ينبغي على الأبحاث المستقبلية الاهتمام بمناقشة تلك الفروق والتمايزات التي أفرزتها الأدوات الثلاث على مستوى كل مشكلة.

3- المقارنة بين الطلبة مرتفعي ومنخفضي الأداء في مسار حل كل مشكلة بالاعتماد على شكل المنحنى الذي رُصد عن طريق تتبع آثار وحركات الأفراد التي سجلها البرنامج في كل مشكلة على حدة:

يُنتج البرنامج هذه الآثار في شكل منحنى محوره الرأسي يعبر عن الحد الأدنى للحركات الصحيحة المطلوبة لحل المشكلة، تبدأ من الخطوة الأولى إلى الخطوة الأخيرة، ومحوره الأفقي يعبر عن الحركات التي يقوم بها الفرد في اتجاهه للحل التي قد تكون صحيحة أو خاطئة، بحيث توضع الإشارات الخاصة بمهارات الفرد في كل تقاطع بين حركة المفحوص وأقرب حركة صحيحة قام بها بحيث تكون الإشارة إما موجبة إذا كانت المهارة نشطة، والحركة صحيحة أو سالبة إذا كانت المهارة غير نشطة والحركة غير صحيحة.

توفر تلك المنحنيات معلومات ذات قيمة عن مهارات الفرد ونشاطها وتتابعاتها من خلال محاولاته وإجراءاته والتي تتوزع خلال مدى من الحركات المحتملة. وبذلك يمكن استخدامها للمقارنة بين الأفراد مختلفي الأداء في شكل المهارات ونشاطها في كل مرحلة من مراحل الحل، كما أنها توفر معلومات عن أين ومتى وقعت معظم الحركات الزائدة غير الصحيحة، وفي أي خطوة من خطوات الحل، وهل هناك خطوات محورية تميز بين الأفراد مختلفي الأداء.

ونظرًا لوجود عقبات كثيرة تواجه التمثيلات البيانية كما تذكر هادوين (Hadwin, et al., 2007) وهي أنها صعبة التجميع خلال المشاركين، فقد جرت الخطوات الآتية للحصول على منحنى تجميعي لمجموعة الطلبة مرتفعي ومنخفضي الأداء في كل مشكلة على حدة:

أ- فحص منحنيات مجموعة الطلبة مرتفعي الأداء في كل مشكلة على حدة وذلك لحصر تكرارات الإشارات الموجبة والسالبة في كل محاولة على حدة ... على سبيل المثال فحص المحاولة الأولى لدى كل طلبة المجموعة، وتحديد عدد الإشارات الموجبة والسالبة وتسجيلها في شكل بياني.

- محوره الرأسي: يمثل عدد الحركات الصحيحة اللازمة لحل المشكلة، تبدأ من الخطوة الأولى إلى الخطوة الأخيرة، ويمثل ذلك الحد الأدنى للخطوات اللازمة للحل.
- محوره الأفقي: مفتوح لأكبر عدد من الحركات جرى تسجيلها لدى مجموعة الطلبة مرتفعي الأداء

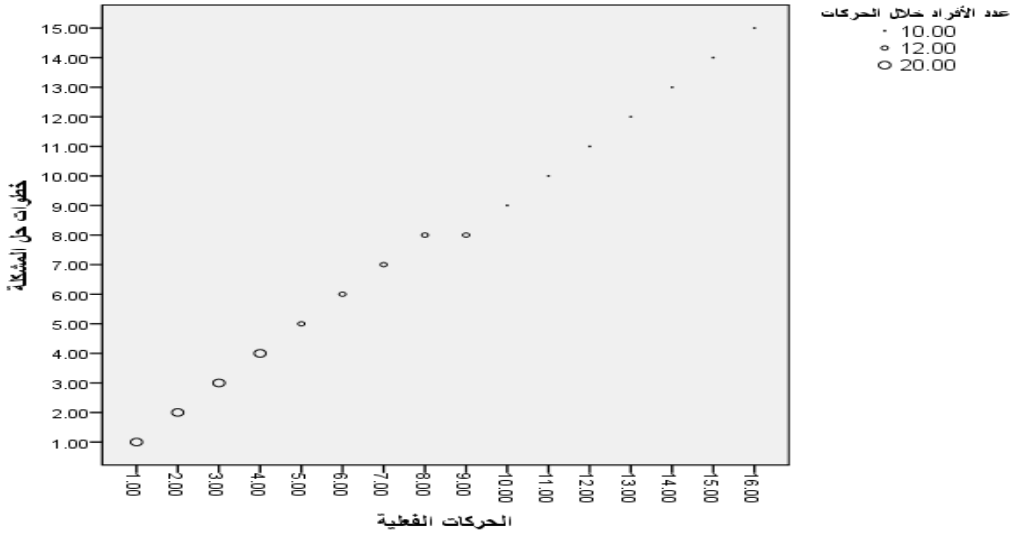
ب- حصر تكرارات الإشارات الموجبة والسالبة في كل محاولة من محاولات طلبة المجموعة ورصدها في التقاطعات الشبكية بين المحوري، بحيث تتضمن كل خانة عدد الحاصلين على إشارة موجبة والحاصلين على إشارة سالبة على المهارة في تلك الخطوة.

ج- فحص منحنيات مجموعة الطلبة منخفضي الأداء على كل مشكلة على حدة بنفس الطريقة السابقة.

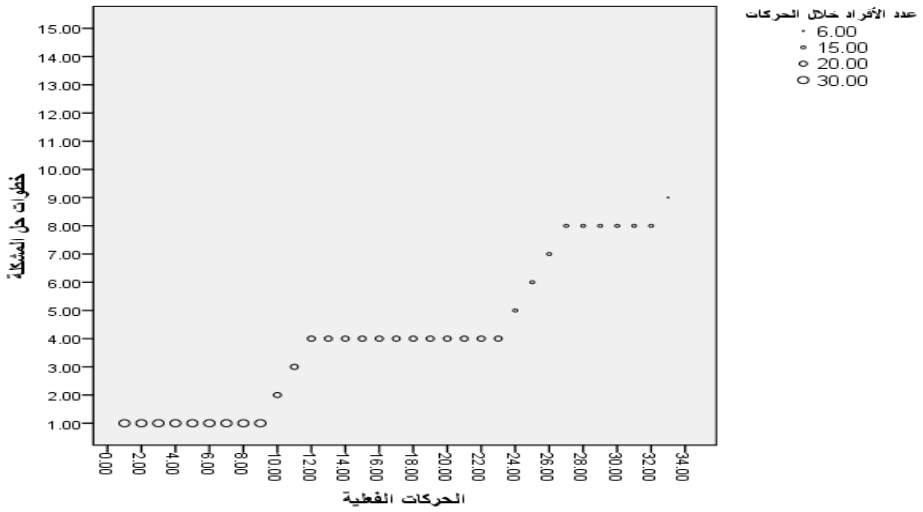
د- حصر عدد الإشارات الموجبة في كل مجموعة أثناء مراحل الحل وفقًا لأقل عدد من الحركات جرى تسجيلها لدى المجموعة المعنية، وذلك لتسهيل عملية المقارنة بين طلبة المجموعتين وذلك لتحديد الخطوات التي يمكن أن تميز بين المجموعتين والتي حدثت فيها معظم الأخطاء لدى مجموعة الطلبة منخفضي الأداء، وقد قامت الباحثة بهذه الخطوات نظرًا لأن عدد الحركات كبيرٌ لدى مجموعة الطلبة منخفضي الأداء بلغ في بعض المشكلات (198) حركة، مما يجعل من الصعب تمثيله في رسم بياني والمقارنة من خلاله.

https://doi.org/10.20428/ijtd.v14i1.2309

ويوضح شكل (1، 2) مثالًا للمنحنيات التجميعية التي جرى الحصول عليها لكل فئة من الطلبة على المشكلة الأولى كمثال، وقد نوقشت المنحنيات التجميعية لكل فئة من الطلبة في كل مشكلة على حدة.



شكل (1): منحنى تجميعي لمسار حركات الطلبة مرتفعي الأداء أثناء حل المشكلة الأولى



شكل (2): منحنى تجميعي لمسار حركات الطلبة منخفضي الأداء أثناء حل المشكلة الأولى

يتضح من شكل (1) أن:

- بدأ معظم الطلبة مرتفعي الأداء (20) طالبًا وطالبة من بين (30) طالبًا وطالبة بخطة صحيحة، وكانت المهارات في حالتها الموجبة، بينما انهمك الباقون (10) طلاب بخطة خاطئة استغرقت منهم عدد من الحركات.
- استمرت المهارات لدى العشرين طالب وطالبة نشطة إلى الخطوة الرابعة من الحل الصحيح والتي انحرف فيها ثمانية من الطلبة عن مسار الحل الصحيح، وزادت الحركات الخاطئة والتي تعني نقص المراقبة، بينما استمر باقي الطلبة وعددهم اثنا عشر طالبًا في المسار الصحيح للحل، واستمرت المهارات نشطة موجبة إلى الخطوة الثامنة من الحل الصحيح والتي

انحرف فيها كل الطلبة بحركة خاطئة واستغرق اثنان منهم في عدد من الحركات الخاطئة والتي تعني نقص في المراقبة والتخطيط، بينما استمرت المهارات نشطة وموجبة لدى العشرة الطلبة الباقين.

- توصل أفضل عشرة طلبة من مجموعة الطلبة مرتفعي الأداء إلى الحل الكامل للمشكلة في عدد من الحركات بلغت (16) حركة، وكان الحد الأدنى للحركات اللازمة لحل المشكلة (15) حركة، واستكمل الباقي خطواتهم في عدد حركاتٍ أكبر كان أقصاها (30) حركة.

يتضح من شكل (2) أن:

- بدأ جميع الطلبة منخفضي الأداء بخطة خاطئة (30) طالبًا وطالبة، وكانت المهارات في حالتها السالبة واستغرق معهم ذلك عدد من الحركات وصلت إلى (9) حركات لدى أفضلهم الذين تمكنوا من تخطي هذه الخطوة وعددهم (20) طالبًا وطالبة، بينما انهمك العشرة الباقون في تكرار الخطوات الخاطئة والتي تعني نقصًا واضحًا في مهارات التخطيط.
- استمرت المهارات لدى العشرين طالب وطالبة نشطة إلى الخطوة الرابعة، ولم يتمكن أحدهم من الانتقال إلى الخطوة التالية، واستمرت الحركات الخاطئة في الزيادة والتي تعني نقص مهارات المراقبة وضياع الهدف، حيث إن بعضهم قد رجع إلى حالة البداية.
- تمكن خمسة عشر طالبًا وطالبة من الانتقال إلى الخطوة التالية بعد عدد من الحركات الزائدة، واستمرت مهاراتهم نشطة موجبة إلى أنه جرى الوصول إلى الخطوة الثامنة، والتي لم يتمكن أحدهم من الانتقال منها إلى الخطوة التالية الصحيحة في مسار الحل، مما أدى إلى ظهور الإشارات السالبة للمهارات، واستمر ذلك في عدد من الحركات إلى أن تمكن ستة من الطلبة فقط من الانتقال إلى الخطوة التالية وهم أفضل ستة طلبة في المجموعة، أما باقي الطلبة استمروا في إعادة الخطوات وتكرار الأخطاء، حيث بلغ أكبر عدد من الحركات (110) حركة وقد كانت معظم الحركات الزائدة في الخطوة الأولى والرابعة والثامنة، والتي اتضح فيها نقص مهارات التخطيط والمراقبة.

ومن ذلك يتضح أن البيانات التي جُمِعت من خلال الآثار توفر معلومات ذات قيمة عن مهارات الطلبة ونشاطها وتتابعاتها خلال كل خطوة من خطوات الحل؛ كما توفر معلومات عن أين ومتى وقعت معظم الحركات الزائدة غير الصحيحة، وتحدد أي الخطوات تمثل عتبات يمكن أن تفرِّق بين مجموعات الطلبة مرتفعي ومنخفضي الأداء.

4- دراسة قدرة مقياس المعايرة القبلية والبعدية بوصفها مؤشرًا لمهارات المراقبة ما وراء المعرفية على التمييز بين فئتي الطلبة مرتفعي ومنخفضي الأداء في نمط أحكامهم سواءً القبلية أو البعدية.

وللحصول على نمط المعايرة الخاص بكل فئة من الطلبة في كل مشكلة على حدة جرت مقارنة متوسطات درجات المعايرة القبلية والبعدية لفئتي الطلبة مرتفعي ومنخفضي الأداء باستخدام اختبار (ت) لعينتين مستقلتين.ويوضح جدول (10) النتائج التي جرى الحصول عليها:

جدول (10): دلالة الفروق بين متوسطات درجات الطلبة مرتفعي الأداء والطلبة منخفضي الأداء على كل من المعايرة القبلية والبعدية

الطلبة مرتفعو الأداء		الطلبة منخفضو الأداء		قيمة (ت)
م	ع	م	ع	

https://doi.org/10.20428/ijtd.v14i1.2309

المشكلة الأولى	المعايرة القبلية	2.73-	1.60	3.27	2.89	9.960**
	المعايرة البعدية	0.27-	0.58	3.43	2.88	6.89**
المشكلة الثانية	المعايرة القبلية	0.93-	1.80	0.20-	2.30	1.377
	المعايرة البعدية	0.07-	0.64	0.43-	1.01	1.68
المشكلة الثالثة	المعايرة القبلية	0.23-	0.43	0.40-	1.73	0.511
	المعايرة البعدية	0.07-	0.25	1.87	3.17	3.33**
المشكلة الرابعة	المعايرة القبلية	0.30-	0.47	0.43	0.86	4.112**
	المعايرة البعدية	0.00	0.69	1.90	1.21	7.44**
المشكلة الخامسة	المعايرة القبلية	0.12-	0.25	1.18-	0.28	15.57**
	المعايرة البعدية	0.00	0.13	0.60	1.74	1.88
المشكلة السادسة	المعايرة القبلية	0.43-	0.50	1.53	0.90	10.477**
	المعايرة البعدية	0.13-	0.51	0.30	0.75	2.62*

يتضح من هذه النتائج أن:

- **المعايرة القبلية لدى فئة الطلبة مرتفعي الأداء كانت في كل المشكلات الست سالبة، فقد كان نمطهم السائد في التقدير أنهم مقللون في درجاتهم المتوقعة عن درجاتهم الفعلية، وقد كان هذا ملحوظًا في المشكلة الأولى والثانية، حيث كان متوسط معايرتهم على الترتيب (-2.73)، (-0.93).**
- **أما بالنسبة لفئة الطلبة منخفضي الأداء فقد اتسم نمط معايرتهم القبلية بأنهم مبالغون في تقديراتهم في كل من المشكلة الأولى والرابعة والسادسة، حيث توقعوا الحصول على درجات أكبر من الدرجات الفعلية التي حصلوا عليها، حيث كان متوسط معايرتهم القبلية على الترتيب (3.27، 0.43، 1.53) بينما كانوا مقللين في كل من المشكلة الثانية والثالثة والخامسة، حيث كان متوسط معايرتهم القبلية على الترتيب (-0.20، -0.40، -1.18).**

- وربما يمكن تفسير الاختلاف الواضح بين فئتي الطلبة المرتفعين والمنخفضين في نمط معايرتهم القبلية أن فئة الطلبة مرتفعي الأداء أكثر وعيًا وتقديرًا لمدى صعوبة المهمة خصوصًا في المشكلة الأولى والرابعة والسادسة والتي كان معامل صعوبتها على الترتيب (75٪، 70٪، 80٪) حيث كان نمط معايرة الطلبة مرتفعي الأداء عليها سالبًا، وتوقعوا الحصول على درجات أقل من الدرجات الفعلية التي حصلوا عليها، وقد يرجع ذلك إلى أنهم لم يقوموا بالمحاولة الفعلية، وإنما كان حكمهم تنبؤيًا، ونظرًا لوعيهم بمدى صعوبة المهمة فقد توقعوا درجات قد تكون أقل من الدرجات الذين يعتقدون الحصول عليها وذلك ليعطوا أنفسهم فرصة أكبر للعمل دون ضغط ولتجنب الإحباط، على العكس تمامًا كان نمط معايرة الطلبة منخفضي الأداء موجبًا، وكانوا مبالغين في تقديراتهم على نفس المشكلات الثلاث ذوات معاملات الصعوبة الأكبر مقارنة بباقي المشكلات ذوات معاملات الصعوبة الأقل منها، وبذلك فهم أقل وعيًا بقدراتهم وبمدى صعوبة المهمة.
- اتسمت المعايرة البعدية لدى فئة الطلبة مرتفعي الأداء بالدقة مقارنة بالمعايرة القبلية، وكانت الدقة في أقصاها في المشكلة الرابعة والخامسة، حيث بلغ متوسط معايرتهم البعدية صفرًا، مما يعني أن الدرجات التي توقع الطلبة الحصول عليها كانت تمامًا هي الدرجات الفعلية التي حصلوا عليها، كما كانت في باقي المشكلات أيضًا أقرب إلى الواقعية على الرغم من أنها ليست تامة، حيث كانت الانحرافات بين الدرجات المتوقعة والدرجات الفعلية في باقي المشكلات ضئيلة تراوحت بين (-0.07 إلى -0.27). وقد اتضح من المقابلات مع الطلبة بعد انتهائهم من حل المشكلات أنهم غالبًا ما يقررون في مقياس المعايرة درجات أقل مما يتوقعونها، وبسؤالهم عن سبب ذلك أجاب بعضهم *"أحب أن أضع فرصة لاحتمالات الخطأ الواردة أو لمعايير التصحيح العالية"*. وبعضهم الآخر *"حتى أسمح لنفسي بالفرحة عندما أرى النتائج النهائية أكبر مما توقعت"*. وبعضهم الآخر *"حتى لا أحبط أضع لنفسي مساحة كحماية للذات"* وبذلك فكانت معايرتهم ليست تامة لأسباب ترجع لسمات شخصية معينة، ومع ذلك فقد كانت أقرب للواقعية.
- أما بالنسبة لفئة الطلبة منخفضي الأداء فقد اتسم نمط معايرتهم البعدية بعدم الدقة، فقد كانوا مبالغين في معظم المشكلات فيما عدا المشكلة الثانية، حيث تراوحت متوسطات معايرتهم في المشكلات الخمس من (0.30 إلى 3.43). وقد يرجع ذلك إلى نقص مهارات المراجعة والتقويم حيث إنهم توقعوا درجات أكبر من الدرجات الفعلية التي حصلوا عليها. وبسؤالهم من المقابلات المتعمقة عن سبب ذلك أقر بعضهم ... أنهم يعلمون من البداية بأن الدرجات على خطوات الحل المتدرجة، وهم قاموا بكثير من الخطوات الصحيحة من وجهة نظرهم.؛ أما بالنسبة للمشكلة الثانية (السودوكو) فقد كان نمط معايرتهم سالبًا، حيث كان المتوسط (-0.43) أي أنهم توقعوا درجات أقل مما حصلوا عليها. وبسؤالهم عند ذلك أقر معظمهم بأن سبب ذلك يرجع إلى خلفياتهم عن هذه النوعية من المشكلات الصعبة والتي لا يحبها الكثير منهم. ويمكن أن يسمى هذا بتأثير الهالة (Halo effect) (Dompnier,et al., 2006).

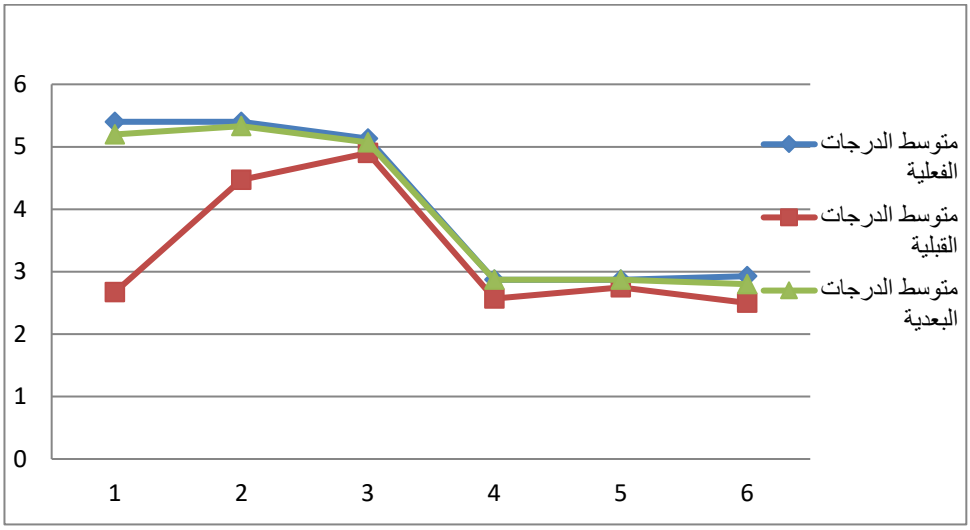
ولتحديد شكل الانحرافات في دقة المعايرة القبلية والبعدية في علاقتها بالأداء جرى الآتي:

1- حساب متوسط الدرجات الفعلية والمقدرة قبليًا والمقدرة بعديًا لدى فئتي الطلبة مرتفعي ومنخفضي الأداء في كل مشكلة على حدة.
2- رسم منحنى المعايرة لفئتي الطلبة مرتفعي ومنخفضي الأداء في المشكلات الست، وهو عبارة عن شكل بياني محوره الأفقي يعبر عن رقم كل مشكلة من المشكلات الست، ومحوره الرأسي يعبر عن الدرجة على

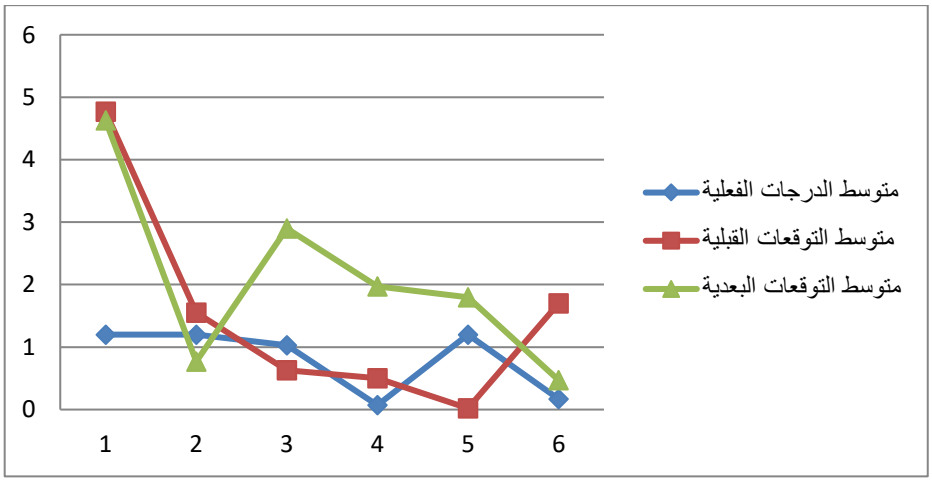
https://doi.org/10.20428/ijtd.v14i1.2309

المشكلات، مع العلم بأن الدرجة العظمى للمشكلات الثلاث الأولى تساوي (6) وللمشكلات الثلاث الأخيرة تساوي (3)؛ ثم جرى تمثيل متوسط الدرجات الفعلية والتوقعات القبلية والبعدية لكل مجموعة من الطلبة في كل مشكلة على حدة والمقارنة بينهما في انحرافات الدرجات المتوقعة في ضوء الدرجات الفعلية. بحيث إذا تطابقت درجات الطلبة المتوقعة مع الدرجات الفعلية وكانت على نفس الخط يعبر ذلك عن المعايرة التامة (Perfect Colibration) وتعبر النقاط التي تقع أعلى المعايرة التامة عن زيادة في الثقة (Over Confident) أما النقاط التي تقع أسفل المعايرة التامة تعبر عن نقص في الثقة (Under Confident) (Hacker, et al., 2008b).

وتقدم منحنيات المعايرة بوصفها طريقةً لتمثيل البيانات تمثيلات سهلة التفسير للكيفية التي تتغير بها الدقة عبر مستويات الأداء، بالإضافة إلى أنها توضح الكيفية التي تتغير بها زيادة أو نقص الثقة في الأحكام مع الأداء (Krug, 2007). ويوضح شكلا (3، 4) منحنيات المعايرة الخاصة بكل من مرتفعي ومنخفضي الأداء في المشكلات الست.



شكل (3): متوسط الدرجات الفعلية والتوقعات القبلية والبعدية للطلبة مرتفعي الأداء



شكل (4): متوسط الدرجات الفعلية والتوقعات القبلية والبعدية للطلبة منخفضي الأداء

ويتضح من شكل (3) أن متوسط التوقعات البعدية لدى فئة الطلبة مرتفعي الأداء متطابق تقريبًا مع المعايرة التامة والتي تقترب من الدرجات العظمى في معظم المشكلات، ومتوسط التوقعات القبلية يقع بانحرافات متفاوتة خلال المشكلات أسفل المعايرة التامة بحيث كانت أقصى الانحرافات في المشكلة الأولى والثانية، وبذلك كان النمط السائد هو (التقليل) وكانت المعايرة البعدية هي الأدق. ويتفق ذلك مع كل من (Cummings, 2006; Snyder, et al., 2011).

ويتضح من شكل (4) أن متوسط التوقعات البعدية لدى فئة الطلبة منخفضي الأداء يقع أعلى المعايرة التامة بانحرافات متفاوتة في كل المشكلات ماعدا المشكلة الثانية ومتوسط التوقعات القبلية يقع بانحرافات متفاوتة أعلى المعايرة التامة في المشكلة الأولى والثانية والرابعة والسادسة، وأسفل المعايرة التامة في المشكلة الثالثة والخامسة، وبذلك كان النمط السائد هو (المبالغة) حيث أظهر الأفراد منخفضي الأداء إفراطًا في الثقة. وقد يرجع ذلك إلى قصورهم الواضح في مهارات المراقبة، ويتفق ذلك مع ما أشار إليه (Ormrod, 2004) .

الخاتمة:

توجد صعوبات كثيرة تواجه تقييم مهارات ماوراء المعرفة، ونظرًا لكونها وعيًا أو عمليات داخلية ليست سلوكيات قابلة للملاحظة والقياس، ولذلك فقد اقترح كثيرًا من الباحثين استخدام التصميم متعدد الطرق (Multi method design) بحيث يجري فيه استخدام أكثر من أداةٍ في قياس تلك المهارات يُجمع فيها بين الجانب الكمي المتمثل في الاستبيانات، والجانب الكيفي المتمثل في البروتوكولات اللفظية المتزامنه لكي يتم الحصول على صورة كاملة عن عمليات ومهارات الأفراد.

ولكن كشفت نتائج كثير من الدراسات التي استخدمت التصميم متعدد الطرق عن ارتباطات منخفضة بين المقاييس المتزامنة والمقاييس غير المتزامنة وغالبًا ما كان يعزى ذلك إلى قصور الأدوات غير المتزامنه المستخدمة، في الوقت نفسه فالمقاييس غير المتزامنة تحتاج إلى كثير من الوقت والجهد ، بالإضافة إلى الإجراءات المعملية الحساسة، وهذا ما دفع الباحثة إلى السعي للوصول إلى منهجية صادقة وأدوات بديلة تتوافر فيها المعايير السيكومترية المعتمدة في المجال لقياس مهارات ما وراء المعرفة اعتمادًا على مدخل القياس المحدد بالمهمة؛ والذي يبنى في ضوء سياقات حقيقية محددة من خلال بيانات ومعلومات مشتقة من مصادر كيفية متعددة بحيث يجري تقليل مهددات صدقها إلى الحد الأدنى ومقارنتها بالأدوات المتزامنة في ضوء التصميم متعدد الطرق.

وقد كشفت النتائج المتعلقة بالسؤال الأول عن وجود ارتباطٍ موجبٍ ودالٍ إحصائيًا بين درجات الطلبة على القياسات المتزامنة لمهارات ما وراء المعرفة المتمثلة في البروتوكولات اللفظية المتزامنة ودرجاتهم على القياسات غير المتزامنة المتمثلة في كل من الاختبار متعدد الاستجابات، والاستبيان المحدد بالمهمة؛ كما أمكن التنبؤ بسلوك حل المشكلة من خلال المهارات المقاسة بالأدوات الثلاث ولكن كان الاختبار متعدد الاستجابات أكثرها إسهامًا في تفسير الأداء والتنبؤ به. وبذلك يمكن القول إنه من التدخل في تصميم المقاييس غير المتزامنة بحيث تكون على مهام محددة تقرِّب مؤشرات كل من المقاييس المتزامنة وغير المتزامنة بحيث تستهدف شيئًا واحدًا، فإذا كانت المقاييس غير المتزامنة مبنية بدقة وجرى اختبارها بطرق سليمة فقد تصبح بدائل معقولة وكافية للمقاييس المتزامنة المكثفة المعملية في قياس إستراتيجيات ومهارات الأفراد.

كما كشفت النتائج أيضًا أن درجات معايرة الطلبة البعدية كانت أكثر دقة من درجات المعايرة القبلية وكانت بمثابة مؤشرًا جيدًا ومرتبطًا مع مهارات المراقبة المقدرة من خلال الأدوات الثلاث التي تستهدف قياس مهارات ما وراء المعرفة، كما ارتبطت بالأداء وكان ذلك محكًا لصدق الأدوات.

https://doi.org/10.20428/ijtd.v14i1.2309

وقد كشفت النتائج المتعلقة بالسؤال الثاني أن كل من البروتوكولات والاختبار متعدد الاستجابات، والاستبيان المحدد بالمهمة، كان له قدرة على التمييز بين مهارات ما وراء المعرفية النوعية التي استخدمها الأفراد مرتفعو ومنخفضو الأداء في كل مشكلة على حدة، وبذلك يمكن القول إن تطوير المقاييس غير المتزامنة بحيث تكون على مهام محددة مع وضع كل الاحتمالات والمسارات المستخدمة في الحل في الاعتبار كان له دورٌ في زيادة دقة الأدوات وقدرتها على التمييز بين المهارات النوعية التي يستخدمها الأفراد مختلفو الأداء في المشكلات المختلفة.

بالإضافة إلى ذلك، فقد قدمت بيانات الآثار معلومات ذات قيمة عن مهارات الطلبة ونشاطها وتتابعاتها أثناء كل خطوة من خطوات الحل، كما وفرت معلومات عن أين ومتى وقعت معظم الحركات الزائدة غير الصحيحة، وحددت أي الخطوات تمثل عتبات (Thresholds) يمكن أن تفرق بين مجموعات الطلبة مرتفعي ومنخفضي الأداء، فضلاً عن أنها انسجمت مع البيانات التي جُمِعت من البروتوكولات.

كما كشفت النتائج أيضًا اختلاف نمط المعايرة القبلية والبعدية لدى مجموعتي الطلبة مرتفعي ومنخفضي الأداء، وقد اتسمت المعايرة البعدية بالدقة مقارنةً بالمعايرة القبلية، ووضحت منحنيات المعايرة كيف تغيرت الدقة عبر مجموعتي الطلبة ذوي المستويات المختلفة في الأداء.

ومن ذلك كشفت الدراسة عن كثير من النتائج المهمة التي ربما تغيِّر مسار الأبحاث المستقبلية.

توصيات:

1- ينبغي الاهتمام بإعداد المقاييس ذات المهام المحددة للتغلب على مشكلات تحيز الذاكرة، كما ينبغي تجنب المقاييس ذات النمط المتكرر من الاستجابة التي تخضع لكثيرٍ من العوامل التي تهدد صدق نتائجها.
2- ينبغي الاهتمام أثناء عمليات التحقق من صلاحية المقاييس والاختبارات بالمكوِّن الكيفي لعملية صدق المفردات، وتأكيد فكرة جمع الأدلة الكيفية مع الأدلة السيكومترية.
3- توظيف البرمجيات الحديثة لجمع كل الآثار التي يتركها الفرد أثناء انهماكه في المهام، والتي قد توفر كثيرًا من المعلومات ذات القيمة عن مسارات تفكيره الخفية، والتي تفيد في عمليات القياس.
4- الاهتمام بعمليات القياس الفوري للمهارات، مع تقديم الدعم المباشر الذي يحتاجه الفرد باستخدام البرمجيات الحديثة.
5- إعادة صياغة التعلم القائم على حل المشكلة اعتماداً على فهم طبيعة العمليات ومسارات الأفراد مختلفي الأداء أثناء حلهم للمشكلات.

بحوث مقترحة:

1- دراسة فعالية استخدام التصميم متعدد الطرق في قياس مهارات ما وراء المعرفة على مهام محددة ذات متطلبات مختلفة والمقارنة بينها.
2- دراسة الفروق بين فئات الطلبة مرتفعي ومنخفضي الأداء في ترتيب العمليات ما وراء المعرفية الدقيقة، ودراسة كيف يتغير هذا الترتيب عند مواجهة الأخطاء.
3- تطوير البرمجيات التي تساعد على ترجمة الإجراءات التي يقوم بها الفرد مباشرة إلى مهارات ذات مستويات مختلفة على حسب نوع الإجراء وترتيب حدوثه.
4- دراسة قدرة الآثار على التنبؤ بالأداء في سياق سلوك حل المشكلة والتنبؤ بالأداء وراء هذا السياق.

المراجع:

المصادر والمراجع العربية

الأعسر، صفاء. (1998). *تعليم من أجل التفكير*. القاهرة: دار قباء.

البنا، عادل السعيد (1996)، *برنامج للتدريب على إستراتيجيات تجهيز المعلومات بمساعدة الحاسب الآلي وأثره على تنمية سلوك حل المشكلة لدى طلاب كلية التربية*، رسالة دكتوراة غير منشورة، كلية التربية بدمنهور، جامعة الإسكندرية.

جروان، فتحي (1999)، *تعليم التفكير مهارات وتطبيقات*، العين: دار الكتاب الجامعي.

زكي، هناء محمد (2007)، *أثر ما وراء المعرفة وتحمل الغموض وخصائص المهمة على إستراتيجيات الأداء في حل المشكلات الرياضية*، رسالة ماجستير غير منشورة، كلية التربية، جامعة بنها.

عامر، أيمن محمد (2002)، *أثر الوعي بالعمليات الإبداعية والأسلوب الإبداعي في كفاءة حل المشكلات*، رسالة دكتوراة غير منشورة، كلية الآداب، جامعة القاهرة.

عبد القوي، مصطفى محمد (1993)، *أثر استخدام بعض المعالجات لتدريس إستراتيجيات حل المشكلة على تنمية أداء حل المشكلات الرياضية والاتجاه نحوه لدى تلاميذ المرحلة الثانوية*، رسالة دكتوراة غير منشورة، كلية التربية، جامعة الإسكندرية.

عبد القوي، مصطفى محمد (2006)، العلاقة بين وعي التلاميذ بالصف الأول من المرحلة الثانوية بالعمليات "ما وراء المعرفة" المصاحبة لحل المشكلة الرياضية وأدائهم فيها، *مجلة تربويات الرياضيات*، كلية التربية، جامعة بنها، (9)، (47-88).

محفوظ، سهير أنور (1985)، دراسة *تجريبية في تعلم سلوك حل المشكلة*، رسالة دكتوراة غير منشورة، كلية التربية، جامعة عين شمس.

هيبة، محمد أحمد (2007)، *صدق بروتوكولات حل المشكلات الرياضية وعلاقته بالذاكرة العاملة ونوع البروتوكول*، رسالة ماجستير غير منشورة، كلية التربية، جامعة عين شمس.

المصادر والمراجع الأجنبية

Allan, A.(1997). Begging the questionnaire: Instrument effect on readers response to a self-report checklist, *Language Testing*.*12*,133-156.

Baker, S. F. (2010). Calibration analysis within the cognitive and personality domains: Individual differences in confidence, accuracy, and bias. Doctoral dissertation, University of Southern Queensland, Retrieved February 15, 2012 from http://eprints.usq.edu.au/8856/.

https://doi.org/10.20428/ijtd.v14i1.2309

Baker,L.,&Cerro,L.C.(2000). Assessing metacognition in children and adults, *Educational Psychology Review.1*(1),3-30.

Bannert, M.,& Mengelkamp, C. (2008) . Assessment of Metacognitive Skills by Means of Instruction to Think Aloud and Reflect when Prompted. Does the Verbalisation Method Affect Learning?, *Metacognition Learning.3* (1), 39 – 50.

Bråten, I., & Samuelstuen, S.M. (2007). Measuring strategic processing: Comparing task specific self-Reports to traces, *Metacognition learning.2*(2),1-20.

Cascallar,E., Boekaerts,M.,&Costigan,T.(2006). Assessment in the evaluation of self-Regulation as aproces, *Education Psychology review.18*(1),297-306.

Cooper,M.M., Urena,S.S.,&Stevens,R.(2008). Reliable multimethod assessment of metacognition use in chemistry problem solving,*Chemistry education research and practice*,9,18-24.

Cromley, J.G.,& Azevedo, R. (2007). Self-report of Reading Comprehension Strategies: What are we measuring?, *Metacognition Learning*, *1* (3), 229 – 247.

Cromley,J,&Azevedo,R.(2011). Measuring strategy use in context with multiple choice items, *Metacognitiona and learning, 6*(2),155-177.

Cummings, A. M. (2006). The use of item response theory to assess adults' postdiction accuracy, *Educational Psychology and Special Education Dissertations*. Paper 42: Retrieved August19, 2011 from http://digitalarchive.gsu.edu/epse_diss/42.

Desoete,A.(2007). Evaluating and improving the mathematics teaching-learning process through metacognition, *electronic journal of research in educational psychology*,*5*(3),705-730.

Desoete,A.(2008). Multi-method assessment of metacognitive skills in elementary school children :how you test is what you get, *Metacognition Learning, 3*,189–206.

Dompnier, B., Pansu. P.,& Bressoux, P. (2006). An Integrative Model of Scholastic Judgments: Pupils' Characteristics, Class Context, Halo Effect and Internal Attributions, *Journal of Psychology of Education*, *43* (2), 119 – 133.

Ericsson, K.A., & Simon, H.A. (1980). Verbal reports as data, *Psychological Review*,*87*,215-251.

Hacker, D. J., Bol, L., & Keener, M. C. (2008b). Metacognition in education: A focus on calibration, In: Dunlosky, J &. Bjork, R. A. (Eds.), Handbook of metamemory and memory (pp. 429-455). New York: Psychology Press.

Hadwin, A.F, Winne, P.H., Stockley, D.B., Nesbit, J.C., & Woszczyna, C. (2001). Context moderates students self-reports about how they study, *Journal of educational Psychology*,93(3),477-487.

Hadwin,A.F, Nesbit,J.C., Noel,D.J., Code,J.,& Winne, Ph.H. (2007). Examing trace data to explore self-regulated learning, *Metacognition learning*,*2*,107-124.

Jarvela, S., Vermans, M.,& Leinonen, P. (2008). Investigating Student Engagement in Computer-Supported Inquiry: A Process- Oriented Analysis, *Soc Psychol Edu*, *11*, 299 – 322.

Krug, K. (2007). The relationship between confidence and accuracy: Current thoughts of literature and a new area of research, *Applied Psychology in Criminal Justice*. *3*, 7-41.

Long,D.L.,&Bourg,T.(1996). Thinking aloud: Telling a story about astory , *Discourse Processes,21*(3),329-339.

Magliano, J.P., Trabasso, T.& Graesser, A.C.(1999). Strategic processing during comprehension, *Journal of Educational psychology*, *91*(4),615-629.

Mengelkamp, C., & Bannert, M. (2010). Accuracy of confidence judgments: Stability and generality in the learning process and predictive validity for learning outcome, *Memory & Cognition, 38*, 441-451.

Messick,S.(1995). Validity of psychological assessment: Validation of inferences from persons responses and performances as scientific inquiry into score meaning, *American Psychologist*,*50*(9),741-749.

Ormrod, E.J. (2004) . *Human Learning: Metacognition Self-Regulated Learning, and Study Strategies*, New Jersey: Pearson Merrill Prentice Hall.

Richardson,.T.E. (2004). Methodological issues in questionnaire-based research on student learning in higher education, *educational psychology review*,*16*(4),347-358.

Russo, J.E., Johnson, E.J.,& Stephens, D.L. (1989). The validity of verbal protocols, *Memory&Cognition*,*17*(6),759-769.

Sampaio, C., & Brewer, W. F. (2009). The role of unconscious memory errors in judgments of confidence for sentence recognition, *Memory & Cognition*, *37*, 158-163.

Samuelstuen, M.S., & Bråten,I. (2007). Examing the validity of self reports on scales measuring students strategic processing, *British journal of educational psychology*,77,351-378.

Santiago, G.(2008). Design and validation of multimethod assessment of metacognition and study of the effectiveness of metacognitive interventions, *Doctoral Dissertation*, Clemson University.

https://doi.org/10.20428/ijtd.v14i1.2309

Schellings, G.(2011). Applying learning strategy questionnaires: Problems and possibilities, *Metacognition and learning, 6*(2),91-109.

Secolsky, C., Wentland, E., & Denison, B. (2011). The Need for Documenting Validation Transactions : A Qualitative Component of the Testing Validation Process, *Qual Quant*, © Spring Science + Business Media.

Snyder, K. E., Nietfeld, J. L., & Linnenbrink-Garcia, L. (2011). Giftedness and Metacognition: A Short-Term Longitudinal Investigation of Metacognitive Monitoring in the Classroom, *Gifted Child Quarterly*, *55*(3), 181-193.

Sternberg, R. J. (1992). Ability tests, measurements, and markets, *Journal of Educational Psychology*, *84*(2), 134–140.

Sternberg, R.J. (1996). *Cognitive Psychology*, 4th ,Londen: Harcourt Prace.

Stone,J.N.(2000). Exploring the relationship between calibaration and self-regulated learning, *educational psychology review*,*12*(4),437-475.

Van Haut – Wolters, B.H.A.M. (2009). Assessing Self-directed Learning, In : Simons, P.R.J., & Vanderlinden, J., & Duffy, T. (Eds.). New Learning, (pp. 83 – 101). Dordrecht , Kluwer.

Veenman, M.V.J.,Bernadette, M.A.H., Wolters, V.H.,&Afflerbac, P. (2006). Metacognition and learning: conceptual and methodological considerations, *Metacgnition learning*,*1*,3-14.

Veenman,M.V.J.(2005). The assessment of metacognitive skills: What can be learned from multi-method design, *Learning & Instruction*,*14*(1), 89 – 109.

Veenman,M.V.J.(2007). The assessment and instruction of self-regulation in computer-based environments: a discussion, *Metacognition learning*, *2*,177-183.

Veenman,M.V.J.(2011). Alternative assessment of strategy use with self-report instruments: a discussion, *Metacognition learning*, *6*, 205-211.

Veenman,M.V.J.(2013). Assessing metaognitive skills in computerized learning environments, In: Azevedo, R., & Aleven,V. (2013). International Handbook of metacognition and learning technologies, (pp.157-168). Retrieved from:Springer International Handbooks of education.

Veenman,M.V.J.,Prins,F.J.,&Verheij, J.(2003). Learning styles: self-reports versus thinking-aloud measures*, British Journal of educational psychology*,*73*,357-372.

Winne, P. H., Jamieson-Noel, D.,& Muis, K.R. (2002). Exploring students' calibration of self-reports about study tactics and achievement, *Contemporary Educational Psychology,27*(4), 551–572.