التحقق من صدق الحقيبة التقويمية للطالب (البورتفوليو) في التقويم الصفي في الصف الأول الثانوي في مملكة البحرين

راشد حمّاد الدوسري*

جامعة البحرين، مملكة البحرين

استُلم بتاريخ: ٢٠١٣/١٠/٢٥ عُدل بتاريخ: ٢٠١٣/١٢/٦ قُبل بتاريخ: ٢٠١٣/١٢/٣١

هدف هذا البحث إلى التحقق من صدق الحقيبة التقويمية للطالب (البورتفوليو) في المرحلة الثانوية باستخدام نموذج ميسيك الموحد ذي المكونات الستة (صدق المحتوى، والصدق البنائي، وصدق القابلية للتعميم، والصدق الخارجي، والصدق الموسع،وصدق العواقب)، والكشف عن ممارسات المعلمين في الحقيبة التقويمية واستخدامها في تقويم تحصيل الطالب. طُبق البحث في مدرستين ثانويتين في البحرين، واحدة للبنين وأخرى للبنات. شارك في البحث ٣٠ معلماً ومعلمة تم اختيارهم عشوائياً من مجتمع الدراسة المكون من ١٣٠ معلماً ومعلمة. استُخدمت أداتان في البحث، الأولى استبيان مكون من ٢٠ فقرة يتعلق بممارسات المعلمين في التقويم الصفي. والأخرى عبارة عن استمارة مكونة من ٢٢ عنصراً من العناصر المكونة للحقيبة التقويمية للطالب موزعة على خمسة محاور رئيسية. بينت النتائج أن أكثر ممارسات المعلمين في التقويم باستخدام الحقيبة التقويمية للطالب كانت منصبة على الاهتمام بالأنشطة والمشروعات، ووضع تقدير عام للطالب في الحقيبة التقويمية. كما أشارت النتائج أيضاً إلى وجود أدلة على صدق التقويم بالحقيبة التقويمية وفق نموذج ميسيك، وذلك من خلال الارتباطات البينية ومعاملات الثبات بين مكونات الحقيبة التقويمية للطالب حيث تراوحت معاملات الثبات للمحاور الخمسة بين ٠,٨٤ و ٠,٩٢. وتراوحت معاملات الارتباط البينية للمحاور الخمسة بين ٠,٨٥ و ٠,٦٣. كما بينت النتائج أيضاً أن ٤٧٪ من التباين في علامات الطالب في الحقيبة التقويمية يُعزى إلى تقديرات المحكمين على ملاءمة عناصر الحقيبة التقويمية. و بينت النتائج أيضاً وجود دليل على الصدق الخارجي للحقيبة التقويمية تمثل في العلاقة بين علامات الطالب في الحقيبة التقويمية والمعدل التراكمي للطالب (ر = ٠,٧٥)، وبين علامات الطالب في الحقيبة وعلامات الطالب في الاختبار الوطني (ر = ٠,٦٨). وفي ضوء نتائج الدراسة قدم الباحث مجموعة من التوصيات.

الكلمات المفتاحية: نموذج ميسيك الموحد، الصدق، الحقيبة التقويمية للطالب، المرحلة الثانوية، التقويم الصفي.

Investigating the Validity of Student's Portfolio in Classroom Assessment at the First Secondary Class in the Kingdom of Bahrain

Rashed H. Aldoseri*

University of Bahrain, Kingdom of Bahrain

This research aimed to investigate the validity of the secondary student's portfolio by utilizing Messick's six-facet unified model of validity (content, structure, generalizability, external, substantive, consequences), as well as exploring teachers' practices in using student's portfolio for assessing students' achievement. The research was applied in two secondary schools in Bahrain, one for males and one for females. A random sample of 30 teachers from 130 teachers participated in the study. Two instruments were used: a 20-item questionnaire pertaining to teachers' classroom assessment practices; and a 22-item form about the content of students' portfolio, with five dimensions. The results revealed that most teachers' assessment practices of the portfolio focused on students' activities and projects. Also, assigning grades to the students was another priority of teachers. Evidence of validity of students' portfolio in assessment according to Messick's model was found. Inter-correlations ranged between 0.63 to 0.85; and reliability coefficients ranged from 0.84 to 0.92 among the 5 dimensions of the student portfolio. Regarding the generalizability validity, 47% of the variance of students' scores in the portfolio is attributed to raters' judgments about the relevance of portfolio components. Moreover, the results indicated evidence of the generalizability validity of the portfolio as depicted in the relationship between students' portfolio scores and students' GPA (r = 0.75) and between students' portfolio and students' scores in the national exam (r = 0.68). The researcher provided suggestions for further research.

Keywords: Messick's unified model, validity, student's portfolio, secondary stage, classroom assessment.

*rashida882@gmail.com

خضعت جودة التعليم الصفي وتقويم تحصيل الطالب في العقدين الماضيين إلى المزيد من التدقيق والمراقبة؛ وذلك لاعتقاد الكثير من التربويين أن المعلم يفتقر إلى المعرفة الضرورية، والمهارات التدريسية الأساسية، وطرق تقويم تحصيل الطلبة بشكلٍ صحيح ودقيق (Levine, 2006). وقد استجابت برامج تعليم المعلمين وتدريبهم قبل الخدمة وفي أثنائها لمتطلبات المساءلة التربوية Accountability عن جودة التعليم الصفي للمعلم، وجودة أدوات تقويم التحصيل الدراسي للطلبة، وذلك من خلال بناء نُظُمٌ جديدة (Bredo, 2005). وقد صُممت نظم التقويم الجديدة تلك لقياس نواتج برامج تدريب المعلم وتأهيله، مثل مدى إتقان المعلم مهارات التعليم الصفي، وبناء أدوات التقويم واستخدامها، وتحليل نتائجها؛ وذلك بديلاً عن التركيز على عدد المقررات التي درسها المعلم، أو عدد الساعات المعتمدة التي اجتازها في كليات المعلمين (Ma & Rada, 2005). ويُعدُ هذا التحول - من التركيز على مدخلات البرامج التربوية لإعداد المعلمين وتطويرهم إلى التركيز على مخرجات تلك البرامج - ظاهرة جديدة لم تعهدها النُظُم التربوية من قبل (Brookhart, 2011).

ورغم قيام وزارات التربية والتعليم وكليات المعلمين ببناء استراتيجيات ومعايير لقياس وتقويم فاعلية التعليم الصفي للمعلم، وأدوات التقويم التي يستخدمها، فإنه لا يوجد إجماع لدى التربويين بشكلٍ عام، والمتخصصين في القياس والتقويم التربوي بشكلٍ خاص، على أفضل طريقة لقياس وتقويم أداء المعلم، ومهاراته، ومعرفته، يمكن أن تؤدي إلى تعليم صفي أفضل، وتحصيل أجود، ناهيك عن أن تلك الاستراتيجيات والمعايير لم تتحقق من صدق أدوات التقويم الصفي (Burns & Haight, 2005). ولعل الحقيبة التقويمية للطالب (البورتفوليو) Portfolio إحدى أدوات التقويم الصفي التي لم يتم التحقق من صدقها على نطاقٍ واسع في تقويم التحصيل الدراسي للطالب (Birenbaum, 2007).

خلفية الدراســـــــة

مفهوم الحقيبة التقويمية للطالب (البورتفوليو) ومكوناتها وأغراضها: استخدمت الحقيبة التقويمية منذ وقتٍ طويل في مجالات مختلفة مثل الفنون والعمارة مقارنة باستخدامها في المجال التربوي الذي لا يزيد عن عقدين من الزمن (Davis & Ponnamperuma, 2005). الحقيبة التقويمية للطالب تمثل تجميعاً لصيغ مختلفة من الأدلة والبراهين عن تحصيله ونواتج تعلمه، من خلال أعماله في المقرر خلال الفصل الدراسي، أو بعبارةٍ أخرى، يمكن القول إن الحقيبة التقويمية للطالب عبارة عن تجميع مقصود وهادف ومنتظم لعينات من أعماله في المقرر الدراسي توثق تحصيله الدراسي وتقدمه التعليمي في مدة زمنية محددة خلال الفصل الدراسي (Strudler & Wetzel, 2005).

وتتكون محتويات الحقيبة التقويمية من نتائج الطالب في الاختبارات، ونشاطاته الصفية، ومشروعاته، وبحوثه، وإنجازاته العلمية، والشهادات التي حصل عليها في المسابقات المدرسية، وصور، ومقاطع مصورة (فيديو) توثق نشاطاته وأعماله، وتعليقاته وملاحظاته على أعماله، وتأمله الذاتي Reflective thinking في أعماله، ونقده الذاتي لها، ومجموعة من مقاييس التقدير وقواعد التصحيح Scoring rubrics (Goodson, 2007).

ولعل المحرك الأساس لاستخدام الحقيبة التقويمية في تقويم تحصيل الطالب وتعرف مستوى أدائه هو التأكد من مدى تحقيق الأهداف التعليمية الصفية، وأهداف التعلم Learning targets . وقياس التحصيل الدراسي بطرق تقويم متعددة، وتُستخدم الحقيبة التقويمية لأغراض كثيرة، منها متابعة تحصيل الطالب طوال الفصل الدراسي، وتعرف جوانب القوة والضعف في تعلمه، وتشخيص نقاط الضعف وعلاجها أولاً بأول، وتعزيز تعلم الطالب من وقتٍ إلى آخر، وحصول المعلم على تغذية راجعة عن تعلم الطالب وتحصيله، وكذلك عن تعليم المعلم لغرض مراجعة أو تغيير أساليب تعلم الطالب، وطرق تعليم المعلم الصفية، وتحسين طرق تقويم تحصيل الطالب (Donelly, 2005).

يمر استخدام الحقيبة التقويمية وتطبيقها في تقويم تحصيل الطالب عبر خمس مراحل أساسية (Alimemaj & Ahmetaj, 2010)، وهي (١)، تجميع الأدلة الكافية من أعمال الطالب التي تمثل

تحصيله ومستوى أدائه ونواتج التعلم. (٢). التأمل الذاتي للطالب في أعماله وتقييمه الذاتي لها ونقدها. (٣). تقويم صدق (صحة) البراهين أو الأدلة المرتبطة بتحصيل الطالب. (٤). دفاع الطالب عن مدى صدق الأعمال التي وضعها في حقيبته التقويمية، عن طريق مقابلة المعلم للطالب بشكلٍ فردي. (٥). قرار التقويم. وذلك من خلال مقاييس تقدير. وقواعد تصحيح لتقويم البراهين والأدلة في أعمال الطالب.

جوانب القوة والضعف في الحقيبة التقويمية للطالب: تشير الدراسات التي أجريت حول التقويم باستخدام الحقيبة التقويمية للطالب إلى الكثير من جوانب القوة فيها، والتي يمكن إجمالها في التالي: (Darling-Hammond, 2006; Derham & Diperna, 2007; Goodson, 2007; NBPTS, 2006)

١. تقويم التفكير النقدي للطالب وتعزيزه.
٢. تجعل تعلم الطالب مُوجّهاً ذاتياً، ومدعوماً من قبل أقرانه.
٣. تسهل عملية التفكير التأملي والتقويم الذاتي.
٤. تسمح للطالب بتطبيق أنماط مختلفة للتعلم.
٥. تسمح بمراقبة تعلم الطالب وتقويم تحصيله خلال الفصل الدراسي.
٦. تساعد الطالب في تطبيق معرفته النظرية في مواقف طبيعية في الحياة اليومية.
٧. تأخذ في الحسبان تقديرات أكثر من مُقدّر أو مُقوّم.
٨. تستخدم طرق تقويم متعددة لتقويم تحصيل الطالب بوساطة أعماله.
٩. تتمتع بمستوىً عالٍ من الصدق الظاهري، وصدق المحتوى، وصدق السمة.
١٠. تشجع على الإبداع، وحل المشكلات، واستخدام مهارات التعلم ما وراء المعرفية Meta-cognitive .
١١. تستخدم في التقويم التكويني طوال الفصل الدراسي، وفي التقويم الختامي عند نهاية الفصل الدراسي.
١٢. تجمع بين التقويم الكمي والنوعي، وبين التقويم الموضوعي والذاتي.
١٣. يمكن استخدامها لتقويم الاتجاهات والتطور في شخصية الطالب تجاه المعرفة والتحصيل الدراسي.
١٤. تقدم للمعلم معلومات مهمة لتشخيص جوانب القوة والضعف في تعلم الطالب.

أما بالنسبة لجوانب الضعف في استخدام الحقيبة التقويمية للطالب أداة للتقويم، فيمكن تلخيصها في النقاط التالية (Birenbaum, 2007; Bredo, 2005; Levine, 2006; Porter, 2002; Zeichner & Wray, 2001):

١. عندما يستخدم المعلم الحقيبة التقويمية في التقويم الختامي يتردد الطلبة في الكشف عن جوانب الضعف في تعلمهم وتحصيلهم الدراسي طوال الفصل الدراسي.
٢. الحقيبة التقويمية للطالب وثيقة شخصية، وقد يثير استخدامها في التقويم الصفي بعض القضايا الأخلاقية، والقضايا المتعلقة بخصوصية الطالب نفسه.
٣. تبرز بعض الصعوبات في التقويم عندما يريد المعلم التحقق مما إذا كانت الأعمال الموجودة في الحقيبة التقويمية من عمل الطالب نفسه، أو من عمل غيره.
٤. الحقيبة التقويمية تستغرق وقتاً طويلاً لبنائها، وتقويم محتوياتها.
٥. عملية التقويم باستخدام الحقيبة التقويمية قد ينتج عنها انخفاض في قيمة معامل ثبات ما بين المقدرين (المصححين)، وخاصة إذا لم يتم إعداد مقاييس التقدير وقواعد التصحيح بطريقة صحيحة وصادقة، أو تم استخدامها من قبل مقدرين غير مُدرّبين.

صدق الحقيبة التقويمية للطالب: يمكن تعريف الصدق بشكلٍ عام بأنه مدى قدرة أداة القياس على قياس ما هي مصممة لقياسه. ويُعرّفُ الصدقُ أيضاً بأنه الدرجة التي يستطيع بوساطتها كل من البرهان (الدليل) Evidence والنظرية دعم التفسيرات المستهدفة من درجات الاختبارات من قبل من يتوقع استخدامهم نتائج تلك الاختبارات (Popham, 2008). هذا التعريف للصدق ينطبق على كل أدوات القياس والتقويم، وتقويم الأداء، والتي من ضمنها الحقيبة التقويمية للطالب في تقويم الأداء (Messick, 1994a, 1994b, 1995a, 1995b). وهو التعريف الذي يتبناه الباحث في هذه الدراسة.

ويرى بيرنبوم (Birenbaum, 2007) أن جودة التقويم الصفي تحددها مدى ملاءمة وفائدة الاستدلالات التي يستخلصها التربويون من تفسيرات نتائج التقويم، مثل درجات الاختبار الصفي. إن المفهوم التقليدي لأنواع الصدق يشير إلى أن الصدق تم تقسيمه إلى ثلاثة جوانب منفصلة، ولكنها مترابطة سيكومترياً في الأساس، وهي: صدق المحتوى، والصدق المحكي (ويشمل الصدق التلازمي، والصدق التنبؤي)، وصدق السمة ويشمل الصدق التقاربي (Convergent)، والصدق التمايزي (Divergent) (Sherry & Bartlett, 2005). أما في مجال التقويم باستخدام الحقيبة التقويمية للطالب فيُعرّفُ الصدقُ بأنه مدى قدرة الحقيبة التقويمية على تقويم نواتج التعلم في المنهج الدراسي (Alimemaj & Ahmetaj, 2010).

يتمتع التقويم باستخدام الحقيبة التقويمية للطالب بصدقٍ ظاهريّ عالٍ، بسبب قدرتها على تقويم الأداء كما هو في الواقع باستخدام أدوات تقويم متعددة كمية (مثل التقديرات)، ونوعية (مثل التعليق الوصفي للطالب على أعماله، وتعليق المعلم على أعمال الطالب). وإذا نُفذت عملية بناء الحقيبة التقويمية بصورةٍ صحيحة، فإن عملية التقويم ستكون مُحكمة Robust. وستتضمن عينات ممثلة لأعمال الطالب وتحصيله الدراسي وجوانب تعلمه، و ستكون ممثلة محتوى المنهج الدراسي عبر مدىً واسعٍ من نواتج التعلم؛ مما يُعدُ مؤشراً على صدق المحتوى للحقيبة التقويمية. ويمكن للحقيبة التقويمية أيضاً أن تكون مؤشراً على صدق السمة التي يتم تقويمها، عبر التفكير التأملي للطالب في قدراته (Alimemaj & Ahmetaj, 2010).

أما فيما يتعلق بكيفية التحقق من الصدق التلازمي والصدق التنبؤي للتقويم باستخدام الحقيبة التقويمية، فيمكن أن يسبب ذلك إشكالاً لدى المعلم، ذلك أن افتقار المعلم إلى أدوات قياس أخرى لتقويم أداء الطالب وتحصيله يثير بعض المشكلات في قياس الصدق التلازمي، ولا يؤدي كذلك إلى التحقق من استخدام الحقيبة التقويمية في التنبؤ بأداء الطالب في المستقبل المنظور. كما أنه من الصعب أحياناً التحقق من مكونات الحقيبة التقويمية للحصول على دليل أو برهان يؤكد أن تلك المحتويات من عمل الطالب نفسه؛ مما يُعدُ تهديداً لصدق التقويم بالحقيبة.

نموذج ميسيك الموحد لصدق السمة: يشير بيرنبوم (Birenbaum, 2007)، وجودسون (Goodson, 2007) إلى أن التقسيم التقليدي للصدق إلى ثلاثة تصنيفات قاد المتخصصين في القياس والتقويم إلى الاعتقاد بأن الصدق ثلاثة أنواع متميزة بعضها عن بعضها، مما أدى إلى تركيز عملية المصادقة (التحقق من الصدق) في الاختبار نفسه بدلاً من التركيز على الاستدلالات Inferences المفترض استخلاصها من درجات الاختبار. كما أدى هذا التصور عن الصدق إلى التركيز على السلوك البشري الملاحَظ دون الاهتمام بشكلٍ كافٍ بالسمات النظرية الكامنة. ولعل أكثر الدراسات حول الصدق اعتمدت على هذا التقسيم التقليدي للصدق، كما أشار إلى ذلك بورتر (Porter, 2007).

وضع ميسيك (Messick, 1994a, 1994b, 1995a, 1995b) نموذجاً أو تصوراً بديلاً لفهم الصدق وتوصيفه بشكلٍ يختلف عن التقسيمات التقليدية التي تمت الإشارة إليها، ووضع الصدق تحت مظلة واحدة سماها "المفهوم الموحد لصدق السمة". يتكون نموذج ميسيك من ستة وجوه Facets مترابطة فيما بينها، وهي: (أ). صدق المحتوى، (ب). الصدق الموسَّعْ (الحقيقي/الواقعي) Substantive، (ج). الصدق البنائي Structural ، (د). الصدق الخارجي External، (هـ). صدق القابلية للتعميم Generalizability ، (و). صدق النتائج (العواقب) Consequential validity .

ضمن هذا المنظور الموحد للصدق قام ميسيك بتوسيع صدق المحتوى ليشمل وجهين من الصدق هما صدق المحتوى، والصدق الحقيقي/الواقعي؛ حيث يركز صدق المحتوى على المطابقة بين محتوى الاختبار ومحتوى المجال المعرفي المراد تقويم تحصيل الطالب فيه، بينما يركز الصدق الحقيقي/الواقعي على المطابقة بين العمليات العقلية (المعرفية) وعمليات استجابة الطالب لأداة التقويم.

الطريقة الأكثر شيوعاً واستخداماً للتحقق من صدق المحتوى (Birenbaum, 2007; Porter, 2002) ضمن نموذج ميسيك هي مواءمة المحتوى المعرفي للمقرر الدراسي مع المنهج الدراسي بوساطة خبراء في المحتوى المعرفي للمقرر المستهدف بعملية المصادقة. ويتطلب ذلك تحليل محتوى مهمات الأداء في الحقيبة التقويمية للطالب، وفي المنهج الدراسي ضمن المجال المعرفي للمقرر. ويعتمد التحقق من صدق المحتوى في هذه الحالة على أحكام الخبراء على حدود السمة المقاسة، والمنهج الدراسي، والمهارات والمحتوى اللذين تقيسهما مهمات الأداء في الحقيبة التقويمية.

يمكن التحقق من الصدق الموسع من خلال البحث عن الدليل أو البرهان الذي يوضح أن مهمات الأداء في الحقيبة التقويمية تقود الطالب إلى الانخراط في العمليات المعرفية Cognitive processes المستهدفة، وتقليل تأثير العوامل غير المرتبطة بالسمة Construct-irrelevant factors إلى الحد الأدنى. ويرتبط الصدق البنائي بمدى كفاية وملاءمة التقديرات، وعمليات القياس والمقاييس المستخدمة. أما الصدق المرتبط بقابلية التعميم، وقابلية نتائج القياس للتكرار، فيُعبّرُ عن ثبات التقديرات ونتائج التقويم في مهمات الأداء في الحقيبة التقويمية، ويتم التحقق منه عبر ثبات النتائج عند مستويات متعددة، ومدى الترابط الداخلي بين مكونات الحقيبة التقويمية. الصدق الخارجي يمكن التحقق منه بواسطة الكشف عن وجود علاقة ارتباط بين الدرجات (التقديرات) التي يحصل عليها الطالب في مكونات الحقيبة التقويمية، والمتغيرات الخارجية المرتبطة بمهمات الأداء، مثل الطلبة الذين يحصلون على درجات (تقديرات) عالية بوصفها مؤشراً للإتقان في الحقيبة التقويمية، فهؤلاء تكون درجاتهم عالية في مؤشرات الإتقان لوسائل التقويم الأخرى كالاختبارات الوطنية.

أما صدق العواقب (النتائج) فيمكن التحقق منه بوساطة الكشف عن وجود آثار إيجابية مقصودة استهدفتها الحقيبة التقويمية، وآثار سلبية غير مقصودة ولكنها محتملة الحدوث جراء التقويم بالحقيبة التقويمية للطالب؛ وما يترتب على كلا النوعين من الآثار من قرارات تمس تعلم الطالب وتحصيله الدراسي، وتنبثق عن الجوانب التي سيتم استخدام نتائج الطالب فيها، مثل تقرير نجاح الطالب، أو رفعه إلى صف أعلى، أو تعزيز تعلمه، أو علاج جوانب الضعف لديه (Yao, Thomas, Nickens, Downing, Burkett, & Lamson, 2008).

ويُلاحظ أن نموذج ميسيك الموحد للصدق قد دمج أنواع الصدق التقليدية (الصدق التلازمي، والصدق التنبؤي، والصدق التمايزي Discriminant ، والصدق التقاربي Convergent) ضمن وجه مفرد Single facet يمثل الصدق الخارجي، والذي يعتمد في الأساس على الصدق التمايزي والصدق التقاربي (Birenbaum, 2007).

ورغم ما يشاع عن صدق الحقيبة التقويمية للطالب في تقويم تحصيله وأنها أداة تقويم حقيقية (أصيلة) Authentic في التعليم الصفي (Darling-Hammond, 2000; 2006)؛ فإن عملية التحقق من ذلك الصدق (أو المصادقة Validation)، لم يتم بحثها بتوسع، ولم تتم دراستها في مراحل دراسية مختلفة، وصفوف دراسية مختلفة (Derham & Diperna, 2007; Reis & Villaume, 2002). ويرجع ذلك إلى مجموعة من المعوقات تعرقل عملية التقويم باستخدام الحقيبة التقويمية، مثل الوقت الطويل الذي يتطلبه التقويم بالحقيبة التقويمية من جانب المعلم والطالب، وعدم قدرة المعلم على التحكم في بيئة التعلم الصفية، فقد يتردد المعلم في منح الطالب الفرصة لتحمل مسئولية تعلمه في المقرر من خلال الحقيبة التقويمية، عبر التحكم في أعماله ومهمات الأداء التي يضعها في الحقيبة؛ وذلك لأن الطالب لم يتعود هذه الأمور من قبل. كما أن الوعي الذاتي للطالب بمسئولية تعلمه وتفكيره وتأمله في مهمات الأداء والأعمال في حقيبته التقويمية من آنٍ إلى آخر غير موجود في المدارس الثانوية بوصفه إجراءً علمياً، وخاصة فيما

يتعلق بتقويم عمليات التفكير لدى الطالب، وتنظيمها، والتحكم فيها (Gadbury-Amyot, Kim, Palm, Mills, Noble, & Overman, 2003).

تُعدُ دراسة بيرنز وهايت (Burns & Haight, 2005) من الدراسات القليلة التي تصدت لهذا الموضوع، فقد قاما بفحص ومراجعة مجموعة من الحقائب التقويمية للطلبة لغرض التحقق من الصدق التلازمي، وذلك من خلال إيجاد الارتباط بين درجات الطالب في الحقيبة التقويمية ونتائجه في النشاطات الأخرى والاختبارات القصيرة، كما قام الباحثان بإيجاد مقدار الارتباط بين درجات الطلبة في الحقيبة التقويمية ودرجاتهم النهائية في المقررات في نهاية الفصل الدراسي؛ للتحقق من الصدق التنبؤي للحقيبة التقويمية، حصل الباحثان على معامل ارتباط مقداره ٠.٦٠، دليلاً (متوسط القوة) على الصدق التلازمي، وآخر مقداره ٠.٥٣، دليلاً (متوسط القوة) على الصدق التنبؤي للحقيبة التقويمية، أشار الباحثان كذلك إلى أن معامل ثبات ما بين المقدرين (المصححين) Inter-rater reliability لمكونات الحقيبة التقويمية كان عالياً جداً (٠.٩١).

كما حصل بيرنبوم (Birenbaum, 2007) على دليل قوي على الصدق البنائي في الحقيبة التقويمية (ر= ٠.٨٥ – ٠.٨٩) في دراسته التي طبقها على عينة من طلبة الصف الحادي عشر (٤٠ طالباً)، ولكن وجدت دراسات أخرى (Guerrero, 2000; Porter, 2007) دليلاً (متوسط القوة) على الصدق الخارجي (٠.٥٣ – ٠.٥٧) للحقيبة التقويمية، وكان دالاً إحصائياً (ألفا= ٠.٠١).

أما دراسة درهام وديبيرنا (Derham & Diperna, 2007) حول صدق التقويم بالحقيبة التقويمية، فقد طبقاها على ٣٠ طالباً لمعرفة العلاقة بين مستوى الطلبة في أعمالهم في الحقيبة التقويمية وبين درجاتهم في الاختبار الفصلي، وذلك في القراءة والكتابة والرياضيات، ومعدلهم التراكمي الفصلي، توصل الباحثان إلى وجود معاملات ارتباط دالة إحصائياً (عند ألفا=٠.٠٥)، وهي ر = ٠.٣٩، ر= ٠.٣٤، ولكنها ضعيفة، كما وجد الباحثان أن متوسط قيمة كابا – كوهين Cohen's kappa كان K = ٠.١٧ .

وفي دراسة أخرى لأبرامي وباريت (Abrami & Barret, 2005) حول صدق الحقيبة التقويمية للطالب طُبقت على جميع المراحل الدراسية، توصل الباحثان فيها إلى نتائج متباينة، فقد وجدا انخفاضاً في قيم ثبات ما بين المقدرين على الأعمال كلها في الحقيبة التقويمية؛ رغم أن متوسط معامل ارتباط بيرسون (ر) في كل الأعمال في الحقيبة كان ٠.٥٦ للدرجة الكلية، و ٠.٤٥ للمحور الواحد في الحقيبة (قراءة، كتابة، ...)، و ٠.٣٥ لمتوسط درجات الفقرات كلها؛ وذلك في الرياضيات فقط، وللحصول على دليل أو برهان على صدق التقويم بالحقيبة التقويمية، قام الباحثان كذلك بحساب معامل ارتباط بيرسون (ر) بين درجة الطالب على أعماله في الحقيبة التقويمية ككل وبين الاختبار الوطني في مراحل دراسية مختلفة، ووجدا أن معاملات الارتباط تراوحت بين ٠.٤٣ و ٠.٦١.

أما نايزر (Naizer, 1997) فقد فحص صدق الحقيبة التقويمية في مقرري الرياضيات والعلوم بالمرحلة الثانوية، وأوضح التحليل الإحصائي للصدق المتعلق بقابلية تعميم النتائج الذي أجراه نايزر أن التباين المرتبط بأثر المقدر (المصحح) كان ضئيلاً جداً، رغم وجود درجة معينة من عدم الاتفاق بين المقدرين حول الترتيب النسبي لأهمية الموضوعات في الحقيبة التقويمية، فيما يتعلق بالصدق التلازمي للحقيبة التقويمية، فقد تحقق منه نايزر بوساطة التحليل التمايزي Discriminant analysis وذلك بتصنيف الحقائب التقويمية إلى مجموعات طبقاً للدرجة الكلية التي حصل عليها الطلبة في كل منها، اعتمد نايزر على مجموعة من المتغيرات للتنبؤ بالصدق المستهدف، مثل درجات الطالب في الاختبار النهائي، ودرجات الاختبار المقنن على المستوى الوطني، واختبار التفكير المنطقي، واستبانة استراتيجيات محفزات التعلم، وتوصل نايزر إلى نتيجة مؤداها أن ٦٩% من الحالات التي فحصها تنبأت بها تلك المتغيرات بصورة صحيحة.

ركزت الدراسات السابقة على تحليل عينات من الحقائب التقويمية للطلبة، ولكنها أغفلت ضرورة تعرف الباحثين ممارسات المعلم في كيفية استخدام الحقيبة التقويمية للطالب في التقويم الصفي، وهذا ما أضافته الدراسة الحالية؛ ذلك أن تعرف تلك الممارسات يسهم في الكشف عن صدق الحقيبة

التقويمية في التقويم الصفي، كما أشار إلى ذلك ميسيك (Messick, 1994a, 1994b)، فإنه يرى أن تلك الممارسات تدفع المعلم للقيام بعملية المطابقة بين الأهداف التعليمية للمقرر ومحتوى المقرر من جهة، وأهداف الحقيبة التقويمية ومحتوياتها من جهةٍ أخرى خلال الفصل الدراسي، كما تسهم تلك الممارسات في تعرف المعلم على مدى تحقق أهداف التعليم والتعلم في الحقيبة التقويمية، ومدى قياسها لمختلف أبعاد السمة (التحصيل الدراسي المستهدف، مثلاً)، من خلال تنوع النشاطات في الحقيبة وارتباطها بالسمة المقاسة، وذلك من خلال التقويم الذاتي للمعلم لتعليمه الصفي في أثناء تطبيق الحقيبة التقويمية في التقويم، ومن خلال التغذية الراجعة التي يقدمها المعلم للطلبة، وتلك التي يحصل عليها المعلم من الطلبة كذلك.

مشكلة الدراسة

تُستخدم الحقيبة التقويمية للطالب في مدرستين ثانويتين - واحدة للبنين وأخرى للبنات - في المدارس الحكومية في البحرين منذ عامين، بوصفها جزءاً من عملية تقويم تحصيل الطالب في المقررات الدراسية في الصف الأول الثانوي فقط، ويشار هنا إلى أن التقويم الصفي - باستخدام الحقيبة التقويمية للطالب، بالمفهوم العلمي والمعايير العلمية لتطبيق الحقيبة التقويمية - مطبق في المدرستين اللتين تم تطبيق الدراسة فيهما، وذلك إلى حين جمع بيانات هذه الدراسة (وحسب علم الباحث ومعرفته).

أما المدارس الثانوية الأخرى، فلا تطبق التقويم باستخدام الحقيبة التقويمية للطالب بمفهومه العلمي، بل يقوم المعلم بتوزيع العلامة المخصصة لأعمال الطالب، وهي ٣٠% من العلامة الكلية للمقرر، على أعمال الطالب الفصلية في المقرر، مثل المشروعات، والبحوث، والتقارير، ولكن من دون مراجعة دورية لجودة تلك الأعمال، ومن دون تقديم تغذية راجعة مستمرة للطالب فيها، أو الطلب من الطالب عرض أعماله في الصف أمام زملائه، كما لا يُطلبُ من الطالبِ تقديم تغذية راجعة حول تلك الأعمال، وكذلك من دون منح الطالب علامة كلية شاملة على تلك الأعمال كما هو الحال في الحقيبة التقويمية للطالب في الصف الأول الثانوي.

والجدير بالذكر أن وزارة التربية والتعليم في طور التوسع التدريجي في تطبيق التقويم الصفي باستخدام الحقيبة التقويمية للطالب في المدارس الثانوية الأخرى، أما بالنسبة للمرحلتين الابتدائية والإعدادية، فلا يُطبقُ فيهما التقويم الصفي باستخدام الحقيبة التقويمية للطالب بمفهومها العلمي، بل توضعُ أعمال الطالب بعد الانتهاء منها في حافظة، أو ما يُعرف بملف الإنجاز (ملف الأعمال)، دون الرجوع إليه مرةً أخرى، أو استخدامه في التقويم الصفي خلال الفصل الدراسي بعد رصد درجة للطالب عليها.

وتكمن المشكلة في أن المعلم الذي يتولى تقويم الطالب بوساطة الحقيبة التقويمية لم يتم تدريبه تدريباً كافياً على عملية التقويم تلك، فقد أفاد ٦٣.٣% (١٩ معلماً ومعلمة) من أفراد العينة أنهم تلقوا تدريباً حول كيفية استخدام الحقيبة التقويمية للطالب في التقويم الصفي، في حين أن ذلك التدريب لم يكن كافياً، حسب ما أفاد ٨٦.٧% (٢٦ معلماً ومعلمة) منهم، فضلاً عن عدم وجود دراسة ميدانية في البحرين توضح مدى استخدام المعلم الحقيبة التقويمية في تقويم التحصيل، ناهيك عن أنه لم يتم التحقق من صدق الحقيبة التقويمية في تقويم تحصيل الطالب، وخاصة أن نتائج التقويم تلك تنعكس على تحصيل الطالب الدراسي في الصفين الثاني والثالث الثانوي، وأن العلامة المخصصة للتقويم بالحقيبة التقويمية تشكل ٣٠% من علامة المقرر في الصف الأول الثانوي، كل تلك العوامل مجتمعة قد تشكك في صدق نتائج تقويم المعلم للطالب في الحقيبة التقويمية.

أهداف الدراسة وأسئلتها

تهدف هذه الدراسة إلى:

١. الكشف عن ممارسات المعلم في استخدام الحقيبة التقويمية للطالب في تقويم تحصيل الطالب، ومدى استيفاء المعلم لكل العناصر الضرورية للتقويم بالحقيبة التقويمية، وذلك من خلال التقديرات الذاتية للمعلم.

٢. التحقق من صدق الحقيبة التقويمية للطالب في تقويم تحصيل الطالب، بوساطة نموذج ميسيك الموحد للصدق.

وفي ضوء هذين الهدفين، تسعى هذه الدراسة للإجابة عن السؤالين الآتيين:

١. ما ممارسات المعلم للحقيبة التقويمية للطالب في التعليم والتعلم الصفي، وفي تقويم تحصيل الطالب خلال المقرر الدراسي؟

٢. هل تتوافر مؤشرات أو دلائل صدق ملائمة للحقيبة حسب نموذج ميسيك الموحد للصدق؟

أهمية الدراسة

تكمن أهمية الدراسة في أنها تسلط الضوء على موضوع مهم لم تتم دراسته ميدانياً في المدارس الثانوية في البحرين، وهو مدى صدق الحقيبة التقويمية للطالب في تقويم تحصيل الطالب في المقرر الدراسي، ويُتوقع أن تسهم نتائجها في تحسين ممارسات المعلم في استخدام الحقيبة التقويمية في تقويم تحصيل الطالب، وجعلها أكثر صدقاً، أملاً في أن ينعكس ذلك على تحسين قرار نجاح الطالب في المقررات الدراسية، ومتابعة دراسته.

الطريقة والإجراءات

منهج الدراســـة

اتبعت الدراسة المنهج المسحي للتعُّرف على ممارسات المعلمين في استخدام الحقيبة التقويمية للطالب في تقويم تحصيل الطالب في المقرر الدراسي. كما قامت الدراسة أيضاً بتقويم مجموعة عشوائية من الحقائب التقويمية للطلبة، للتحقق من وجود العناصر الأساسية للحقيبة التقويمية للطالب، ثم دراسة العلاقات الارتباطية بين تلك العناصر، وبين مجموعة من المتغيرات المتعلقة بتحصيل الطالب في المقرر، مثل المعدل التراكمي، وتقديرات المتخصصين حول صدق الحقيبة التقويمية للطالب، وعلامات الطالب في الحقيبة التقويمية ككل، وعلامة الطالب في الاختبار الوطني الذي تطبقه وزارة التربية والتعليم.

مجتمع الدراسة وعينتها

تكون مجتمع الدراسة من جميع معلمي ومعلمات الصف الأول الثانوي في مملكة البحرين في المدرستين اللتين تطبقان التقويم باستخدام الحقيبة التقويمية للطالب (٦٠ معلماً، و ٧٠ معلمة). تم اختيار عينة عشوائية من المعلمين في كل مدرسة على حدة (١٥ معلماً من الصف الأول الثانوي لمدرسة البنين، و ١٥ معلمة من الصف الأول الثانوي لمدرسة البنات). كما تم اختيار ٦٠ حقيبة عشوائية من الحقائب التقويمية للطلبة، وذلك في خمسة مقررات في الصف الأول الثانوي وهي اللغة العربية (عرب ١٠١)، والرياضيات (ريض ١٠١)، والفيزياء (فيز ١٠١)، والكيمياء (كيم ١٠١)، والأحياء (حيا ١٠١). ويشار هنا إلى أن معايير وضوابط بناء وتطبيق الحقيبة التقويمية للطالب في تلك المقررات موحدة من قبل وزارة التربية والتعليم؛ كما أن ٣٠% من العلامة الكلية للمقرر الدراسي الواحد مخصصة رسمياً للتقويم بالحقيبة التقويمية للطالب، حسب لوائح وزارة التربية والتعليم.

أدوات الدراسة

استخدم الباحث ثلاث أدوات في الدراسة، وهي موضحة على النحو التالي:

١. **استبانة المعلم حول استخدام الحقيبة التقويمية للطالب في التقويم الصفي:** طبق الباحث استبانة مكونة من ٢٠ فقرة تتعلق بممارسات المعلم في استخدام الحقيبة التقويمية في تقويم تحصيل الطالب في الصف الأول الثانوي. وقد شملت الاستبانة خمسة محاور، هي تخطيط المعلم للحقيبة التقويمية (ثلاث فقرات)، ومحتوى الحقيبة التقويمية (سبع فقرات)، والنمو والتقدم في تحصيل الطالب (ثلاث فقرات)، والتواصل بين المعلم والطالب عبر الحقيبة التقويمية (خمس فقرات)، وتقويم تحصيل الطالب في الحقيبة التقويمية (فقرتان). وقد تم بناء فقرات الاستبانة في ضوء العناصر التي أشارت إليها الدراسات التربوية حول مكونات الحقيبة التقويمية للطالب واستخدامها في التقويم الصفي (Brookhart, 2011; Popham, 1999, 2008; Wiggins, 1998). وقد استخدم الباحث مقياس تقدير ثلاثي (٣= أمارس دائماً، ٢= أمارس أحياناً، ١= لا أمارس أبداً) (Messick, 1994a, 1994b)، لتعرف ممارسات أفراد العينة في الحقيبة التقويمية.

٢. **الحقائب التقويمية للطلبة:** أخذ الباحث عينة عشوائية من الحقائب التقويمية للطلبة والطالبات

(٦٠ حقيبة تقويمية) في المقررات الدراسية الخمسة، مضافاً إليها علامة الطالب الكلية في الحقيبة التقويمية، وعلامته في الاختبار الوطني، ومعدله التراكمي في المقرر.

وتتكون الحقيبة التقويمية للطالب في الصف الأول الثانوي من قائمة بالموضوعات التي تم تقويم الطالب فيها، ومعايير الأداء، وتوزيع الدرجات على الأعمال، والأنشطة والمشروعات والبحوث والتقارير، ونتائج الطالب في الاختبارات، ونماذج من أعمال الطالب على صورة أقراص مدمجة (CD). كما تتضمن الحقيبة التقويمية التغذية الراجعة من المعلم للطالب حول مستوى أدائه في الأعمال التي تشملها الحقيبة، وتقويم الطالب الذاتي لأعماله، وكذلك مقاييس التقدير لتحديد مستوى أداء الطالب في كل عمل من الأعمال التي ينجزها.

يتمثل دور المعلم في التقويم باستخدام الحقيبة التقويمية في التخطيط المُسبق لبناء الحقيبة، واختيار نوع وطبيعة وعمق الموضوعات التي ستشملها الحقيبة، وكيفية تقويم الطالب في كل موضوع، ونوع الأدوات الملائمة لتقويم كل عمل، وكيفية متابعة تقدم الطالب في التحصيل الدراسي في المقرر من خلال أعماله في الحقيبة التقويمية، والتواصل مع الطالب، وتقديم تغذية راجعة كتابية للطالب حول مستوى الأداء في أعماله، وكيفية استخدام مقياس التقدير المناسب لكل عمل من أعماله. كما يقوم المعلم أيضاً بتوجيه الطالب ومتابعته حول كيفية استخدام أعماله التي أنجزها سابقاً في تطوير أعماله الحالية، وتبصيره بكيفية تطوير جوانب القوة في أعماله، وتلافي جوانب القصور فيها.

أما دور الطالب في الحقيبة التقويمية، فيتمثل في إنجاز الأعمال المطلوبة منه، والمخطط لها سابقاً من قبل المعلم، والتحقق من استيفائها للمعايير والتعليمات الموضحة في الحقيبة، حسب كل عمل، والتقويم الذاتي الدوري لأعماله، وتحسينها، وإعادة تقديمها للمعلم لتقييمها مرةً أخرى، وعرضها في قاعة الدرس، والتواصل مع المعلم حول ذلك طوال الفصل الدراسي.

٣. **استمارة الحُكم على صدق التقويم باستخدام الحقيبة التقويمية للطالب**: استخدم الباحث استمارة مكونة من ٢٢ عنصراً (فقرةً) تغطي ٥ محاور، وهي: محتوى الحقيبة التقويمية (خمس فقرات)، والنمو المعرفي والتقدم في تحصيل الطالب (أربع فقرات)، والأعمال الإبداعية وحل المشكلات (أربع فقرات)، والتواصل (أربع فقرات)، والتفكير التأملي للطالب في أعماله (خمس فقرات)، فقرات الاستمارة مستقاة من الدراسات التي عالجت الموضوع نفسه؛ للتحقق من صدق الحقيبة التقويمية للطالب في تقويم تحصيله (Yao et. al, 2008)، وقد استخدمها الباحث للحصول على البيانات المتعلقة بصدق التقويم بالحقيبة التقويمية للطالب.

التعريفات الإجرائية للمصطلحات

فيما يلي تعريف بالمصطلحات التي وردت في هذا البحث والتي تشمل الوجوه الستة لنموذج ميسيك الموحد للصدق، كما عرّفها ميسيك نفسه (Messick, 1994a, 1994b, 1995a, 1995b, 1996).

صدق المحتوى (Content validity): مدى قدرة أداة القياس (الاختبار مثلاً) على تمثيل المحتوى والمهارات في مجال تعليمي محدد (مقرر دراسي مثلاً). ويجب أن تكون معايير المحتوى ملائمة، وتمثل نطاق السمة المقاسة.

الصدق المُوسَّعْ (Substantive validity): التحقق من عمليات النطاق (المجال) المعرفي التي يراد الكشف عنها في مهمات التقويم، من خلال قدرة مهمات الأداء تلك على تقديم المعاينة الملائمة لعمليات النطاق المعرفي، إضافة إلى تغطيتها لمحتوى النطاق المعرفي.

الصدق البنائي (Structural validity): يتعلق بعملية الاتساق المنطقي بين ما هو معروف حالياً عن العلاقات البنائية المتضمنة في جوانب السلوك التي تظهر بوصفها مؤشرات إلى السمة المراد قياسها. لذلك، يجب أن تتسق البنية الداخلية للتقويم وأدواته مع ما هو معروف حالياً عن البُنية الداخلية للنطاق المعرفي للسمة المراد قياسها.

صدق القابلية للتعميم (Generalizability validity): قدرة التقويم وأدواته على تمثيل محتوى وعمليات النطاق المعرفي للسمة المراد قياسها، مما يستوجب أن تكون تفسيرات نتائج التقويم (الدرجات

مثلاً) قابلة للتعميم بتوسع في حدود السمة المراد قياسها.

الصدق الخارجي (External validity): المدى الذي تعكس فيه علاقات نتائج التقويم (الدرجات مثلاً) مع المقاييس الأخرى ومع السلوك غير المتعلق بالتقويم، والعلاقات المتضمنة في السمة المحددة المراد قياسها، وذلك بالمستويات الثلاثة لتلك العلاقات (العالي، والمنخفض، والتفاعلي).

صدق النتائج/العواقب (Consequential validity): قدرة التقويم وأدواته على الحصول على دليل أو مجموعة أدلة لتقويم العواقب (النتائج) المقصودة وغير المقصودة عند تفسير نتائج التقويم (الدرجات مثلاً) واستخدامات تلك النتائج.

صدق وثبات أدوات الدراسة: للتحقق من صدق الاستبانة المكونة من ٢٠ فقرة، طلب الباحث من ١٠ محكمين الحكم على مدى ملاءمتها للغرض المطلوب تحقيقه، وهؤلاء خمسة محكمين متخصصين في القياس والتقويم التربوي، و خمسة محكمين متخصصين في علم النفس التربوي، وكلهم برتبة أستاذ مشارك. وقد كان أقل نسبة اتفاق بين المحكمين على فقرات الاستبانة ٨١,٢% ، وأعلى نسبة اتفاق بينهم كانت ٨٦,٧% وبلغ متوسط الاتفاق بين تقديرات المحكمين ٨٤,٢%. أما بالنسبة لثبات فقرات الاستبانة فقد حصل عليها الباحث من إجابات أفراد العينة من المعلمين والمعلمات (٣٠ فرداً). وقد بلغ معامل ثبات كرونباخ ألفا لتلك التقديرات ٠,٨٦. أما بالنسبة لصدق عناصر الحقيبة التقويمية (٢٢ عنصراً) فقد كان أقل نسبة اتفاق بين المحكمين ٨٢% وأعلى نسبة اتفاق ٨٦% تقريباً، وكان متوسط الاتفاق ٨٣,٧%.

إجراءات الدراسة

١. اجتمع الباحث بأفراد العينة، كل في مدرسته، وشرح لهم الهدف من الدراسة، ثم شرح لهم فقرات الاستبانة التي يريد تطبيقها عليهم، وجمع إجاباتهم بعد ٢٠ دقيقة من بدء تطبيق الاستبانة. كما أطلع الباحث أفراد العينة على استمارة المكونات الأساسية للحقيبة التقويمية (٢٢ عنصراً)، وطلب منهم وضع تقديرات على مدى التطابق بين تلك العناصر ومحتويات الحقيبة التقويمية للطالب من جهة، وبين فقرات الاستبانة (٢٠ فقرة) ومحتويات الحقيبة التقويمية للطالب من جهةٍ أخرى.

طلب الباحث من أفراد العينة كذلك الحصول على إجابة بنعم أو لا، من الطلبة الذين اختيرت حقائبهم التقويمية، مع التعليقات الكتابية للطلبة ومدى تقدمهم في التحصيل الدراسي في المقرر في الفصل الدراسي الثاني في ضوء أدائهم في الحقيبة التقويمية في الفصل الدراسي الأول، وذلك من خلال النموذج الذي ألحقه الباحث بأداة الدراسة التي تستهدف تعرف مدى استخدام المعلم للحقيبة التقويمية للطالب في التقويم الصفي. ثم جمع الباحث تلك المعلومات من أفراد العينة مع نهاية الفصل الدراسي الثاني.

كما طلب الباحث من أفراد العينة تسليمه نموذجين لحقيبتين تقويميتين لطالبين: أحدهما حاصل على أعلى درجة في الحقيبة التقويمية، والآخر حاصل على أدنى درجة فيها، إضافة إلى المعدل التراكمي للطالب، وعلامته في الحقيبة التقويمية (من ٣٠ درجة)، وعلامة الطالب في الاختبار الوطني (من ١٠٠). وبذلك يكون الباحث قد تسلم ٦٠ حقيبة تقويمية من أفراد العينة (٣٠ من المعلمين، و٣٠ من المعلمات). وتجدر الإشارة هنا إلى أن المعدل التراكمي للطالب في المدارس الثانوية في البحرين على شكل نسبة مئوية (يعتمد على العلامة ١٠٠ لكل مقرر)، وليس على شكل نقاط.

٢. طلب الباحث من المحكمين العشرة منح تقدير لكل عنصر من عناصر الحقيبة التقويمية الموضحة في الاستمارة، وذلك على مقياس تقدير رباعي (٤= ممتاز، ٣= جيد جداً، ٢= مقبول، ١= ضعيف) (Yao et al., 2008). طلب الباحث من المحكمين الحكم على مدى صدق الحقيبة التقويمية في تقويم تحصيل الطالب، وكان ذلك من خلال منح درجة (علامة) كلية على الحقيبة التقويمية للعناصر (٢٢ عنصراً) التي تشملها المحاور الخمسة لاستمارة الحكم على صدق التقويم بالحقيبة التقويمية

للطالب. استخدم الباحث مقياس ياو وزملائه (Yao et al., 2008) لمنح المحكم علامة (درجة) كلية على الحقيبة التقويمية، وهو كالتالي: ٤٥-٥٠ درجة (ممتاز)، ٤٠-٤٤ درجة (جيد)، ٣٥-٣٩ درجة (مرضي)، و ٠-٣٤ درجة (غير مرضي). ثم جمع الباحث تقديرات المحكمين على ٦٠ حقيبة تقويمية، بعد منح كل محكم ستة أيام للانتهاء من عملية الأحكام على صدق الحقائب التقويمية الستين في تقويم التحصيل الدراسي للطالب، على أن يسلم الحقائب كلها إلى المحكم التالي حتى انتهاء عملية التحكيم عند آخر محكم.

تحليل البيانات

١. استخدم الباحث النسبة المئوية لتحليل إجابات المعلمين والمعلمات عن فقرات الاستبانة، لتعرف ممارساتهم في الحقيبة التقويمية للطالب.

٢. استُخْدِمَ برنامج GENOVA (University of Iowa, 2012) لتحليل تقديرات المحكمين العشرة للحقائب التقويمية التي جمعها من أفراد العينة، مع قيم المعدل التراكمي للطلبة، ودرجاتهم على الحقائب التقويمية، وكذلك درجاتهم في الاختبار الوطني لحساب الارتباطات البينية ومعاملات الثبات (كرونباخ ألفا)، وذلك للتحقق من صدق الحقيبة التقويمية بوساطة نموذج ميسيك الموحد للصدق، والمكون من الأنواع الستة التي سبق شرحها، والجدير بالذكر أن برنامج التحليل الإحصائي GENOVA يبحث في تحليل الثبات والارتباطات البينية بين فقرات المقياس المطبق في البحث، كما يقوم البرنامج بتحليل المكونات الأساسية للتباين المستهدف، والتفاعلات بين تلك المكونات، والتباين الكلي، ونسبة التباين لكل مكون أساسي، والتفاعل بين المكونات الأساسية (مثل علامة الطالب في الحقيبة التقويمية وتقديرات المحكمين) وبين الخطأ. وهو شبيه بأسلوب التحليل العاملي إلى حدٍ ما، وينطلق البرنامج في فلسفته وبُنيته من نظرية القابلية للتعميم Generalizability theory التي طورها برينان (Brennan, 2001). وأدخل عليها شيو (Chiu, 2001) بعض الإضافات. ونظرية القابلية للتعميم عبارة عن إطار إحصائي Statistical framework لتوصيف المشاهدات الميدانية للظاهرة المقاسة، ثم تقصيها وتحري مدى ثبات مكوناتها. وتُستخدمُ للتحقق من ثبات القياس ضمن شروط أو ظروف محددة، وتتركز فائدتها بخاصة في تقويم الثبات عند دراسة تقويم الأداء. وتتعامل هذه النظرية مع مصادر التباين على أساس أنها وجوه Facets لسمة واحدة، وذلك في الدراسة العامة للسمة (G study)، كما هو الحال في نموذج ميسيك الموحد للصدق. فضلاً عن إمكان استخدام نتائجها فيما بعد في اتخاذ قرار جديد في حالة حدوث تعديل في جوانب الدراسة الميدانية التي قام بها الباحث، وهو ما يُعرفُ بدراسة القرار (D study).

النتائج

تشير البيانات في جدول ١ إلى ممارسات المعلمين والمعلمات في استخدام الحقيبة التقويمية لتقويم تحصيل الطالب، ومدى شمولها للمكونات الأساسية المطلوبة. ويتضح من جدول ١ أن أكثر الممارسات المستخدمة بشكلٍ دائم، أو أقرب إلى ذلك تركزت في الفقرات ٤، ٥، ٦، ١٨، ١٩، والمتعلقة بوضع المعلم الأنشطة والمشروعات والبحوث، ووضع تقدير عام (درجة عامة) ومقياس تقدير في الحقيبة. أما الفقرات الأخرى، فيستخدمها المعلمون "أحيانا" في أثناء التقويم بالحقيبة التقويمية، وشمل ذلك الفقرات ١، ٢، ٧، ١٣، ١٤، ١٥، ١٦، ١٧، ٢٠، والتي تتعلق بالتخطيط لبناء الحقيبة التقويمية قبل بدء العام الدراسي، ووضع قائمة بالموضوعات وشهادات الإنجاز والتقدير في الحقيبة، واللقاء بالطلبة لمناقشة مستوى أدائهم ومنحهم الفرصة للتقويم الذاتي، وعرض ما أنجزوه في الصف، واستخدام الحقيبة في التقويم الصفي اليومي وفي تصحيح مسار التعلم، ووضع مقياس تقدير في الحقيبة لتقويم مهمات الأداء.

أما الفقرات المتبقية فقد تأرجحت ممارسات المعلمين فيها بين "أحياناً" و "لا"، مع ملاحظة أن أكثر من ٤٠% من المعلمين أفادوا أنهم لا يدرجون في الحقيبة التقويمية رأي الطالب (الفقرة ٨)، ولا نماذج من أعمال الطالب (الفقرة ١٠)، ولا أي خطط للطالب مع زملائه لإنجاز المشروعات (الفقرة ١٢). ويُلاحظ بشكلٍ عام أن أكثر ممارسات المعلمين لاستخدام الحقيبة التقويمية في التقويم الصفي كان في الاهتمام بالأنشطة والمشروعات، ووضع تقدير عام (درجة عامة) للطالب في الحقيبة، وللتحقق من وجود مؤشرات أو أدلة على صدق الحقيبة التقويمية في تقويم تحصيل الطالب، من خلال نموذج ميسيك الموحد للصدق (السؤال الثاني)، فإن البيانات الموضحة في الجداول التالية (٢، ٣، ٤، ٥) تجيب عن السؤال الثاني المتعلق بكيفية التحقق من تلك الأنواع من الصدق، وذلك وفق المعايير التي وضعها ميسيك في نموذجه الموحد للصدق. يرى ميسيك أن **دليل صدق المحتوى** يعتمد أساساً على أحكام الخبراء (المحكمين) حول حدود السمة، والمحتوى، والمهارات، التي تقيسها مهمات الأداء في تقويم الأداء بشكلٍ عام، ومنها الحقيبة التقويمية للطالب (Messick, 1996).

يوضح جدول ٢ الارتباطات البينية بين أحكام (تقديرات) المحكمين العشرة، ومعاملات الثبات (كرونباخ ألفا) التي توضح الاتساق الداخلي بين تقديرات المحكمين في المحور الواحد.

جدول ١

النسبة المئوية لتقديرات المعلمين الذاتية لممارستهم الصفية في استخدام الحقيبة التقويمية للطالب في التقويم الصفي (ن = ٣٠)

أبعاد المقياس وفقراته الفقرة	دائماً (%)	أحياناً (%)	لا (%)
البُعد الأول: تخطيط المعلم للحقيبة التقويمية	٢٥,٥	٥١,١	٢٣,٤
١. أخطط لبناء الحقيبة التقويمية للطالب قبل بداية الفصل الدراسي.	٢٣,٣	٥٣,٤	٢٣,٣
٢. أضع قائمة بالموضوعات التي ستشملها الحقيبة التقويمية.	٣٣,٣	٥٦,٧	١٠
٣. أضع في الحقيبة التقويمية معايير الأداء لتحديد مستوى أداء الطالب في مهمات الأداء التي ينفذها.	٢٠	٤٣,٣	٣٦,٧
البُعد الثاني: محتوى الحقيبة التقويمية	٤٥,٧	٢٩	٢٥,٣
٤. تتضمن الحقيبة التقويمية للطالب الأنشطة التي ينفذها الطالب.	٧٣,٤	٢٣,٣	٣,٣
٥. تتضمن الحقيبة التقويمية للطالب المشروعات التي ينفذها الطالب.	٦٦,٧	٢٠	١٣,٣
٦. تتضمن الحقيبة التقويمية للطالب البحوث التي يقوم بها الطالب.	٥٠	٤٣,٣	٦,٧
٧. . تتضمن الحقيبة التقويمية للطالب شهادات الإنجاز والتقدير المدرسية وغير المدرسية التي حصل عليها الطالب.	٥٣,٤	١٣,٣	٣٣,٣
٨. تتضمن الحقيبة التقويمية للطالب صفحات خاصة برأي الطالب في مهمات الأداء التي نفذها (من حيث جودتها وكيفية تحسينها لاحقاً).	٢٦,٧	٢٦,٧	٤٦,٦
٩. تتضمن الحقيبة التقويمية للطالب تقارير وصفية يكتبها المعلم عن مستوى أداء الطالب في مهمات الأداء التي ينفذها الطالب.	٢٦,٧	٤٣,٣	٣٠
١٠. تتضمن الحقيبة التقويمية للطالب نماذج من أعمال الطالب في المقر على صورة أشرطة سمعية وبصرية وأقراص مدمجة لموضوعات/مشروعات المقرر أو زيارات ميدانية قام بها الطالب.	٢٣,٤	٣٣,٣	٤٣,٣
البُعد الثالث: النمو والتقدم في تحصيل الطالب	٢١.١	٤٤,٤	٣٤,٥
١١. تتضمن الحقيبة التقويمية للطالب خطط الطالب الفردية لكيفية تنفيذ مهمات الأداء المطلوبة.	١٠	٥٣,٣	٣٦,٧
١٢. تتضمن الحقيبة التقويمية للطالب خطط الطالب مع زملائه في المشروعات والأنشطة المشتركة لكيفية تنفيذ مهمات الأداء المطلوبة.	١٦,٧	٤٠	٤٣,٣
١٣. أمنح الطالب الفرصة الكافية للتقييم الذاتي في مهمات الأداء المطلوبة في الحقيبة التقويمية.	٣٦,٧	٤٠	٢٣,٣
البُعد الرابع: التواصل بين المعلم والطالب عبر الحقيبة التقويمية	٣٦,٦	٥٠,٧	١٢,٧
١٤. ألتقي الطالب لمناقشة مستوى أدائه في مهمات الأداء في الحقيبة التقويمية، وخاصة عندما يكون مستوى الطالب منخفضاً.	٤٠	٤٦,٧	١٣,٣
١٥. أطلب من الطالب عرض مهمات الأداء التي أنجزها لمناقشتها أمام زملائه في الصف.	٣٦,٧	٥٣,٣	١٠
١٦. استخدم الحقيبة التقويمية للطالب في التعليم الصفي اليومي.	٢٣,٣	٦٠	١٦,٧
١٧. أعود إلى مهمات الأداء التي أنجزها الطالب في الحقيبة التقويمية، لتصحيح مسار تعلمه في الدروس اليومية.	٤٣,٣	٣٣,٣	٢٣,٤
١٨. تفيدني الحقيبة التقويمية للطالب في تصحيح مسار تعليمي الصفي ومعرفة أخطائي في التدريس.	٤٠	٦٠	٠
البُعد الخامس: تقويم تحصيل الطالب في الحقيبة التقويمية	٥٠	٣١,٦	١٨,٤
١٩. أضع تقديراً عاماً/ درجة عامة للطالب في الحقيبة التقويمية.	٦٣,٣	٣٠	٦,٧
٢٠. يوجد في الحقيبة التقويمية للطالب مقياس تقدير يختلف حسب نوع مهمة الأداء (بحث/نشاط/مشروع) التي ينفذها الطالب.	٣٦,٧	٣٣,٣	٣٠

جدول ٢

الارتباطات البينية ومعاملات الثبات (كرونباخ ألفا) لتقديرات المحكمين على الحقائب التقويمية للطلبة (دليل صدق المحتوى)

المحور	الأول	الثاني	الثالث	الرابع	الخامس
الأول: محتويات الحقيبة التقويمية	(٠,٨٤)				
الثاني: النمو والتقدم في تحصيل الطالب	٠,٧٩	(٠,٩٢)			
الثالث: الأعمال الإبداعية وحل المشكلات	٠,٧٥	٠,٧٢	(٠,٩٠)		
الرابع: التواصل	٠,٧٨	٠,٧٧	٠,٨٥	(٠,٨٩)	
الخامس: التفكير التأملي للطالب في أعماله	٠,٨٠	٠,٧٣	٠,٨١	٠,٦٣	(٠,٨٧)

ملاحظة: المعاملات في المحور هي معاملات الثبات؛ القيم الأخرى هي معاملات الارتباط البينية بين المحاور الخمسة.

جدول ٣

معاملات الارتباط بين كل محور وعناصره من محاور مكونات الحقيبة التقويمية للطالب، ومعاملات الارتباط بين كل عنصر بالمحاور الأخرى وعناصرها (دليل الصدق البنائي)

المحور وعناصره (فقراته)	الأول	الثاني	الثالث	الرابع	الخامس
الأول: محتوى الحقيبة التقويمية					
١. المشروعات والبحوث مرتبطة بأهداف التعلم في المقرر	٠,٨٣	٠,٦٤	٠,٧١	٠,٤٧	٠,٦٦
٢. الأنشطة الصفية متنوعة وتغطي موضوعات المقرر	٠,٧٩	٠,٦٣	٠,٧٧	٠,٦٥	٠,٧٢
٣. هناك تنوع في الاختبارات الصفية (مقالية، موضوعية، .... الخ)	٠,٨١	٠,٥٨	٠,٦٨	٠,٥١	٠,٦٨
٤. إرشادات تنفيذ مهمات الأداء واضحة ومحددة للطالب	٠,٨٥	٠,٥٢	٠,٧٤	٠,٤٨	٠,٥٥
٥. مقاييس تقدير مستوى أداء الطالب واضحة ومرتبطة بمهمات الأداء	٠,٨٢	٠,٦١	٠,٦٩	٠,٥٦	٠,٥٨
الثاني: النمو المعرفي والتقدم في تحصيل الطالب					
٦. مهمات الأداء تُظهر تقدماً في تحصيل الطالب من حين إلى آخر	٠,٦٣	٠,٩٠	٠,٤١	٠,٦٠	٠,٦٦
٧. يتعقد محتوى مهمات الأداء طبقاً لتقدم الطالب في التحصيل	٠,٨١	٠,٨٥	٠,٥٧	٠,٦٠	٠,٧٢
٨. تقوم مهمات الأداء مرتبط بطبيعة المهمات وأنواعها	٠,٧٤	٠,٩١	٠,٤٨	٠,٦٩	٠,٧٠
٩. توجد أدلة على تقدم الطالب في التحصيل الدراسي عبر الزمن	٠,٦٩	٠,٨٧	٠,٤٦	٠,٥٧	+٠,٥٩
المحور الثالث: الأعمال الإبداعية وحل المشكلات					
١٠. يُظهر الطالب قدرة واضحة على حل المشكلات خلال مهمات الأداء	٠,٧١	٠,٦١	٠,٨٨	+٠,٥٨	٠,٧٤
١١. أعمال الطالب فيها إبداع (إضافات جديدة، اقتراح حل جديد)	٠,٧١	٠,٦٠	٠,٨٩	٠,٤٧	٠,٨٠
١٢. خصائص الطالب ومهاراته الشخصية واضحة تماماً في أعماله	٠,٦٣	٠,٦٩	٠,٨٤	٠,٤٤	٠,٧٧
١٣. لدى الطالب قدرة على تطبيق المعرفة في مواقف جديدة	٠,٦٢	+٠,٥٨	٠,٨٨	٠,٧٠	٠,٧٢
المحور الرابع: التواصل					
١٤. مقدمة كل أعمال الطالب واضحة وبسهل فهمها من عدة أطراف	٠,٧٨	٠,٤٩	٠,٤٣	٠,٨٦	٠,٤٥
١٥. تنظيم الأعمال في الحقيبة التقويمية مرتب منطقياً، ويسهل تتبعه	٠,٦٩	٠,٥٩	٠,٤٧	٠,٩١	٠,٥٩
١٦. أعمال الطالب تظهر قدرته على ربط الأعمال بعضها ببعض وبالواقع	٠,٧٠	٠,٧١	٠,٥١	٠,٧٨	٠,٧٣
١٧. اللغة المستخدمة في عرض الأعمال توصل أفكار الطالب إلى الآخرين	٠,٧٢	+٠,٦٣	٠,٦٢	٠,٨٩	٠,٦٧
المحور الخامس: التفكير التأملي للطالب في أعماله					
١٨. يحدد الطالب جوانب القوة في أعماله	٠,٥٧	٠,٧٢	+٠,٥٦	٠,٦٥	٠,٧٧
١٩. يحدد الطالب جوانب الضعف في أعماله	٠,٦٦	٠,٧٠	٠,٦٠	٠,٥٧	٠,٧٢
٢٠. يشرح الطالب كيفية تطوير جوانب القوة في أعماله	٠,٧٠	٠,٦٦	٠,٦٣	٠,٦٠	٠,٨٤
٢١. يشرح الطالب كيفية علاج جوانب الضعف في أعماله	٠,٥٩	٠,٦٠	٠,٥٩	٠,٦٠	٠,٨٧
٢٢. يوثق الطالب التقويم الذاتي لأعماله باستخدام دليل خارجي	٠,٦٩	٠,٧٠	٠,٤٩	+٠,٢٧	٠,٨١

ملحوظة: القيم المكتوبة باللون الأسود الغامق والمؤكدة بخط تحتها هي معاملات ارتباط كل محور بعناصره، وبقية القيم هي معاملات الارتباط (ر) بين كل عنصر والمحاور الأخرى. كل القيم دالة إحصائياً عند مستوى ٠,٠٥، باستثناء المشار إليها بعلامة + .

جدول ٤

المكونات الأساسية للتباين المقدر من علامة الطالب في الحقيبة التقويمية ودرجات المحكمين على ملاءمة عناصر الحقيبة وتقديرات المحكمين الكلية على الحقيبة التقويمية (دليل صدق القابلية للتعميم)

مصدر التباين	مجموع المربعات	د ح	متوسط المربعات	التباين المقدر	% من التباين الكلي
علامة الطالب في الحقيبة التقويمية (١)	٥٤٩٢,٦٢	٥٩	٩٣,٠٩	٥,١٣	٩,٠١
درجات المحكمين على ملاءمة عناصر الحقيبة التقويمية (٢)	٤٢٦٢٠,١٢	٩	٤٧٣٥,٥٧	٤٦,٧٦	٨٢,١٨
تقديرات المحكمين الكلية على الحقيبة التقويمية (٣)	٩٤٨,٩٠	٩	١٠٥,٤٣	٠,٨٤	١,٤٨
التفاعل بين (١) و (٢)	٨٧٢,٤٠	٦٣١	١,٣٨	٠,٦٢	١,٠٩
التفاعل بين (١) و (٣)	٢٥٨٠,٧٥	٦٣١	٤,٠٩	٢,١٨	٣,٨٣
التفاعل بين (٢) و (٣)	٧٤٦,٥٣	٨١	٩,٢٢	١,٠٧	١,٨٨
التفاعل بين (١. ٢ . ٣) و الخطأ	١٤٧٢,٨١	٤٧٧٩	٠,٣١	٠,٣٠	٠,٥٣

ويتضح من جدول ٢ أن معاملات ثبات الاتساق الداخلي بين أحكام المحكمين لعناصر الحقيبة التقويمية للطالب كانت عالية جداً، فقد تراوحت بين ٠,٩٢ (المحور الثاني)، و ٠,٨٤ (المحور الأول). أما الارتباطات البينية بين المحاور الخمسة للحقيبة التقويمية فقد كانت عالية (٠,٧٢ إلى ٠,٨٥)، باستثناء الارتباط بين المحور الخامس (التفكير التأملي للطالب في أعماله) والمحور الرابع (التواصل) فقد كان متوسطاً (٠,٦٣).

أما **دليل الصدق البنائي**، والمكون أساساً من الصدق التقاربي والصدق التمايزي، فيعتمد على فحص نظام تقدير الدرجات، والعلامات التي تُمنح للطالب في أدوات تقويم مهمات الأداء في الحقيبة التقويمية للطالب. ويرى ميسيك (Messick, 1996) أن طريقة تقدير الدرجات (العلامات) يجب أن تكون متسقة مع المجال المعرفي للسمة التي تقع في نطاقه مهمات الأداء في الحقيبة التقويمية، على أساس أن الصدق البنائي أحد جوانب صدق السمة.

ولما كان التحليل الأولي للسمة من خلال تقديرات المحكمين على جميع عناصر أداة التقويم (الحقيبة التقويمية) نظاماً فعالاً للتحقق من وجود دليل على الصدق البنائي - كما يذهب إلى ذلك ميسيك - فقد قام الباحث بإجراء عملية تحليل إحصائي للعناصر المكونة للحقيبة التقويمية في ضوء تقديرات المحكمين والعلامات التي منحوها لكل عنصر من تلك العناصر، وقد شمل ذلك التحليل مدى ارتباط العنصر مع محوره (الصدق التقاربي)، ومدى ارتباطه بالمحاور الأخرى (الصدق التمايزي)، كما تشير إلى ذلك النتائج في جدول ٣.

ويتضح من جدول ٣ أن الارتباطات (ر) بين عناصر المحور الواحد، في كل المحاور الخمسة، كانت عالية جداً، فقد كان معظمها أعلى من ر = ٠,٨٠ . ومعظمها دال إحصائياً عند مستوى الدلالة ألفا= ٠,٠٥، مما يشير إلى توافر دليل على وجود الصدق التقاربي. أما بالنسبة للعلاقات الارتباطية بين كل عنصر وعناصر المحاور الأخرى (الصدق التمايزي)، فقد كانت متوسطة بشكلٍ عام لكل المحاور الخمسة (بين ر= ٠,٦٠ و ر= ٠,٦٩)، مع استثناءات قليلة، فقد كانت معاملات الارتباط ضعيفة نسبياً (أقل من ٠,٥٠) . مع ملاحظة أن أعلى ارتباط من بين كل تلك الارتباطات كان بين العنصر ١٤ في محور التواصل، والمحور الثاني (النمو المعرفي والتقدم في تحصيل الطالب) (ر= ٠,٧٨)؛ وأدنى ارتباط (ر= ٠,٢٧) كان بين العنصر ٢٢ في المحور الخامس (التفكير التأملي والتقويم الذاتي للطالب في أعماله)، والمحور الرابع (التواصل). وهذه الارتباطات جميعها تشمل دليل الصدق التمايزي.

أما بالنسبة إلى دليل **صدق القابلية للتعميم** لاستخدام الحقيبة التقويمية في التقويم الصفي، كونه جزءاً من صدق السمة، فقد حصل عليه الباحث عن طريق استخدام التحليل الخاص بنظرية القابلية للتعميم، والذي يتطلب تفكيك التباين الكلي إلى مكوناته الأساسية، ومعرفة التفاعل بين تلك المكونات، وكذلك التفاعل بينها وبين خطأ التباين (Brennan, 2001; Chiu, 2001).

ويوضح جدول ٤ نتائج تفكيك التباين الكلي وتحليله إلى مكوناته الأساسية، والناتج عن معالجة البيانات المتعلقة بعلامات الطلبة في الحقيبة التقويمية، ودرجات المحكمين على مدى ملاءمة عناصر الحقيبة التقويمية، وتقديرات المحكمين العامة على جودة الحقيبة التقويمية؛ والتي تُشكل في مجموعها المكونات الثلاثة الأساسية للتباين الكلي. وتشير النتائج في جدول ٤ إلى أن أكبر مكون من مكونات التباين (٤٦,٧٦)، والمتعلق بدرجات المحكمين على مدى ملاءمة الحقيبة التقويمية، يشكل ٨٢,١٨% من التباين الكلي، بينما كان نصيب علامات الطالب في الحقيبة التقويمية ٩,٠١ من التباين الكلي. ويُلاحظ أن أعلى تباين نتج عن التفاعل بين المكونات الثلاثة الأساسية كان من نصيب التفاعل بين المكونين ١ (علامات الطالب في الحقيبة التقويمية)، و ٣ (تقديرات المحكمين)، فقد بلغ ٣,٨٣ ، وأقل تباين من تلك التفاعلات كان بين المكونين ١، و ٢ (انظر جدول ٤).

وأما **دليل الصدق الخارجي**، فإنه يعتمد على الارتباط بين علامات الطالب في الحقيبة التقويمية التي قدّرها له المعلم، مع متغيرات خارجية، كما في هذه الدراسة، وهي المعدل التراكمي للطالب، وعلامة الطالب في الاختبار الوطني السنوي الذي تعده وتجريه وزارة التربية والتعليم.

ويرى ميسيك (Messick, 1996) أن الأفراد الذين يحصلون على علامات عالية في الحقيبة التقويمية يُتوقع أن يحصلوا على درجات عالية أيضاً على مؤشرات خارجية أخرى للسمة المقاسة. ويُعبِّر عن كل ذلك بمعاملات الارتباط بين علامة الطالب في الحقيبة التقويمية وعلاماته في المتغيرات الخارجية، فإنها دليلٌ على الصدق الخارجي. ويشير جدول ٥ إلى علاقات ارتباطية موجبة ومتوسطة القوة بين علامة الطالب في الحقيبة التقويمية كما حددها المعلم، وعلامته في الاختبار الوطني، ومعدله التراكمي في الصف الأول الثانوي؛ وهي دليل على توافر الصدق الخارجي إلى حدٍ كبير.

والجدير بالذكر أن متوسط علامات الطلبة في الحقيبة التقويمية (م= ٢٩، ع = ٣,٢)، والاختبار الوطني (م= ٩٣,٢، ع = ٠,٣١) والمعدل التراكمي (م = ٨٩,٤، ع = ٠,١٥) كانت عالية جداً للطلبة الأعلى تحصيلاً. أما الطلبة الأدنى تحصيلاً، فقد أظهرت تلك النتائج انخفاضاً كبيراً في متوسط أدائهم في كل من الحقيبة التقويمية (م = ١٦، ع = ٠,١٢)، والاختبار الوطني (م = ٦٨,٢، ع = ٠,٤٩)، والمعدل التراكمي (م = ٦٥,٧، ع = ٠,١٢). ويُلاحظُ كذلك أن هناك تقارباً كبيراً في مستوى أداء الطلبة الأعلى تحصيلاً في كل من الحقيبة التقويمية والاختبار الوطني والمعدل التراكمي، كما هو واضح من قيم الانحراف المعياري. وكذلك الحال بالنسبة للطلبة الأدنى تحصيلاً في العناصر الثلاثة نفسها.

جدول ٥

معاملات الارتباط بين علامة الطالب في الحقيبة التقويمية والمعدل التراكمي للطالب وعلامة الطالب في الاختبار الوطني (دليل الصدق الخارجي)

العنصر	١	٢	٣
علامة الطالب في الحقيبة التقويمية (١)	١,٠٠	٠,٧٥*	٠,٦٨*
المعدل التراكمي للطالب (٢)		١,٠٠	٠,٦٤*
علامة الطالب في الاختبار الوطني (٣)			١,٠٠

*دالة عند مستوى دلالة ٠,٠٥

يشير ميسيك (Messick, 1996) إلى أنه يمكن التحقق من وجود دليل الصدق الموسع (الواقعي/الحقيقي) من خلال ملاحظة أن العمليات المستخدمة في إنجاز مهمة الأداء، في إطار التقويم بالحقيبة التقويمية، وممارسات المعلم التقويمية وأثره في توضيح تلك العمليات للطالب، ومراقبته له باستمرار، وان ذلك كله يجب أن يكون ممثلاً وملائماً لنظيراته التي تمثل السمة المراد قياسها. ويرى ميسيك أنه يمكن تحقيق ذلك عن طريق معرفة مدى قرب أو بعد ممارسات المعلم في التقويم بالحقيبة التقويمية من عناصر ومكونات الحقيبة التقويمية التي حددتها الأبحاث في هذا الموضوع.

وكذلك يمكن التحقق من وجود دليل على الصدق الموسع من خلال المطابقة بين مكونات الحقيبة التقويمية وعناصرها، وبين أعمال الطالب كما هي في الواقع في حقيبته التقويمية. ويُلاحظ أن الكثير من الفقرات التي تمثل مقدار ممارسات المعلمين في التقويم باستخدام الحقيبة التقويمية (جدول ١) تطابق تقريباً العناصر الأساسية التي حددتها الأبحاث السابقة (جدول ٣). وقد اتضح ذلك من خلال قيام المحكمين العشرة بعملية المطابقة بين الفقرات في جدول ١، والعناصر في جدول ٣. فقد وجد الباحث أن نسبة الاتفاق (المطابقة) كانت ٨٦,٤%، وهو الدليل الأول على الصدق الموسع.

أما الدليل الثاني على الصدق الموسع، فقد تم الحصول عليه من خلال الطلب من جميع أفراد العينة (٣٠ معلماً ومعلمة) المطابقة بين مكونات الحقيبة التقويمية وعناصرها (جدول ٣)، وبين أعمال الطلبة في حقائبهم التقويمية، وتم تكرار هذه العملية مع المحكمين العشرة أيضاً. وقد أشارت نتائج المطابقة إلى أن ٧٨,٤% من المعلمين اتفقوا على تطابق عناصر الحقيبة التقويمية مع أعمال الطالب، بينما كانت نسبة الاتفاق بين المحكمين ٨٩% تقريباً. وهذا هو الدليل الثاني على الصدق الموسع.

آخر نوع من أنواع الصدق الستة في نموذج ميسيك الموحد للصدق هو **صدق العواقب (النتائج)**، دليل صدق العواقب يتحقق من مدى وجود العواقب المقصودة وغير المقصودة التي تترتب على استخدامات نتائج تقويم تحصيل الطالب في الحقيبة التقويمية، وتأثير ذلك في تفسير الدرجات واستخداماتها في صنع قرارات تتعلق بالطالب في المستقبل المنظور، وخاصةً عندما يُنقل إلى صف أعلى (Messick, 1996).

ويرى ياو وزملاؤه (Yao et al., 2008) أنه من المهم التحقق من العواقب المحتملة لاستخدامات نتائج

التقويم بالحقيبة التقويمية لأغراض متعددة، مثل تقديم تعليم علاجي للطالب، أو تعليم إثرائي يعزز تعلمه، أو عد الحقيبة التقويمية معياراً هاماً أو مكوناً رئيساً من مكونات تقويم تحصيل الطالب؛ ومتابعة تأثير ذلك بشكلٍ مباشر أو غير مباشر في تعلمه الحالي والقادم في المقرر الدراسي. وللتحقق من ذلك، طلب الباحث من أفراد عينة الدراسة من المعلمين والمعلمات أن يطلبوا من الطلبة الذين اختيرت حقائبهم التقويمية للدراسة أن يجيبوا عن سؤالين، أحدهما مغلق الإجابة (نعم أو لا)، يتعلق بتقدم الطالب في التحصيل الدراسي في المقرر جراء أدائه في الحقيبة التقويمية؛ والآخر مفتوح النهاية يتعلق بتأثير تعلم الطلبة وتقويمهم بالحقيبة التقويمية في تحصيلهم الدراسي في الفصل التالي. وقد أشارت النتائج، بشكلٍ عام، إلى أن معظم الطلبة (٨١.٧%) قد تقدم تحصيلهم في المقرر جراء أدائهم في أعمالهم في الحقيبة التقويمية الخاصة بمقرراتهم.

وقد بينت نتائج تعليقات الطلبة الكتابية حول السؤال المفتوح النهاية، أن ٧٥% من الطلبة ذوي التحصيل العالي في الحقيبة التقويمية أشاروا إلى تعلمهم مهارات جديدة، وقدم لهم معلموهم نشاطات وأعمال تتحدى مستواهم الحالي، بينما أشار الباقون (٢٥%) إلى أن تحصيلهم ارتفع لكنهم لم يحصلوا من معلميهم على نشاطات أو أعمال تتحدى مستواهم. أما الطلبة ذوي العلامات المنخفضة في الحقيبة التقويمية، فلم يتغير مستواهم كثيراً، حيث أشار ٩٠% منهم إلى حاجتهم إلى تعليم علاجي، وأشار الباقون (١٠%) إلى وجود تقدم طفيف في مستواهم الدراسي في المقرر. وتبيّن العبارات التالية خلاصة لبعض النماذج الإيجابية والسلبية من التعليقات الكتابية للطلبة حول تقدمهم الدراسي في المقرر في الفصل الدراسي الثاني؛ في ضوء نتائج أدائهم في أعمالهم في الحقيبة التقويمية.

اللغة العربية (عرب ١٠١): "أستطيع الآن ربط بعض المهارات التي تعلمتها سابقاً في النحو، مع المهارات الأعلى منها، بسبب تركيز المعلم معي على الرجوع إلى عملي السابق في حقيبتي التقويمية، وتصحيح عملي في النحو بشكل مستمر في الصف". "أشعر كثيراً أنني لا أستفيد من أعمالي في حقيبتي في اللغة العربية، لأني لا أرى زيادة في علاماتي في المقرر بعد ذلك، وعندي الأخطاء تتكرر كثيراً في الإملاء والتعبير".

الرياضيات (ريض ١٠١): " علاماتي في المقرر ارتفعت كثيراً لأنني استفدتُ من تكرار المعلم لأخطائي معي عند مراجعته معي لها في حقيبتي، وكذلك عند مناقشتي معه، وعند طلبه منا البحث عن أخطائنا بأنفسنا في التمارين في الصف". " مستواي في الرياضيات لم يتغير في الفصل الثاني؛ لأنني حتى الآن لا أعرف كيف أحل المسألة الرياضية إذا كانت مكتوبة لي. لكن المعلم علمني ذلك بعض المرات في الحقيبة، ولكني أكرر الأخطاء نفسها".

الكيمياء (كيم ١٠١): " أحس أنني الآن أقدر على التمييز بين المعادلات الكيميائية والرموز الكيميائية لكثير من العناصر الكيميائية التي علمنا المدرس في الصف. كان المدرس يجلس معي ويطلب مني أن أقارن بين الصيغ الكيميائية للمعادلات والرموز، ثم يرجع معي إلى حقيبتي ويريني الصحيح من الخطأ، ثم يطلب مني عرض ذلك أمام زملائي. بعد ذلك يطلب مني المدرس أن أكتب تقريراً عن موضوع يشمل تلك العناصر والصيغ، ويصححه لي، ثم يعود معي مرة أخرى إلى حقيبتي ويقارن عملي الآن مع عملي السابق". " أحس أنني قادر على إجراء التجارب الكيميائية في الفصل الثاني، بعد أن علمني المعلم كيف أستفيد من عرض الأدوات وطريقة التجربة التي توجد في حقيبتي السابقة، وعرض أمامنا فلماً عن إجراء التجارب في المختبر، لكن المشكلة أني حتى الآن أخطأ كثيراً في المعادلات الكيميائية ولا أستطيع أحياناً التفريق بين بعض الرموز".

الفيزياء (فيز ١٠١): "تحسنتُ كثيراً في كتابة التقارير العلمية في المقرر؛ لأن المدرس كان دائماً يقارن تقاريري الحالية مع تقاريري التي عملتها قبل ذلك، ويريني أخطائي، وأخطاء زملائي في تلك التقارير في حقيبة كل منا". " لا أحس أني تقدمت في دراستي في مقرر الفيزياء في الفصل الثاني، وخاصةً أنني لم اتبع ما قاله المدرس لي وما كتبه من ملاحظات حول طريقتي في عمل تجارب المقرر في حقيبتي. ما تحسن

عندي فقط هو كتابة الملاحظات فقط بعد نهاية التجربة في المختبر".

الأحياء (حيا ١٠١): " حصلت على علامات عالية في الأحياء في الفصل الثاني؛ لأني عرفتُ كيف أحلل الموضوع إلى أجزاء وأفهما، وأبتعد عن الحفظ. حيث علمني مدرس الأحياء طريقة فهم الموضوع، لأنه يجلس معنا في مجموعات صغيره، ويجعلنا نحدد أجزاء الموضوع التي حفظناها في حقيبة كل منا، ويرينا علاماتنا المنخفضة بسبب حفظ الموضوع، وعلاماتنا المرتفعة في نقاط الموضوع بسبب الفهم". " مادة الأحياء صعبة جداً لي، ولم أحصل على علامات عالية فيها في الفصل الثاني، رغم أن المدرس راجع معي أخطائي في الاختبارات القصيرة، والبحوث، والمشروع، عند نهاية كل موضوع".

مناقشة النتائج

هدفت هذه الدراسة إلى تعرف ممارسات المعلمين في الصف الأول الثانوي في استخدام الحقيبة التقويمية للطالب في التقويم الصفي، وكذلك التحقق من مدى صدق التقويم بالحقيبة التقويمية، وذلك من خلال التوصل إلى دليل يؤيد كل نوع من الأنواع الستة لنموذج ميسيك الموحد للصدق. وقد أشارت نتائج الدراسة إلى أن المعلمين بشكل عام يستخدمون الحقيبة التقويمية للطالب في التقويم الصفي "أحياناً" أو "دائماً" طوال الفصل الدراسي.

كما أظهر أفراد العينة اتفاقاً كبيراً بين عناصر الحقيبة التقويمية (٢٢ عنصراً)، وبين محتويات الحقائب التقويمية التي جمعها الباحث منهم. وقد أيدت أحكام المحكمين، حول الموضوع نفسه، صدق أحكام المعلمين، مما يُعدُ مؤشراً على الصدق الموسع من جانب، وصدق المحتوى من جانبٍ آخر (Yao et al., 2008).

وأظهرت نتائج هذه الدراسة دليلاً على صدق المحتوى (جدول ٢) من خلال القيم المرتفعة لمعاملات ثبات الاتساق الداخلي، ومعاملات الارتباط بين المحاور الخمسة التي تتضمن ٢٢ عنصراً بوصفها مكونات أساسية للحقيبة التقويمية للطالب.

وهذه النتائج أعلى بكثير من نظيراتها التي توصل إليها ياو وزملاؤه (Yao et al., 2008)، والتي تراوحت بين ٠,٦٢ و ٠,٨١ (لمعاملات الثبات)، وبين ٠,٣٢ و ٠,٧٢ (للارتباطات البينية)؛ وأعلى أيضاً من النتائج التي توصل إليها بيرنز وهايت (Burns & Haight, 2005). فقد كان أعلى معامل ارتباط حصل عليه الباحثان هو ٠,٦٠ وأعلى معامل ثبات هو ٠,٥٣، وعليه فإن الحقيبة التقويمية في هذه الدراسة تتمتع بدليل قوي على صدق المحتوى (جدول ٢).

الدراسات التي أجراها جادبوري-أميوت وزملاؤه (Gadbury-Amyot et al., 2003)، وديفيز وبونامبيروما (Davis & Ponnamperuma, 2005)، وجيريرو (Guerrero, 2000)، و درهام ودايبيرنا (Derham & Diperna, 2007)، وبيرنبوم (Birenbaum, 2007) أشارت إلى أنه يمكن الحصول على دليل قوي للصدق البنائي (بشقيه التقاربي والتمايزي) في الحقيبة التقويمية للطالب بشكلٍ خاص، وتقويم الأداء بشكلٍ عام، إذا تم استخدام العناصر الرئيسة المكونة للحقيبة التقويمية التي أشارت إليها الدراسات. فقد وجدت تلك الدراسات أن قيم الصدق التقاربي بين محاور الحقيبة التقويمية تراوحت بين ٠,٨٠ و ٠,٨٩ . مع اختلاف طفيف من دراسة إلى أخرى.

أما بالنسبة لقيم الصدق التمايزي لعناصر الحقيبة التقويمية فقد تراوحت بين ٠,٧٣ و ٠,٣١ . هذه النتيجة قريبة من تلك التي توصلت إليها هذه الدراسة (جدول ٣). مع ملاحظة أن بعض تلك الدراسات مثل (Gadbury-Amyot et al., 2003). و (Birenbaum, 2007) وجدت أن كل تلك الارتباطات المتعلقة بالصدق التمايزي والتقاربي دالة إحصائياً (ألفا= ٠,٠٥). على العكس من هذه الدراسة التي بينت وجود ارتباطات غير دالة إحصائياً في بعض العناصر المتعلقة بالصدق التمايزي فقط.

والنتيجة التي يمكن استخلاصها من كل ذلك أن هذه الدراسة بينت وجود دليل قوي على الصدق البنائي. ولكن من جانبٍ آخر، يرى ديفيز وبونامبيروما (Davis & Ponnamperuma, 2005) أنه يجب عدم الوثوق بشكلٍ مطلق بالنتائج التي يتم التوصل إليها حول الصدق البنائي، فإن عدم وجود أدوات تقويم أخرى، غير الحقيبة التقويمية، لقياس أداء الطالب يسبب إشكالاً في قياس هذا النوع من الصدق. وذلك أنه في هذه الحالة قد يلجأ بعض الطلبة إلى النسخ أو الغش من زملائهم، أو نقل بعض الأعمال من مصادر أخرى مثل شبكة

الإنترنت، مما يصعّب من عملية التحقق من صدق التقويم.

أما بالنسبة إلى صدق القابلية للتعميم، فقد توصلت هذه الدراسة (جدول ٤) إلى أن الدرجات التي منحها المحكمون على مدى ملاءمة عناصر الحقيبة التقويمية استحوذت على أكبر تباين (٨٢,١٨%) كمصدر من المصادر الثلاثة الرئيسة التي أدت إلى التباين في درجات المحكمين. هذه النتيجة تختلف عن نظيراتها التي توصلت إليها بعض الدراسات (Gadbury Amyot et al., 2003; Guerrero, 2000)، والتي وجدت أن تقديرات المحكمين العامة لعناصر الحقيبة التقويمية أخذت النصيب الأكبر من التباين الكلي (حوالي٧٨%).

وقد بينت نتائج هذه الدراسة أن هناك دليلاً قوياً إلى حدٍ كبير على الصدق الخارجي (جدول ٥) يتمثل في وجود ارتباط قوي نسبياً بين علامة الطالب في الحقيبة التقويمية و المعدل التراكمي للطالب (ر= ٠,٧٥) من جهة، والحقيبة التقويمية وعلامة الطالب في الاختبار الوطني (ر= ٠,٦٨)، من جهةٍ أخرى. وكذلك بين المعدل التراكمي للطالب وعلامة الطالب في الاختبار الوطني (ر= ٠,٦٤). هذه النتيجة قريبة جداً من النتيجة التي توصل إليها جادبوري-أميوت وزملاؤه (Gadbury-Amyot et al., 2003). فهم حصلوا على ارتباط قيمته ٠,٧٠ بين علامة الطالب في الحقيبة التقويمية ومعدله التراكمي؛ وارتباط قيمته ٠,٦٠ بين علامة الطالب في الحقيبة التقويمية والاختبار الوطني؛ وكلا الارتباطين دالان إحصائياً عند مستوى الدلالة ٠,٠٥ . والنتائج التي توصلت إليها هذه الدراسة تُعد دليلاً أمبيريقياً على الصدق الخارجي، في ضوء مقارنتها بنتائج دراسات أخرى مماثلة، مثل دراسة جادبوري-أميوت وزملاؤه (Gadbury-Amyot et al., 2003)، ودراسة الميماج وأهميتاج (Alimemaj & Ahmetaj, 2010)، ودراسة ياو وزملاؤه (Yao et al., 2008).

وقد وجدت هذه الدراسة دليلاً قوياً على الصدق الموسع اتضح في نسبة الاتفاق المرتفعة (أعلى من ٨٠%) التي أظهرها المعلمون والمحكمون حول مدى ملاءمة عناصر الحقيبة التقويمة (٢٢ عنصراً) مع محتويات الحقائب التقويمية للطلبة (٦٠ حقيبة) مما اطلعوا عليها. هذه النتائج مقاربة لنظيراتها التي توصل إليها جادبوري-أميوت وزملاؤه (Gadbury-Amyot et al., 2003) (٠,٨٥)، وكذلك بورتر (Porter, 2007) (٠,٧٩). ولكن من جهةٍ أخرى، تشير دارلنج-هاموند (Darling-Hammond, 2000; 2006)، وبورتر (Porter, 2007) إلى أنه رغم إمكانية التحقق من الصدق الموسع من خلال الاتفاق بين المحكمين أو المعلمين على مدى ملاءمة عناصر الحقيبة التقويمية مع محتوياتها، فإن ذلك يبقى أمراً صعباً على الصعيد الواقعي أحياناً، وذلك بسبب إمكانية وضع الطالب لمجموعة أعمال في الحقيبة تكون بعيدة أحياناً عن المهمات المطلوبة، أو بسبب وجود أعمال جماعية ينفذها الطالب مع بعض زملائه في الصف، مما يؤدي إلى صعوبة تقدير الدرجات لكل فرد أسهم في مهمة الأداء.

ويبقى دليل صدق العواقب (النتائج) هو الأصعب في عملية التحقق من وجوده في أعمال الطالب في الحقيبة التقويمية. فإنه رغم إشارة هذه الدراسة إلى أن ٧٥% من مجموعة الطلبة المتفوقين حصلوا من معلمهم على تعليم إثرائي يعزز تعلمهم، فإن ٩٠% من الطلبة ذوي التحصيل المنخفض في الحقيبة التقويمية أشاروا إلى حاجتهم لتعليم علاجي في الفصل الثاني، مما يعني أنهم لم يتقدموا في التحصيل الدراسي، وهذه إحدى العواقب السلبية التي تُعدُ دليلاً على صدق العواقب (Messick, 1996). ولكن يجب الأخذ في الحسبان هنا أن انتقال الطالب إلى صف أعلى من صفه الحالي قد يُصعّبُ عملية التحقق من وجود دليل على صدق العواقب، وخاصةً إذا تم تسليم حقيبة الطالب التقويمية الحالية إلى معلم جديد في الصف الذي انتقل إليه الطالب، كي يتعرف المعلم النتائج المترتبة على انتقال الطالب إلى الصف الجديد، من حيث تقدم الطالب في التحصيل الدراسي أو تأخره، أو بقاؤه في مستواه السابق. هذه قضية هامة تجعل الحصول على دليل دامغ لصدق العواقب أمراً شاقاً (Ma & Rada, 2005). وقد تتعقد المشكلة أكثر إذا تم استخدام نتائج تقويم الطالب في الحقيبة التقويمية في صنع قرارات تتعلق بمصير الطالب التعليمي، مثل تسكينه في مجموعة المتفوقين أو المتأخرين دراسياً (Abrami & Barret, 2005).

الخلاصة

خلصت نتائج هذه الدراسة إلى وجود أدلة على صدق التقويم بالحقيبة التقويمية في المرحلة الثانوية، وذلك وفق نموذج ميسيك الموحد للصدق، مع أخذ المحاذير - التي أشار إليها الباحث في الحسبان - عند إصدار حُكم عام على صدق الحقيبة التقويمية. إن جانب القصور في هذه الدراسة هو عدم استطلاع رأي الطلبة فيما يتعلق بممارساتهم في الحقيبة التقويمية خلال تنفيذهم مهمات الأداء المطلوبة في المقررات الدراسية.

مقترحات لدراسات مستقبلية حول الموضوع

١. إجراء دراسات موسعة حول الموضوع تشمل قطاعاً أكبر من المعلمين، وعينات أكثر من الحقائب التقويمية للطلبة، واستخدام أدوات أخرى مثل استطلاع آراء الطلبة حول تحصيلهم الدراسي من خلال الحقيبة التقويمية.

٢. التحقق من ممارسات المعلمين في التقويم الصفي باستخدام الحقيبة التقويمية للطالب من خلال اللقاءات والمقابلات والملاحظة المباشرة لممارسات المعلمين، والعروض التي يقدمها المعلمون حول طريقة تقويمهم للطلبة باستخدام الحقيبة التقويمية.

٣. البحث عن العلاقة بين مكونات الحقيبة التقويمية وممارسات الطلبة فيها، وتأثير ذلك في تحصيلهم الدراسي من خلال الدراسات التتبعية (الطولية) Longitudinal من صفٍ إلى آخر في المرحلة الثانوية، وعبر المراحل الدراسية الثلاث (الابتدائية والإعدادية والثانوية).

٤. إجراء دراسات تقويمية مقارنة للكشف عن أثر كل من الحقيبة التقويمية التقليدية، والحقيبة التقويمية الإلكترونية e-portfolio في تحصيل الطالب.

المراجع

Abrami, P., & Barret, H. (2005). Directions for research and development on electronic portfolio. *Canadian Journal of Learning and Teaching*, *31* (3), 1-15.

Alimemaj, Z., & Ahmetaj, L. (2010, July-August). Portfolio assessment, a valid tool to evaluate students' achievements.*AngloHigher*, *2*(4), 9-11.

Birenbaum, M. (2007). Evaluating the assessment: Sources of evidence for quality assurance. *Studies in Educational Evaluation, 33*, 29-49.

Bredo, E. (2005). Addressing the social foundations "accountability dilemma". *Educational Studies, 38*(3), 230-241.

Brennan, R. (2001). *Genralizability theory*. New York: Springer-Verlag.

Brookhart, S. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice, 30*(1), 3-12.

Burns, M., & Haight, S. (2005). Psychometric properties and instructional utility of assessing special education teacher candidate knowledge with portfolios. *Teacher Education and Special Education, 23*(3), 185-194.

Chiu, C. (2001). *Scoring performance assessments based on judgments: Generalizability theory*. New York: Kluwer.

Darling-Hammond, L. (2000). Authentic assessment of teaching in context. *Teaching and Teacher Education, 16*(5), 523-545.

Darling-Hammond, L. (2006). *Powerful teacher education: Lessons from exemplary programs*. San Francisco: Jossey-Bass.

Davis, M., & Ponnamperuma, G. (2005). Portfolio assessment. *JVME, 32*(3), 279-284.

Derham, C., & Diperna, J. (2007). Digital professional portfolios of preservice teaching: An initial study of score reliability and validity. *Journal of Technology and Teacher Education, 15*(3), 363-381.

Donelly, A. (2005). Let me show you my portfolio: Demonstrating competency through peer interview. *Action in Teacher Education, 27*(3), 55-63.

Gadbury-Amyot, C., Kim, J., Palm, R., Mills, G., Noble, E., & Overman, P. (2003). Validity and reliability of portfolio assessment of competency in a baccalaureate dental hygiene program. *Journal of Dental Education, 67*(9), 991-1002.

Goodson, F. (2007). The electronic portfolio: Shaping an emerging genre. *Journal of Adolescent & Adult Literacy, 50*(6), 432-434.

Guerrero, M. (2000). The unified validity of the four skills exam: Applying Messick's framework. *Language Testing, 17*(4), 397-421.

Levine, A. (2006). *Educating school teachers*. Washington, DC: The Education schools Project.

Ma, X., & Rada, R. (2005). Building a web-based accountability system in a teacher education program. *Interactive Learning Environments, 13*(1), 93-119.

Messick, S. (1994a). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher, 23*(2), 13-23.

Messick, S. (1994b). *Alternative modes of assessment, uniform standards of validity*. Paper presented at a Conference on Evaluating Alternatives to Traditional Testing for Selection. Bowling Green, Ohio.

Messick, S. (1995a). Standards of validity and validity of standards in performance assessment. *Educational Measurement: Issues and Practice, 14*(4), 5-8.

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

Messick, S. (1996). *Standards-based score interpretation: Establishing valid grounds for valid inferences.* Proceedings of the joint conference on standard setting for large scale assessments. Washington, DC: Government Printing Office.

Naizer, G. (1997). Validity and reliability issues of performance-portfolio assessment. *Action in Teacher Education, 18*(4), 1-9.

NBPTS (2006). *Portfolio instructions: Middle childhood generalist*. Retrieved June 3, 2012, from: http://www.nbpts.org/for_candidates/the_portfolio?ID27&x=38&y=10

Popham, W. (1999). *Classroom assessment: What teacher's need to know*. Boston: Allyn and Bacon.

Popham, W. (2008). *Transformative assessment*. Alexandria: ASCD.

Porter, A. (2002). Measuring the content of instruction: Uses in research and practice. *Educational Researcher, 31*(7), 3-14.

Porter, A. (2007). Alignment as a teacher variable. *Applied Measurement in Education, 20*(1), 27-51.

Reis, N., & Villaume, S. (2002). The benefits, tensions, and visions of portfolios as a wide-scale assessment for teacher education. *Action in Teacher Education, 23*(4), 10-17.

Sherry, A., & Bartlett, A. (2005). Worth of electronic portfolios to education majors: A "two by four" perspective. *Journal of Educational Technology Systems, 33*(4), 399-419.

Strudler, N., & Wetzel, K. (2005). The diffusion of electronic portfolios in teacher education: Issues of initiation and implementation. *Journal of Research on Technology in Education, 37*(4), 411-433.

University of Iowa (2012). *GENOVA*. Retrieved July 13, 2012, from: http://www.uiowa.edu/casma/computer_programs.htm

Wiggins, G. (1998). *Educative assessment*. San Francisco: Jossey-Bass.

Yao, Y., Thomas, M., Nickens, N., Downing, J., Burkett, R., & Lamson, S. (2008). Validity evidence of an electronic portfolio for preservice teachers.*Educational Measurement: Issues and Practice, 27*(1), 10-24.

Zeichner, K., & Wray, S. (2001). The teaching portfolio in US teacher education programs: What we know and what we need to know. *Teaching and Teacher Education, 17*, 613-621.