ثبات تقييمات الكفاءة الاجتماعية لدى الأطفال باستخدام النسخة العربية لمقياس الكفاءة الاجتماعية المُعَّدل في ضوء نظرية إمكانية التعميم

إعـــداد

د/ صبري محمود عبد الفتاح أمين

قسم علم النفس التربوي

كلية التربية- جامعة المنيا- مصر

ثبات تقييمات الكفاءة الاجتماعية لدى الأطفال باستخدام النسخة العربية لمقياس الكفاءة الاجتماعية المُعَّدل في ضوء نظرية إمكانية التعميم

د/ صبري محمود عبد الفتاح أمين *

المقدمة:

يمثل الانتقال من بيئة البيت إلى بيئة المدرسة نقطة تحول مهمة في حياة كل طفل، ليس فقط على المستوى النفسي حيث يبتعد الطفل عن والديه فيقل اعتماده عليهما ويصبح أكثر اعتمادا على ذاته، لكن أيضا على المستوى الاجتماعي حيث يواجه الطفل كثيرا من التحديات والمتطلبات الاجتماعية التي يفرضها الواقع المدرسي الجديد. فالأطفال الملتحقون بالمدارس مطالبون بأن يتكيفوا اجتماعيا مع البيئة المدرسية من خلال تكوين شبكات علاقات اجتماعية ناجحة وفعالة تسهم في تشكيل هويتهم الاجتماعية وتساعدهم في تحقيق التكيف الاجتماعي المنشود وذلك في إطار مجموعة من القواعد والنظم واللوائح التي تحدد وتنظم العلاقات بينهم وبين أقرانهم، وبينهم وبين معلميهم. وينبغي أن يتعاون الأطفال الجدد مع أقرانهم في إنجاز ما يُسند إليهم من تكليفات وأعمال صفية ولا صفية، وهم في نفس الوقت مشغولون بحالة التنافس الأكاديمي فيما بينهم وهو تنافس لا يخلو من صبغة اجتماعية بهدف الحصول على تقبل ودعم الآخرين وفي مقدمتهم الوالدين والمعلمين. وفي ظل كل هذه الضغوط والتحديات الاجتماعية الكبيرة، فقد بات واضحا أن تحقيق الطفل لانتقال سَلِس من بيئة البيت إلى بيئة المدرسة يتطلب أن يكون الطفل متمتعا بقدر كاف من الكفاءة الاجتماعية (الحسيني وشاكر، ٢٠١٤ Cassidy, (2005; Dockett & Perry, 2007.

وعلى الرغم من أن دراسات الكفاءة الاجتماعية قد بدأت منذ العقد الثالث للقرن العشرين من خلال البحث في طبيعة جماعات الأقران وخصائص الأطفال ووضعهم النسبي داخل هذه الجماعات وهو ما أفرز كما هائلا من التعريفات

* **د/ صبري محمود عبد الفتاح أمين**: قسم علم النفس التربوي- كلية التربية- جامعة المنيا.

لمفهوم الكفاءة الاجتماعية -ربما بسبب تعدد الأطر النظرية المفسرة لهذا المفهوم- إلا أن أي من هذه التعريفات لم يحظ باتفاق الباحثين (Bornstein, Hahn, & Haynes, 2010; Cavell, 1990). فأحد تعريفات الكفاءة الاجتماعية يصفها بأنها القدرة على الإيفاء بمتطلبات المواقف الاجتماعية من خلال تنظيم السلوك والإدراك والوجدان (Conduct Problems Prevention Research Group; CPPRG, 1999) وهناك تعريف آخر أحدث للكفاءة الاجتماعية يصفها بأنها مجموعة من المهارات والسلوكيات الضرورية حتى يتوافق الفرد مع الآخرين ويحظى بقبولهم (Blumberg, Carle, O'Connor, Moore, & Lippman, 2008). والدراسة الحالية تتبني هذا التعريف للكفاءة الاجتماعية لأنه يتوافق مع عدة تعريفات أخرى قدمها عدد من الباحثين وجميعها تصف الكفاءة الاجتماعية على أنها التفاعل الناجح مع الآخرين (Long, & Masten, 2008; Burt, Obradovic, Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011).

ومن هنا ترى سمرود-كليكمن (Semrud-Clikeman, 2007) أن المهارات والسلوكيات الاجتماعية المكونة للكفاءة الاجتماعية تعتمد على قدرات عقلية محددة مثل القدرة على فهم مشاعر الآخرين وتوجهاتهم، والقدرة على معرفة القواعد التي تحكم الحياة الاجتماعية وأدوار الأفراد، والقدرة على تبصر المواقف الاجتماعية وفهم أبعادها، وهذه المهارات والسلوكيات تعبر عن نفسها في صورة قدرة عامة على التفاعل الناجح مع الآخرين واستخدام الفرد للأدوات والأساليب الملائمة عند محاولة التأثير في سلوكيات المحيطين به.

وفي هذا الإطار يؤكد بعض الباحثين Missal & Hojnoski, 2008; Stoiber, (2011) على ضرورة التمييز بين نوعين من المهارات الاجتماعية عند الحديث عن الكفاءة الاجتماعية لدى الأطفال الملتحقين حديثا بالمدرسة الابتدائية وهذان النوعان هما: (١) المهارات الاجتماعية المرتبطة بالأقران و(٢) المهارات الاجتماعية المرتبطة بالتعلم. وتتضمن المهارات الاجتماعية المرتبطة بالأقران، مهارات التفاعل الشخصي الإيجابي مع الآخرين ومنها المشاركة واللعب التعاوني واحترام الأقران. أما المهارات المرتبطة بالتعلم فتتضمن المهارات التي يجب أن يمتلكها الطفل حتى يمكنه إنجاز ما يوكل إليه من أعمال صفية ولا صفية، ومنها الإنصات وإتباع التوجيهات والتعليمات والمشاركة الفعالة مع مجموعات الأقران ومهارات التفاوض. ويؤكد ميسال وهوجنوسكي (Missal & Hojnoski,2008)على

أن المهارات الاجتماعية المرتبطة بالأقران أو بالتعلم تتأثر بالخبرات السابقة التي عايشها الطفل في البيت والروضة والحضانة وأن هذين النوعين من المهارات يمكنهما تقرير مدى استعداد الطفل للالتحاق بالمدرسة الابتدائية.

وعلى الرغم من اختلاف الباحثين حول ماهية الكفاءة الاجتماعية، إلا أن هناك شبه إجماع على أن الكفاءة الاجتماعية تمثل مكونا نفسيا مهما، فنتائج العديد من الدراسات تشير إلى أن الكفاءة الاجتماعية تسهم بصورة إيجابية في الفاعلية الاجتماعية للأطفال وفي تكيفهم الأكاديمي وتحصيلهم الدراسي وإكمالهم لمرحلة الدراسة الابتدائية (Welsh, Parke, Widaman, & O 'Neil, 2001)، وأن ضعف أو غياب الكفاءة الاجتماعية يمكن أن يؤثر سلبا في قدرة الأطفال على تطوير أو الاحتفاظ بعلاقات شخصية سوية مع الآخرين (Ladd, 1999)، فيصبح الطفل أكثر عرضة لانخفاض التحصيل الأكاديمي (Welsh et al., 2001)، وضعف التكيف الاجتماعي، وجنوح الأحداث والأمراض النفسية والاكتئاب (Gresham, Elliott, Cook, Vance & Kettler, 2010). وفي إطار نظرية التعلم الاجتماعي، يري باندورا (Bandura, 2008) أن السلوك العدواني والانسحاب الاجتماعي والخجل هي سمات شخصية نشأت نتيجة افتقار الفرد إلى المهارات الاجتماعية الضرورية وهي مهارات يبدو أن الفرد لم يطورها وينميها أثناء طفولته لذا فإنه يصبح أكثر ميلا إلى السلوكيات المناقضة للكفاءة الاجتماعية مثل التخريب وعدم تقبل الآخرين.

استخدام المعلمين لموازين تقدير السلوك في قياس الكفاءة الاجتماعية:

وفي ضوء تراكم الأدلة البحثية التي تدعم أهمية الكفاءة الاجتماعية وارتباطها بالجوانب الاجتماعية والنفسية والأكاديمية السوية لحياة الأطفال، فإنه لم يكن مستغربا أن يبذل الباحثون جهودا حثيثة في محاولة لقياس مستوى ما يتمتع به الأطفال من كفاءة اجتماعية. وقد أثمرت هذه الجهود عن تطوير عدة طرق يمكن استخدمها في تقييم الكفاءة الاجتماعية لدى الأطفال ومنها الملاحظة، والمقابلة، والطرق السوسيومترية "القياسية الاجتماعية"، وطريقة حل المشكلات الاجتماعية، وأدوات التقرير الذاتي، وموازين تقدير السلوك (Semrud-Clikeman, 2007). وفيما يتعلق بموازين تقدير السلوك Behavioral Rating Scale، يشير علام (٢٠٠٢) إلى أن التقدير يتضمن الحكم على سلوك الفرد أو أحد خصائصه وسماته

في مواقف مضبوطة باستخدام موازين تقدير متدرجة ومحددة بحيث لا يتدخل الفرد في عملية التقدير أو الحكم، وبحيث يحدد القائم بالملاحظة قيمة رقمية للسلوك أو السمة أو الخاصية الملاحظة على ميزان التقدير. فميزان تقدير السلوك يُعد أسلوبا منظما للحكم على السلوك الملاحظ وتسجيله وتقريره، وبذلك يمكن التمييز بدرجة أكثر دقة بين الأفراد في سمة معينة.

وهناك عدة أسباب جعلت من طريقة موازين تقدير السلوك إحدى الطرق المهمة والأكثر شيوعا واستخداما في قياس الكفاءة الاجتماعية (Brown, Odom, & McConnell, 2008; Merrell, 2001; Semrud-Clikeman, 2007) منها: (١) يتطلب استخدام موازين تقدير السلوك وقتا قصيرا ولا تحتاج إلى تدريب مكثف قبل استخدامها مما يقلل من تكلفة استخدامها، (٢) تزودنا موازين تقدير السلوك بمعلومات عن أنماط السلوك المهمة والتي تحدث بصورة غير منتظمة والتي ربما لا يمكن رصدها بدقة من خلال أدوات قياس أخرى، (٣) تزودنا موازين تقدير السلوك ببيانات كمية موضوعية صادقة وثابتة أكثر من طرق القياس الأخرى وذلك من خلال ملاحظة السلوك أكثر من مرة في بيئته الطبيعية وعبر فترات زمنية ممتدة، (٤) يمكن استخدام موازين تقدير السلوك مع الأفراد الذين لا يستطيعون تزويدنا بمعلومات عن أنفسهم مثل الأطفال، (٥) يمكن لأي شخص مثل الوالدين، والمعلمين، والأقران وحتى المفحوصين أنفسهم أن يستخدم موازين تقدير السلوك في تقييم السلوك شريطة أن يكون على معرفة وثيقة بالمفحوص وهو ما يوفر معلومات تتمتع بقدر كبير من الصدق، (٦) تزودنا موازين تقدير السلوك ببيانات معيارية يمكن استخدامها في الحكم على مدى توفر السمة أو الخاصية لدى الطفل مقارنة بأقرانه.

وتهتم الدراسة الحالية بقياس الكفاءة الاجتماعية لدى الأطفال من خلال موازين تقدير السلوك التي يستخدمها المعلمون لعدة أسباب منها: (١) تتوفر لدى المعلم فرصة ملاحظة الطفل عبر فترات زمنية ممتدة خلال تفاعله مع الأطفال الآخرين داخل وخارج الحصص الدراسية مما يساعده على معرفة المهارات والسلوكيات الاجتماعية التي يتمتع بها الطفل بوجه عام، (٢) تتوفر لدى المعلم فرصة ملاحظة أطفال آخرين في نفس عمر الطفل وهو ما يتيح للمعلم فرصة الحكم على مدى ما يتمتع به الطفل من مهارات وسلوكيات اجتماعية مقارنة بأقرانه (بيانات معيارية)، (٣) قد لا يثير عامل ألفة الطفل وارتباطه بالمعلم أي

مخاوف لدى الطفل بأنه يتعرض للملاحظة والتقييم فيلجأ إلى تغيير سلوكه، (٤) يُعد المعلم مؤهلا تربويا ونفسيا لرصد وملاحظة سلوكيات الطفل خلال مراحل نمائية محددة بحيث يستطيع التمييز بين السلوكيات التي تعكس خصائص وسمات نفسية واجتماعية متأصلة لدى الطفل وتلك التي تمثل مجرد سلوكيات عابرة (Eisenberg, Fabes, & Spinrad, 2006; Milfort & Greenfield, 2002; Renk .& Phares, 2004; Semrud-Clikeman, 2007)

مقياس الكفاءة الاجتماعية المُعَّدل:

طور ريدل وهاجيكول وبولن (Rydell, Hagekull, & Bohlin 1997) النسخة الأصلية الأولى لمقياس الكفاءة الاجتماعية Social Competence Inventory والذي يمكن أن يستخدمه المعلمون والوالدون في تقييم الكفاءة الاجتماعية لدى الأطفال من عمر ٧ إلى١٠ سنوات. ويتكون المقياس من ٢٥ عبارة موزعة على عاملين هما: التوجه الاجتماعي الإيجابي (١٧ عبارة)، والمبادرة الاجتماعية (٨ عبارات). وتصف العبارات المكونة لعامل التوجه الاجتماعي الإيجابي مجموعة من السلوكيات التي تدعم التفاعل الاجتماعي الإيجابي للطفل ومنها حب المساعدة والكرم والتعاطف والتفهم الاجتماعي والتعاون ومعالجة الصراعات. ومن أمثلة عبارات عامل التوجه الاجتماعي الإيجابي "*غالبا ما يكون قادرا على إيجاد الحلول أو التسويات عندما يتدخل في صراع*" وتصف العبارات المكونة لعامل المبادرة الاجتماعية سلوكيات المبادرة في مقابل الانسحاب من المواقف الاجتماعية لدى الطفل. ومن أمثلة عبارات عامل المبادرة الاجتماعية "*غالبا ما يقترح أنشطة وألعاب للهو مع أقرانه*". ويتم تقيم الطفل على عبارات المقياس على تدريج خماسي يتكون من خمس نقاط هي "لا تنطبق مطلقا- لا تنطبق – تنطبق إلى حد ما- تنطبق – تنطبق تماما" ويقابلها الدرجات ١، ٢ ، ٣ ، ٤ ، ٥ على الترتيب للعبارات الموجبة ٥، ٤، ٣، ٢، ١ للعبارات السالبة، وتشير الدرجة المرتفعة إلى أن الطفل يتمتع بمستوى عال من الكفاءة الاجتماعية.

وحديثا قام ريدل وهاجيكول (Rydell & Hagekull, 2012) بتطوير مقياس الكفاءة الاجتماعية المُعَّدل Social Competence Inventory-Revised وذلك حتى يتم استيعاب التطورات النظرية الجديدة لمفهوم الكفاءة الاجتماعية لدى الأطفال

ولتحسين الخصائص السيكومترية للمقياس. فقد استخدم الباحثان التحليل العاملي التوكيدي لحساب صدق البناء، ومصفوفة السمات المتعددة والطرق المتعددة Multitrait-Multimethod Matrix لحساب الصدق التقاربي، والصدق التمييزي، ومعاملات الارتباط لحساب ثبات إعادة التطبيق، ومعامل كرونباخ الفا لحساب ثبات الاتساق الداخلي وذلك كله من خلال عينات زادت عن ١١٠٠ طفل تتراوح أعمارهم بين ٧ إلى١٠ سنوات. وقد أبقى الباحثان في النسخة المُعَّدله من المقياس على عبارات عامل التوجه الاجتماعي الإيجابي كما هي دون تغيير (١٧ عبارة)، ولكنهما قاما بإضافة ٧ عبارات جديدة لعامل المبادرة الاجتماعية وبذلك يصبح المقياس المُعَّدل مكون من ٣٢ عبارة موزعة على عاملين هما: التوجه الاجتماعي الإيجابي (١٧ عبارة) والمبادرة الاجتماعية (١٥ عبارة). ويتم تقييم الطفل على عبارات المقياس على نفس التدريج الخماسي المستخدم في النسخة الأصلية.

وتجدر الإشارة إلى أنه توجد في أدبيات علم النفس عدد من المقاييس الشهيرة والتي يمكن للمعلم استخدامها في تقييم الكفاءة الاجتماعية لدى الأطفال ممن هم في سن ٧ إلى١٠ سنوات ومنها على سبيل المثال نظام تقدير المهارات الاجتماعية Social Skills Rating System (SSRS, Gresham & Elliott, 1990)، ومقاييس السلوك الاجتماعي المدرسي School Social Behavior Scales (SSBS, Merrell, 1993)، ومقاييس تقديرات المعلم Teacher Rating Scales (TRS) كأحد المقاييس الفرعية في نظام التقييم السلوكي للأطفال- الإصدار الثاني (Reynolds &Kamphaus, 2014, Behavior Assessment System for Children-2, BASC-2) إلا أن الباحث في الدراسة الحالية فضّل استخدام مقياس الكفاءة الاجتماعية المُعَّدل لعدة أسباب: (١) سهولة استخدامه ومن ثم فهو لا يحتاج إلى تدريب خاص مما يوفر الوقت والجهد والتكلفة، (٢) قلة عدد فقراته نسبيا ومن ثم فهو لا يستغرق وقت طويل في تطبيقه، (٣) يأخذ في اعتباره أخر التطورات النظرية لمفهوم الكفاءة الاجتماعية لدى الأطفال، (٤) تركز عباراته على النجاح الاجتماعي الذي يحققه الطفل من خلال التفاعل الإيجابي مع المحيطين به وخاصة أقرانه ومن ثم فإن محتوى العبارات يتناسب مع التعريف الذي تتبناه الدراسة الحالية للكفاءة الاجتماعية بوصفها مجموعة من المهارات والسلوكيات الضرورية حتى يتوافق الفرد مع الآخرين ويحظى بقبولهم (Blumberg et al., 2008)، (٥)

عباراته مصاغة بطريقة سهلة وسلسة وواضحة مما سهل من عملية ترجمته من اللغة الإنجليزية إلى اللغة العربية.

نظرية القابلية للتعميم Generalizability Theory:

صاغ كرونباخ وزملاؤه (Cronbach, Gleser, Nanda, & Rajaratnam, 1972; Cronbach, Rajaratnam, & Gleser, 1963) نظرية القابلية للتعميم بوصفها نظرية إحصائية تهدف في الأساس إلى التحقق من ثبات Dependability (Reliability) القياسات التي نحصل عليها باستخدام أدوات القياس المختلفة مثل الاختبارات، والاستبيانات، وموازين التقدير، واستمارات الملاحظة. ويشرح برينن (Brennan, 2001) مفهوم الثبات في إطار نظرية إمكانية التعميم بأنه دقة التعميم من الدرجة الملاحظة Observed Score للمفحوص على الاختبار مثلا أو أي أداة قياس أخرى إلى متوسط الدرجة التي سيحصل عليها المفحوص تحت كل الشروط الممكنة مثل كل صور الاختبار الممكنة، أو كل مرات تطبيق الاختبار الممكنة، أو كل أسئلة الاختبار الممكنة. ويطلق على متوسط الدرجة هذه الدرجة الشاملة Universe Score وهي تشبه الدرجة الحقيقة True Score في النظرية الكلاسيكية للقياس، وعليه فإن نظرية إمكانية التعميم ترى أن الدرجات الملاحظة ثابتة إذا ما أمكن من خلال هذه الدرجات الوصول إلى استنتاجات دقيقة عن نطاق الملاحظات المقبولة Universe of Admissible Observation التي تمثلها.

وفي إطار نظرية إمكانية التعميم، فإن الدرجة الملاحظة هي مجرد عينة من نطاق الملاحظات المقبولة، وهذا النطاق يتضمن كل الملاحظات الممكنة التي ستكون مقبولة كبديل للملاحظة الحالية. فمثلا، لو طبق أحد الباحثين مقياسا لتقدير الذات يتكون من ١٥ عبارة على أحد المفحوصين ولكن الباحث أراد أن يستبدل هذه العبارات بـــ ١٥ عبارة أخرى تقيس أيضا مفهوم الذات فإن نطاق الملاحظات المقبولة هو كل العبارات التي تقيس تقدير الذات والتي يمكن أن يقبل بها الباحث، وبالطبع لن يستطيع المفحوص أن يجيب على كل العبارات التي تقيس تقدير الذات والتي يشملها النطاق ولكن نظرية التعميم وتحديدا من خلال معامل إمكانية التعميم Generalizability Coefficient يمكنها تقدير إلى مدى تقترب الدرجة الملاحظة (الدرجة التي يحصل عليها المفحوص من خلال الإجابة على ١٥عبارة تقيس تقدير الذات) من الدرجة الشاملة (الدرجة التي من الممكن أن يحصل عليها

المفحوص لو حدث وأجاب على كل العبارات التي تقيس تقدير الذات والتي يقبلها الباحث (Brennan, 2001, 1997; O'brien,1995).

ومن ناحية أخرى ترى النظرية الكلاسيكية في القياس أن الدرجة الملاحظة للمفحوص من خلال أي وسيلة قياس يمكن تقسيمها إلى درجة حقيقية واحدة وخطأ أحادي عشوائي غير مفسر، ومن ثم تهتم كل طريقة من طرق حساب الثبات بتقدير تأثير مصدر واحد فقط من مصادر هذا الخطأ العشوائي في كل مرة لمعرفة تأثيره على الدرجة الملاحظة. فمثلا، تهتم طريقة إعادة التطبيق Test-Retest بتأثير مرات التطبيق Occasions، في حين تهتم طريقة الصور المتكافئة Parallel Forms بتأثير صور الاختبار Forms، وتهتم طريقة الاتساق الداخلي Internal Consistency بتأثير العبارات Items (Anastasi, *1976; Crocker & Algina, (2006;* Nunnally & Bernstein, 1994). وفي كل الأحوال، فإن النظرية الكلاسيكية للقياس تزودنا بمعلومات محدودة للغاية عن الثبات، حيث لا يمكن استخدام المعلومات المتوفرة عن نوع ما من أنواع الثبات (مثلا ثبات إعادة التطبيق) للوصول إلى استنتاجات خاصة بنوع آخر (مثلا ثبات الاتساق الداخلي)، وحتى في حالة حساب أكثر من نوع من أنواع الثبات فإنه يظل من الصعب استخدام هذه المعلومات مجتمعة لتقرير عدد مرات تطبيق الاختبار، أو صور الاختبار، أو عبارات الاختبار التي نحتاجها للحصول على قياسات ثابتة. ومن الملاحظات المهمة في النظرية الكلاسيكية للقياس أن للمفحوص درجة حقيقة True Score واحدة فقط تمثل متوسط الدرجات التي يحصل عليها إذا طُبق عليه عدد لا نهائي من الاختبارات المتوازية. والاختبارات المتوازية أو المتكافئة Parallel Measures هي اختبارات يتماثل محتواها ويتساوى متوسطات وتباين درجاتها، وكذلك تتساوي الارتباطات البينية لدرجات مفرداتها المتناظرة (*Hand, 2010;* Nunnally & Bernstein, 1994). وعليه فإن نظرية إمكانية التعميم هي بمثابة امتداد للنظرية الكلاسيكية في القياس من حيث إنها تقدم إطارًا لتقدير تأثير مصادر عديدة لخطأ القياس بصورة متزامنة. ويطلق على مصادر خطأ القياس أوجه Facets، ويطلق على مستويات هذه الأوجه شروط Conditions، ونلاحظ أن مصطلحا "أوجه" و"شروط" يشبها مصطلحا "العوامل" Factors و"المستويات" Levels في أدبيات التصميمات التجريبية (Brennan, 2000). ويشير كون (Cone, 1977) إلى ستة أوجه يجب أن يهتم الباحثون بالتعميم خلالها في إطار

نظرية إمكانية التعميم وهي: الفاحص (المُقيّم)، والعبارة، والوقت، والسياق، والطريقة، والبعد، وهذه الأوجه لها نظائرها أو ما يقابلها في النظرية الكلاسيكية للقياس. فالتعميم من خلال الفاحص في نظرية إمكانية التعميم يقابله في النظرية الكلاسيكية ثبات الملاحظ (المحكم) Inter-Observer reliability أو الاتفاق بين الملاحظين (المحكمين) Inter-Observer Agreement، أما تعميم الوقت فيناظره ثبات إعادة التطبيق Test-Retest Reliability، ويناظر تعميم السياق الصدق المرتبط بالمحك Criterion-Related Validity، ويناظر تعميم العبارة ثبات الاتساق الداخلي أو صدق البناء Internal Consistency Reliability or Construct Validity، أما تعميم الطريقة فيناظره الصدق التقاربي Convergent Validity، وأخيرا يناظر تعميم البُعد الصدق التمييزي Discriminant Validity (Gresham &Carey, 1988).

وتؤكد نظرية إمكانية التعميم على أن هناك مصادر متعددة للخطأ في موقف القياس منها مثلا خصائص المفحوصين، وخصائص الفاحصين، والنسخة المطبقة من الاختبار، ووقت تطبيق الاختبار، وعدد مرات تطبيق الاختبار، وظروف تطبيق الاختبار، ونوع فقرات الاختبار ...إلخ وأن هذه الأخطاء قد تحدث معا بصورة متزامنة أثناء عملية القياس، فنظرية إمكانية التعميم تري أن الدرجة الملاحظة يمكن تقسيمها إلى درجة شاملة وواحد أو أكثر من مكونات الخطأ (Shavelson & Webb, 1991). ومن هذا المنطلق لن يكون من المناسب للباحث الاعتماد على درجة واحدة أُخذت من موقف واحد ووفق ظروف دون ظروف واعتمدت على وجه واحد (مصدر خطأ) دون بقية الأوجه مثلما هو الحال في النظرية الكلاسيكية للقياس، بل لابد من الأخذ في الاعتبار مفهوم شامل للأوجه التي يشتمل عليها موقف الاختبار فنأخذ في الاعتبار كل الأوجه أو بعضها في ضوء مبررات قوية وواضحة، وعليه فإن الدرجة الحقيقة في هذه الحالة ستكون مرتبطة بل ومعتمدة على نوعية النطاق الذي يحدده الباحث (نوعية الأوجه أو الأبعاد التي يأخذها الباحث في اعتباره) كمنطلق من أجل الحكم على درجة الفرد.

فوفق تصور نظرية إمكانية التعميم سيكون لكل مقياس عدة درجات حقيقة وكل درجة ستكون مرتبطة بنطاق محدد (واحد أو أكثر من الأوجه أو الأبعاد إلى يحددها الباحث). فقد تُنبي الدرجة الحقيقة على أساس جنس المفحوصين، أو

عمرهم، أو مؤهلاتهم العلمية، أو جنس الفاحصين، أو نوع الاختبار، أو نوعية الأسئلة، أو زمن تطبيق الاختبار، أو عدد مرات تطبيق الاختبار، أو النسخة المطبقة من الاختبار، أو فترة أو مكان التطبيق، ويجب أن يكون لدى الباحث مبرر وتفسير قوي عند انتقائه لمصادر أو أوجه معينة دون غيرها بهدف فحص تأثيرها في الدرجة الملاحظة (Brennan, 2001; Shavelson & Webb, 1991) . وهنا ينوه كرونباخ وزملاؤه (Cronbach et al., 1972) إلى أن فشل الباحث في تضمين أي وجه ضروري ومهم من أوجه خطأ القياس في عملية تحليل الثبات قد ينتج عنه زيادة التباين غير المفسر Undifferentiated Variance ويسمي بالبواقي Residuals، فعلى الباحث أن يحدد وبصورة دقيقة الطريقة التي يرغب بها تعميم نتائج القياس، فهل يهمه التعميم من موقف واحد إلى مواقف متعددة؟ من مُقيّم واحد إلى عدة مُقيّمين؟ من عدد محدود من العبارات إلى عدد أكبر من العبارات؟ فالإجابة على هذه الأسئلة ستختلف قطعا من بحث إلى أخر ومن ثم ستحدد تصميم دراسة التعميم.

وتعتمد نظرية إمكانية التعميم على استخدام الأسلوب الإحصائي تحليل التباين Analysis of Variance ليس بهدف اختبار الفروض ولكن بهدف تقدير قيمة مكونات تباين الدرجة الملاحظة Observed Score Variance والتي يمكن أن يسهم فيه كل من التباين الحقيقي (وهو تباين مصدره الفرد ويمثل الفروق الحقيقة الفعلية بين الأفراد في السمة المقاسة ويطلق عليه (Measurement Object) وتباين الخطأ وهو تباين مصدره كل وجه من أوجه الخطأ، والتفاعل بين أوجه الخطأ مع بعضها البعض، والتفاعل بين أوجه الخطأ وبين مصدر التباين الحقيقي (الفرد). ويأمل الباحث أن يسهم التباين الحقيقي بأكبر قدر ممكن في التباين الكلي للدرجات الملاحظة مقارنة بتباين الخطأ وهو ما يعني ضمنيا انخفاض معدلات خطأ القياس وارتفاع مستوى الثبات (Brennan, 2001; Shavelson & Webb, 1981).

دراسات القرارات (D):

تهدف دراسات القرارات إلى الوصول إلى أفصل إجراء قياس يتمتع بالثبات في موقف معين وذلك اعتمادا على المعلومات المتوفرة من دراسات إمكانية التعميم G Studies. وبمعني آخر فإن دراسات القرارات تعمل على معالجة Manipulation مستويات وجه أو أوجه خطأ القياس التي ضمنها الباحث في دراسات إمكانية التعميم بصورة متزامنة بهدف الوصول إلى أفضل الطرق

وأكثرها كفاءة لاستخدام أداة القياس – أي بهدف تقليل أخطأ القياس لتحقيق هدف معين(Brennan, 2001) . فالتساؤل الذي تهتم دراسات القرارات بالإجابة عليه هو كيف الوصول بثبات الاختبار إلى أفضل مستوى؟ فمثلا، من الممكن باستخدام دراسات القرارات أن نعرف التأثيرات في ثبات القياسات إذا ما تم استخدام مقياس يشتمل على ٢٠ عبارات بدلا من ١٠ عبارة و/أو إذا كان عدد المفحوصين ١٠٠ مفحوص بدل من ٥٠ و/أو إذا كان عدد المُقيِّمين ثلاثة بدلا من اثنين...إلخ- أي أن التساؤل المهم بالنسبة لدراسات القرارات هو: ماذا يمكننا القيام به إذا ما أردنا الاعتماد على هذه القياسات لاتخاذ قرارات ما في المستقبل؟ فمثلا، إذا ما أظهرت نتائج دراسات إمكانية التعميم أن العبارات تسهم بقدر ضئيل في خطأ القياس فربما يلجأ الباحث من خلال دراسات القرارات إلى أسلوب يعمل من خلاله على تقليل عدد هذه العبارات أو أن يغير من طريقة جمع البيانات من التصميم المتقاطع Crossed Design وهو تصميم يعتمد على أن يجيب كل فرد من أفراد عينة الدراسة على عبارات الاستبيان إلى التصميم المتداخل Nested Design وهو تصميم يعتمد على أن يجيب كل فرد من أفراد عينة الدراسة على مجموعة مختلفة من عبارات الاستبيان. وفي حالة لم يجد الباحث من خلال دراسات القرارات ما يمكن أن يعدله في القياسات في ضوء نتائج دراسات إمكانية التعميم فإنه يمكن القول بأن دراسات إمكانية التعميم ودراسات القرارات تؤدي نفس الوظيفة (Shavelson & Webb, 1981). وينتج من دراسات القرارات معامل إمكانية التعميم Coefficients of Generalizability وهو بمثابة نسبة تباين الدرجة الشاملة Universe Score إلى تباين الدرجة الملاحظة Observed Score، وتتراوح قيمه ما بين ٠ و+١ بحيث يعكس اقتراب قيمة المعامل من +١ ثبات القياسات- أي أنه يمكننا اعتبار معامل إمكانية التعميم بديلا عن معامل الثبات الذي نحصل عليه باستخدام النظرية الكلاسيكية للقياس. وكل اختبار يكون له أكثر من معامل إمكانية تعميم واحد وذلك على حسب مصادر الخطأ التي يتم أخذها في الاعتبار عند إجراء دراسة إمكانية التعميم (Brennan, 2001; O'brien, 1995).

الدراسات السابقة:

أجرى كرست (Christ, 2012) دراسة استخدم فيها نظرية إمكانية التعميم بهدف التحقق من ثبات تقييمات معلمتين لبعض المهارات الاجتماعية لدى عينة

شملت ٢٠ طفلا (ذكور = ١٢، اناث = ٨) تراوحت أعمارهم بين ٧ و١١ سنة ومسجلين في فصلين في مدرسة حكومية في الولايات المتحدة الأمريكية. وقامت المعلمتان بتقييم ستة مهارات اجتماعية لدى عينة الدراسة مرتين وبفاصل زمني قدره ٣ أسابيع بين التقييمين. وشملت المهارات الاجتماعية ٦ مهارات أساسية هي التعاون، والتواصل، والإيثار، وحل الخلافات، وإدارة الصراعات والتفاعل مع الغرباء وجميعها جزء من موازين تقدير نظام تحسين المهارات الاجتماعية Social Skills Improvement System – Rating Scales (Gresham & Elliott, 2008). وأظهرت نتائج دراسات إمكانية التعميم G- Studies أن نسبة التباين الحقيقي الذي يصف الفروق الحقيقية بين الأطفال في هذه المهارات كانت ٣٧,٦% من التباين الكلي في الفصل الأول و٣٤,٢% في الفصل الثاني، وأن المُقيّمتين (المعلمتين) لم يفسرا أي نسبة من التباين الكلي في أي من الفصلين، وأن أبرز التفاعلات كان بين الفرد ووقت التطبيق والذي فسر ١٣,٢% من التباين في الفصل الأول و٩,٦% في الفصل الثاني، وأن أقل التفاعلات تأثيرا كانت بين المُقيّمتين والعبارات حيث فسر هذا التفاعل ٢%. وأظهرت نتائج دراسات القرارات D-Studies أنه على مستوي الفصلين فإن استخدام مُقيّمتين وتطبيق المقياس (٦ مهارات) مرتين يُعد إجراءً كافيا للحصول على أعلى قيم معاملات إمكانية التعميم، وأن قيم معاملات إمكانية التعميم تكون أقل ما يمكن عند استخدام مُقيّمه واحدة وتطبيق المقياس مرتين.

وهدفت دراسة كارمنلا (Carminla, 2010) إلى مقارنة ثبات تقييمات المعلمات لأربع مهارات اجتماعية ضمن موازين التقدير الخاصة بنظام تحسين المهارات الاجتماعية Social Skills Improvement System – Rating Scales (Gresham & Elliott, 2008) وذلك بالنسبة للأطفال في مرحلة الروضة (٤-٥ سنوات) والأطفال في المدرسة الابتدائية (٦-٧ سنوات) في بريطانيا. وشملت عينة أطفال الروضة 17 طفلا (ذكور = ٩، إناث = ٨) قامت معلمتان بتقييمهم على أربع مهارات أساسية هي المحادثة، والتفهم الاجتماعي، ومشاركة المجموعة، والتعزيز الاجتماعي مرتين وبفاصل زمني قدره شهرين بين التقييمين. وشملت عينة أطفال المدرسة الابتدائية ١٩ طفلا (ذكور = ٩، إناث = ١٠) قامت معلمتين بتقييمهم على نفس المهارات الاجتماعية مرتين وبفاصل زمني قدره شهرين بين التقييمين. وأظهرت نتائج دراسات إمكانية التعميم G-Studies أن المُقيّمات

(المعلمات) ووقت التطبيق فسروا نسبة ضئيلة من التباين في عينتي أطفال الروضة والمدرسة الابتدائية، وأن العبارات فسرت نسبة أعلى من التباين في عينة أطفال المدرسة الابتدائية (١٦,٥%) منه في عينة أطفال الروضة (١٣,٤%). وكان التفاعل بين المُقيّمتين ووقت التطبيق أكبر مصادر الخطأ حيث فسر ٢٠,٢% من التباين. وأظهرت نتائج دراسات القرارات D-Studies أنه أمكن في كلتا العينتين الحصول على معاملات تعميم مرتفعة من خلال استخدام ثلاث مُقيّمات في كل عينة وتطبيق أداة الدراسة مرتين وليس العكس.

وقام فريق من الباحثين من مركز تطوير التعلم والمهارات الذاتية في جامعة أوساكا في اليابان (Lueng, Lui, & Yim, 2010) بفحص ثبات تقييمات ثلاثة معلمين للمهارات الاجتماعية لدى الأطفال الملتحقين حديثا بالمدرسة (متوسط العمر = ٦,٤ سنة) وذلك على عينة بلغت ١٧٢ طفلا (ذكور = ٨٠، إناث = ٩٠) في احدي المدارس الحكومية مستخدمين النسخة اليابانية لمقياس قائمة أداء الكفاءة الاجتماعية(Social Competence Performance Checklist (Stoiber, 2004 . وتمت عملية التقييم مرتين وبفاصل زمني قدره شهر بين التقييمين. وأظهرت نتائج دراسات إمكانية التعميم G-Studies أن نسبة التباين الحقيقي الذي يصف الفروق الحقيقية بين الأطفال في هذه المهارات كانت ٣٣,٦% من التباين الكلي، وأن عبارات المقياس لم تفسر سوى ٣% من التباين، وأن أبرز التفاعلات كان بين المُقيّم (المعلم) ووقت التطبيق والذي فسر ١٤,٣% من التباين. وأظهرت نتائج دراسات القرارات D-Studies أن استخدام ثلاثة مُقيّمين وتطبيق المقياس مرتين يُعد إجراءً كافيا للحصول على قيم مرتفعة لمعاملات إمكانية التعميم وذلك على الرغم من أن استخدام ثلاثة مُقيّمين وتطبيق المقياس ثلاث مرات يمكن أن يزيد من قيم معاملات إمكانية التعميم بمقدار ٢% إلا أن الباحثين اعتبروا أن هذه النسبة ضئيلة وغير مؤثرة ولا تستحق زيادة عدد مرات التطبيق من مرتين إلى ثلاث مرات. وكذلك أظهرت نتائج دراسات القرارات D-Studies أن قيم معاملات إمكانية التعميم تكون أقل ما يمكن عند تطبيق المقياس مرة واحدة حتى في حالة وجود أثنين أو ثلاثة مُقيّمين.

وفي دراسة أخري أجرها بروجر وسانتمين وكلفن (Broger, Saintmeen, & Keleven, 2012) وهدفت إلى التحقق من ثبات تقييمات ثلاثة معلمين للمهارات

الاجتماعية لدى عينة بلغت ٩٦ طفلا (ذكور= ٤٣، إناث = ٤٣) تتراوح أعمارهم ١١ و١٢ سنة في احدي المدارس الحكومية في السويد وذلك بهدف استخدام هذه التقييمات في اختيار فرق المساعدة والكشافة في المدرسة. وقد قام أربعة معلمين بتقييم الأطفال على موازين تقدير المهارات الاجتماعية Social Skills Rating System (Gresham & Elliott, 1990). وقد تم جمع البيانات على ثلاث جولات وبفاصل زمني قدره شهرين بين كل جولة وأخرى. وأظهرت نتائج دراسات إمكانية التعميم G-Studies أن نسبة التباين الحقيقي الذي يصف الفروق الحقيقية بين الأطفال في هذه المهارات كانت ٣٧,٨% من التباين الكلي، وأن مرات التطبيق فسرت نسبة ضئيلة من التباين لم تتجاوز ٦%، أما المُقيّمين (المعلمين) فقد فسروا ٩% من التباين. وكان أبرز التفاعلات بين المُقيّمين والفرد والذي فسر ١٣,٨% من التباين يليه التفاعل بين المُقيّمين والعبارات والذي فسر ١٠,٤%، بينما كان أقل التفاعلات تأثيرا بين المُقيّمين ووقت التطبيق حيث فسر ١,٧%. ووجد الباحثون أن نسبة التباين غير المفسر Residual هي ٢٢,٦%. وأظهرت نتائج دراسات القرارات D-Studies أن استخدام أربعة مُقيّمين وتطبيق المقياس ثلاث مرات كان إجراءً غير كافيا للحصول على قيم مرتفعة لمعاملات إمكانية التعميم وأنه من الأفضل زيادة عدد المُقيّمين إلى خمسة مُقيّمين وزيادة عدد مرات التطبيق إلى أربع مرات حيث ستبلغ قيمة معامل إمكانية التعميم عندئذ ٠,٩٣٢ وكذلك أظهرت نتائج دراسات إمكانية التعميم أن قيم معاملات إمكانية التعميم تكون أقل ما يمكن عند تطبيق المقياس مرة واحدة حتى في حالة وجود ما بين واحد إلى أربعة مُقيّمين.

وأجرى ليان وهارمند (Lian & Harmand, 2015) دراسة استخدما فيها نظرية إمكانية التعميم بهدف التحقق من ثبات تقييمات ثلاثة من الأخصائيين الاجتماعيين في إحدى دور رعاية الأيتام لبعض المهارات الاجتماعية لدى عينة شملت ٢٧ طفلا (٧-٩ سنوات) في سنغافورة. وتضمنت هذه المهارات التنظيم السلوكي، التعاون الجمعي، والتخطيط مع المجموعة، والوصول إلى تسويات في الصراعات وذلك ضمن موازين التقدير الخاصة بنظام تحسين المهارات الاجتماعية Social Skills Improvement System – Rating Scales (Gresham & Elliot, 2008). وقد تم تقييم الأطفال مرتين وبفاصل زمني قدره شهرين بين التقييمين. وأظهرت نتائج دراسات إمكانية التعميم G-Studies أن نسبة التباين

الحقيقي الذي يصف الفروق الحقيقية بين الأطفال في هذه المهارات كانت ٣٥,٢% من التباين الكلي، وأن مرات التطبيق فسرت نسبة ضئيلة من التباين لم تتجاوز ٤%، أما المُقيّمين (المعلمين) فقد فسروا نسبة ملحوظة من التباين بلغت ١٣%. وقد كان لافتا للانتباه أن حجم التباين غير المفسر Residual بلغ نسبة مرتفعة قدرها ٣٠%. وكان أبرز التفاعلات بين العبارات ووقت التطبيق والذي فسر ١١,٥% من التباين يليه التفاعل بين المُقيّمين ومرات التطبيق والذي فسر ٨,٢%. وأظهرت نتائج دراسات القرارات D-Studies أن استخدام ثلاثة مُقيّمين وتطبيق المقياس مرتين كان إجراءً كافيا للحصول على قيم مرتفعة لمعاملات إمكانية التعميم وأن تحسين الثبات يمكن أن يتم عن طريق زيادة عدد المُقيّمين وليس زيادة عدد مرات تطبيق المقياس. وكذلك أظهرت نتائج دراسات إمكانية التعميم أن قيم معاملات إمكانية التعميم تكون أقل ما يمكن عند تطبيق المقياس مرة واحدة حتى مع زيادة عدد المُقيّمين إلى خمسة مُقيّمين.

وهدفت دراسة لوكا وشينج (Luka & Cheung, 2013) إلى التحقق من ثبات تقييمات المعلمين للمهارات الاجتماعية في مواقف المنافسة والمسابقات الرياضية لدى عينة مكونة من ٤٢ طفلا (ذكور = ٢٤ وإناث = ١٨) في إحدى المدارس الحكومية في شمال غرب الصين حيث قام أربعة معلمين بتقييم الأطفال على نظام تقدير المهارات الاجتماعية Social Kills Rating System (Gresham & Elliott, 1990) وذلك بعد إدخال بعض التعديلات الطفيفة على المقياس حتى يتناسب وموقف جمع البيانات. وقد تمت عملية تقييم المهارات الاجتماعية للأطفال مرتين؛ في بداية ونهاية الفصل الدراسي. وأظهرت نتائج دراسات إمكانية التعميم - G Studies أن نسبة التباين الحقيقي الذي يصف الفروق الحقيقية بين الأطفال في هذه المهارات كانت ٤١% من التباين الكلي، وأن التفاعل بين العبارات والمُقيّمين (المعلمين) فسر ٢٧% من التباين يليه التفاعل بين مرات التطبيق والعبارات والذي فسر ١٩% من التباين. وفسرت العبارات نسبة ضئيلة من التباين لم تتعد ٢%. وأظهرت نتائج دراسات القرارات أن معامل الثبات يزداد بصورة ملحوظة مع ازدياد عدد المُقيّمين وأن أقصي معامل للثبات بلغ ٠,٩٢٥ مع ازدياد أعداد المُقيّمين إلى ٦ مقيمين وزيادة عدد مرات التطبيق إلى ٣ مرات. كما أظهرت

نتائج دراسة التعميم أنه بهدف تحسين الثبات فإنه من الأفضل زيادة عدد المُقيّمين بدلا من زيادة عدد مرات التطبيق.

وأجرت ريبر (Ribar, 2005) ثلاث دراسات مترابطة استخدمت فيها نظرية إمكانية التعميم بهدف التحقق من ثبات تقييمات المعلمات والمعلمات المساعدات للكفاءة الاجتماعية لدى الأطفال في الولايات المتحدة الأمريكية. وشملت الدراستين الأولي والثانية كل على حدة ١٩ طفلا مسجلين بمرحلة الروضة وموزعين على فصلين، وشملت الدراسة الثالثة ١٠ من طلبة الصف الرابع الأساسي في احدي المدارس الحكومية موزعين على فصلين. وقد قامت المعلمة والمعلمة المساعدة في كل فصل في كل عينة بتقييم الأطفال باستخدام قائمة أداء الكفاءة الاجتماعية Social Competence Performance Checklist (Stoiber, 2004) مرتين وبفاصل زمني قدره شهر بين التقييمين. وأظهرت نتائج إمكانية التعميم G-Studies أن إسهام المُقيم ووقت التطبيق والعبارات في التباين الكلي قد اختلف بين الدراسات الثلاث. وبوجه عام فإن التأثيرات الأساسية لهذه الأوجه الثلاثة لم تفسر نسب كبيرة من التباين، وأن أكثر مصادر الخطأ تأثيرا هو التفاعل بين الفرد والعبارة، يليه التباين غير المفسر Residual. كما أظهرت نتائج دراسات القرارات D-Studies أنه لصالح زيادة ثبات أداة القياس المستخدمة فإنه يجب زيادة عدد المقيميين أكثر منه زيادة عدد مرات التطبيق، وأن معاملات إمكانية التعميم ستكون مرتفعة مع زيادة أعداد المُقيمين وتطبيق المقياس مرة واحدة أكثر من زيادة عدد مرات التطبيق ووجود مُقيّم واحد فقط.

وأجرت مابي (Mabe, 2014) دراسة استخدمت فيها نظرية إمكانية التعميم بهدف تطوير أداة قياس للمهارات الاجتماعية. وشملت عينة الدراسة ٢٠ من طلبة الصف السادس حتى الثاني عشر (ذكور = ١٧، إناث = ٣) في احدي المدارس الحكومية في الولايات المتحدة الأمريكية. وقام ثلاثة معلمين بتقييم الطلاب أربع مرات على ٦ مهارات اجتماعية تمثل -وفقا للباحثة- سمة الكفاءة الاجتماعية وجميعها مشتقة من موازين التقدير الخاصة بنظام تحسين المهارات الاجتماعية Social Skills Improvement System – Rating Scales Gresham &) (Elliot, 2008، وهذه المهارات هي المحادثة الجماعية، المناقشة والإقناع، والمشاركة الصفية، ومشاركة المجموعة، والتفاعل مع الآخرين، وفض المنازعات. واستخدمت الباحثة تصميمًا ثلاثيًا الأوجه متداخل جزئيا حيث ضم مرات التطبيق

(٤) والمهارات (٦) كأوجه متقاطعة Crossed Facets والطلاب (٢٠) كأوجه متداخلة Nested Facets مع المُقيّمين (المعلمين) (٣). وأظهرت نتائج دراسات إمكانية التعميم G-Studies أن معامل إمكانية التعميم بلغ ٠,٩٣٦ في حين بلغت نسبة التباين غير المفسر Residual ٣٠%، وفسر الطلاب ٤٣% من التباين. وأظهرت نتائج دراسات القراراتD-Studies أن الوصول إلى أفضل مستويات الثبات لأداة القياس يتطلب زيادة عدد المُقيّمين إلى١٠ مُقيّمين وتطبيق المقياس ١٠ مرات حيث إن معامل الثبات المتوقع هو ٠,٩٤٨

مشكلة الدراسة:

تزخر أدبيات التربية وعلم النفس بالعديد من الدراسات التي اهتمت بتطوير برامج تدريبية خاصة بالكفاءة الاجتماعية لدى الأطفال بهدف إكسابهم مجموعة من المهارات الاجتماعية الأساسية التي يمكن أن تساعدهم في تحقيق التكيف الاجتماعي المنشود ومجابهة الضغوط والمتطلبات الاجتماعية الملقاة على عاتقهم في هذا السن المبكر لاسيما عند الانتقال من بيئة البيت إلى البيئة المدرسية. وتشمل هذه المهارات الاندماج مع الآخرين، والمشاركة، والتفاوض، والتفاعل الإيجابي، وفض النزاعات والخلافات، ومهارات التواصل والاتصال، واللعب التعاوني، واحترام الأقران (Durlak & Wells, 1997). وعلى الرغم من أن نتائج العديد من هذه الدراسات قد أكدت على فاعلية هذه البرامج التدريبية إلا أنه من المهم قبل الشروع في تطبيق مثل هذه البرامج وتحمل الوقت والجهد والتكلفة أن نتأكد أن لدينا أدوات قياس ذات خصائص سيكومترية جيدة يمكن استخدامها في تقييم الكفاءة الاجتماعية لدى الأطفال وبالطبع تقييم فاعلية هذه البرامج عقب الانتهاء من تطبيقها. وأحد الطرق التي يمكن أن تمكننا من توفير أدوات القياس الجيدة هو معرفة وتحديد كم خطأ القياس الذي يمكن أن يسهم به كل وجه من أوجه خطأ القياس والتفاعل بين هذه الأوجه في الدرجات الملاحظة التي نحصل عليها باستخدام أدوات القياس. ويمكن لنظرية إمكانية التعميم أن تزودنا بالإطار المفاهيمي والأساليب الإحصائية المناسبة لتحقيق هذا الهدف. وفي الدراسة الحالية فإن هذه الأوجه تشمل المعلمات والمعلمات المساعدات كمُقيّمات، وعبارات المقياس، ووقت التطبيق.

فمقاييس تقدير الكفاءة الاجتماعية لدى الأطفال والتي تعتمد على المعلمين كمُقيّمين لها أهمية خاصة لأغراض تحديد وتشخيص نواحي القصور في الكفاءة الاجتماعية لدى الأطفال حيث يُعد المعلمين مسؤولين عن معظم حالات الإحالة Referrals التي تتم في المدارس والتي تعاني قصورا ملحوظا في نواحي اجتماعية وسلوكية Caldarella & Merrell, 1997; Foster, Inderbitzen, & Nangle, 1993; Webster-Stratton) Lindsay, 1999،&) ومع هذا فهناك أدلة بحثية عديدة تشير إلى أن المعلمين كمُقيّمين يتدخلوا في عملية التقييم ومن ثم فهم أنفسهم قد يمثلون أحد أوجه خطأ القياس لاسيما إذا ما تطلب موقف التقييم أن ينظر المعلم أبعد من مجرد الملاحظة الفعلية للسلوك حتى يصدر حكمه وتقييمه. فتقييمات المعلم للطفل يمكن أن تتأثر بعلاقته الشخصية وتفاعله مع الطفل المُقيِم أو بتاريخ هذا الطفل وشهرته وإنجازاته الأكاديمية وهو ما يعرف بخطأ أثر الهالة Halo Effect Error، كما يمكن أن تتأثر تقييمات المعلم بسلوكيات الطفل نحو المعلم ونحو أقرانه على الرغم من أن نتائج الأبحاث تشير إلى اختلاف سلوك الطفل نحو أقرانه عنه نحو الكبار من حيث الشكل والوظيفة (Webster-Stratton & Lindsay, 1999). ومن ناحية أخرى يمكن أن تغيب الموضوعية عن تقييمات المعلم فنجدها تتصف بالليونة والتساهل مقارنة بزملائه وهو ما يعرف بخطأ التساهل Error of Leniency، أو أن تختلف تقييمات المعلم عن زملائه بسبب اختلاف البيئة الثقافية والاجتماعية والمستوى التعليمي وعامل الخبرة. فعلى سبيل المثال، عندما يجلس طفل هادئ ساكن أثناء فترة الاستراحة المدرسية (الفسحة) فإن ذلك قد يعتبر أمرًا طبيعيا بالنسبة لأحد المعلمين ولكنه قد يؤشر على وجود مشكلة في الكفاءة الاجتماعية لدى الطفل وفقا لوجهة نظر معلم آخر (Gresham, 1986, 1987; Renk & Phares, 2004; Semrud-Clikeman, 2007; Shah, & Morgan, 1996).

وبالنسبة للعبارات فهي قد تمثل مصدرا لخطأ القياس كون أن تقييمات المعلمين للكفاءة الاجتماعية لدى الأطفال يمكن أن تختلف في هذه المجموعة من العبارات عن مجموعة أخرى يتم معاينتها من نفس النطاق الواسع والشامل للعبارات. وكذلك فإن عينة العبارات قد تمثل مصدرا لخطأ القياس عند محاولة تعميم تقييمات الأطفال من هذه العينة من العبارات إلى مستوى نطاق أوسع وأشمل من العبارات وهو ما يعرف في أدبيات القياس النفسي بخطأ معاينة المحتوي Content Sampling Error (أبو علام، ٢٠٠٢؛ Crocker & Algina, 2006).

وبالنسبة لمرات التطبيق Occasions فإن إعادة تطبيق الاختبار يمثل عينة من نطاق أوسع واشمل لعدد مرات التطبيق Universe of Occasions. وتنظر النظرية الكلاسيكية في القياس إلى اتساق الأداء عبر الفترة الزمنية كمؤشر على ثبات واستقرار أداة القياس، أما نظرية إمكانية التعميم فتنظر لمرات التطبيق على أنها أحد أوجه خطأ القياس حيث إن المعلمين قد لا يُقيمون الكفاءة الاجتماعية لدى الأطفال بصورة متسقة عبر مرات التطبيق. فالسلوك الاجتماعي للطفل يمكن أن يختلف تبعا للموقف حيث إن أحد مظاهر الكفاءة الاجتماعية للطفل تتمثل في قدرته على التفاعل المناسب مع الآخرين في مواقف مختلفة. فالطفل قد يمتلك المهارات الاجتماعية الصحيحة ولكنه قد يستخدمها في التوقيت الخطأ أو قد يستخدمها فقط عندما يتعلق الأمر بفائدة وعائد شخصي. وبوجه عام فإن الطفل الكفء اجتماعيا هو من يستطيع أن يسلك على نحو يتناسب وطبيعة الموقف (Dennis & Stocka, 2015).

وفي ضوء ما سبق ذكره فإنه يتضح لنا أن قياس الكفاءة الاجتماعية لدى الأطفال محفوف بالعديد من أوجه خطأ القياس ومنها المعلمين كمُقيّمين، وعبارات المقياس، ووقت التطبيق، وأن الوصول إلى تقييمات حقيقية تعكس واقع مستوى الكفاءة الاجتماعية لدى الأطفال يتطلب بالدرجة الأولي تحديد ومعرفة أوجه خطأ القياس هذه، ومدى إسهام كل منها منفردا أو من خلال التفاعل فيما بينها في تباين الدرجات الملاحظة التي نحصل عليها عند تطبيق أدوات القياس المناسبة. وفي ضوء ما تم طرحه فإن مشكلة الدراسة الحالية تتمثل في تقدير قيم تباين الخطأ الذي يمكن أن يفسره كل من المعلمات والمعلمات المساعدات كمُقيّمات Raters، والعبارات (Items)، ووقت التطبيق (Occasion) كأوجه محتملة لخطأ القياس في التباين الكلي لدرجات الأطفال على النسخة العربية لمقياس الكفاءة الاجتماعية المُعَّدل، كما تشمل مشكلة الدراسة الحالية فحص الشروط التي تتيح لهذا المقياس تحقيق أفضل مستويات الثبات.

أسئلة الدراسة:

تحاول الدراسة الحالية الإجابة عن السؤالين التاليين:

(١) ما حجم تباين الخطأ الذي يفسره كل من أوجه (المُقيّمات (المعلمات والمعلمات المساعدات)، والعبارات، ووقت التطبيق) في التباين الكلي لدرجات الأطفال على النسخة العربية لمقياس الكفاءة الاجتماعية المُعَّدل؟

(٢) ما الشروط تحقيق النسخة العربية لمقياس الكفاءة الاجتماعية المُعَّدل لأفضل مستويات الثبات؟

أهداف الدراسة:

للدراسة الحالية هدفين رئيسين هما:

(١) الكشف عن حجم تباين الخطأ الذي يفسره كل من أوجه المُقيّمات (المعلمات والمعلمات المساعدات)، والعبارات، ووقت التطبيق في التباين الكلي لدرجات الأطفال على النسخة العربية لمقياس الكفاءة الاجتماعية المُعَّدل.

(٢) الكشف عن الشروط التي يمكن من خلالها أن تحقق النسخة العربية لمقياس الكفاءة الاجتماعية المُعَّدل أفضل مستويات الثبات.

أهمية الدراسة:

تتبلور أهمية الدراسة الحالية من خلال:

(١) تمثل الدراسة الحالية إضافة نوعية للمكتبة العربية في مجال القياس النفسي والتربوي نظرا لندرة البحوث العربية التي تناولت تطبيقات نظرية إمكانية التعميم في مجال قياس السمات النفسية والتربوية.

(٢) من مزايا الدراسة الحالية أنها تهتم بفحص ثبات أدوات القياس من منظور نظرية إمكانية التعميم، باعتبار أن الثبات يمثل أحد القضايا المحورية في مجال القياس النفسي والتربوي، وأن نظرية إمكانية التعميم تتمتع بمزايا عديدة في تقدير الثبات مقارنة بالنظرية الكلاسيكية في القياس.

(٣) تعزو أهمية الدراسة الحالية إلى أن قياس الكفاءة الاجتماعية لدى الأطفال يتم من خلال تقييمات المعلمات والمعلمات المساعدات وذلك لما لهذه الطريقة من مزايا مقارنة بطريقة التقرير الذاتي التقليدية.

(٤) تستمد الدراسة الحالية أهميتها من كون سمة الكفاءة الاجتماعية تمثل مكونا نفسيا مهما يمكن أن يساعد الطلاب الملتحقين حديثا بالمدرسة في تحقيق انتقال سلس من بيئة البيت إلى البيئة المدرسية مما يزيد من دافعيتهم للتعلم وتحصيلهم الدراسي ويدعم توافقهم النفسي والاجتماعي (Gresham et al., 2010).

التعريفات الإجرائية لمصطلحات الدراسة:

الكفاءة الاجتماعية: مجموع الدرجات التي يحصل عليها الطالب على النسخة العربية لمقياس الكفاءة الاجتماعية المُعَّدل حيث تدل الدرجة المرتفعة على مستوى مرتفع من الكفاءة الاجتماعية.

منهجية الدراسة:

منهج الدراسة:

تم استخدام المنهج الوصفي في الدراسة الحالية، حيث يهتم هذا المنهج بدراسة الظاهرة كما هي في الواقع، ووصفها وصفا دقيقا، ويعبر عنها تعبيرا كيفيا أو كميا. فالتعبير الكيفي يصف الظاهرة وخصائصها، أما التعبير الكمي فيقدم وصفا رقميا يوضح مقدار هذه الظاهرة أو حجمها ودرجة ارتباطها مع الظواهر الأخرى (عبيدات وعدس وعبد الحق، ٢٠٠٥).

العينة الاستطلاعية:

شملت العينة الاستطلاعية ١١٥ طفلا (ذكور = ٦٠، إناث = ٥٥، متوسط العمر = ٩,٤٧، انحراف معياري = ٠,٦٢) موزعين على ثلاثة فصول بالصف الثالث الأساسي بإحدى المدارس الحكومية للتعليم الأساسي(١-٤) في محافظة الداخلية بسلطنة عمان. وكذلك شملت العينة الاستطلاعية أربع معلمات فصل (متوسط عدد سنوات الخبرة التدريسية = ٦,٤، انحراف معياري = ٠,٨٥) حيث قامت كل معلمة فصل بتقييم الكفاءة الاجتماعية لدى الأطفال في الفصل المسئولة عن تدريسه باستخدام النسخة العربية لمقياس الكفاءة الاجتماعية-المُعَّدل وذلك لأن قوانين سلطنة عمان تسمح للإناث فقط بالتدريس في مرحلة التعليم الأساسي (١-٤). وتم اختيار جميع معلمات الفصل حاصلات على درجة البكالوريوس في التربية، وأن اللغة العربية هي لغتهن الأم.

ترجمة مقياس الكفاءة الاجتماعية المُعَّدل:

تم ترجمة مقياس الكفاءة الاجتماعية المُعَّدل (Rydell & Hagekull, 2012) من اللغة الإنجليزية إلى اللغة العربية باستخدام طريقة الترجمة المعكوسة (Maneersriwongul & Dixon, 2004). ووفقا لهذه الطريقة، قام اثنين من أعضاء هيئة التدريس المتخصصين في اللغة الإنجليزية وممن لغتهما الأصلية هي اللغة العربية وبطريقة مستقلة عن بعضهما البعض بترجمة المقياس من اللغة الإنجليزية

إلى اللغة العربية، ثم قام اثنين آخرين من أعضاء هيئة تدريس المتخصصين في اللغة الإنجليزية وممن لغتهما الأصلية هي اللغة العربية وبطريقة مستقلة عن بعضهما البعض بإعادة ترجمة المقياس من اللغة العربية إلى اللغة الإنجليزية. وقام اثنين من أعضاء هيئة التدريس المتخصصين في علم النفس التربوي وممن لغتهما الأصلية هي اللغة الإنجليزية وبطريقة مستقلة عن بعضهما البعض بمقارنة الصورة الإنجليزية الأصلية للمقياس بالصورة الإنجليزية المترجمة حيث تم الحكم على مدى تطابق معاني العبارات في النسختين وذلك باختيار أحد بديلين: "مطابق" ويعطى الدرجة ١ أو "غير مطابق" ويعطى الدرجة صفر. وقد تم حساب إحصائية كابا Kappa Statistics بهدف التحقق من ثبات المحكمين Interrater reliability. وتتميز إحصائية كابا عن النسبة المئوية بأنها تأخذ في اعتبارها الاتفاق بين المحكمين بمحض الصدفة وهي تصلح للاستخدام في حالة اثنين من المحكمين والبيانات التصنيفية الثنائية وتنحصر قيمها بين -١ و+١ (Cohen, 1960). وأوضحت لاندس وكوش (1977) Landis and Koch القيم الاسترشادية لإحصائية كابا وهي: أقل من صفر لا يوجد اتفاق مطلقا، من صفر إلى ٠,٢٠ اتفاق ضعيف، من ٠,٢١ إلى ٠,٤٠ اتفاق مقبول، من ٠,٤١ إلى ٠,٦٠ اتفاق متوسط، من ٠,٦١ إلى ٠,٨٠ اتفاق كبير، أعلى من ٠,٨١ اتفاق شبه تام. وقد بلغت قيمة كابا في الدراسة الحالية ٠,٩٤ وهو ما يشير إلى درجة مرتفعة جدا من ثبات المحكمين بما يعكس تشابه شبه تام بين الصورة الإنجليزية الأصلية والصورة الإنجليزية المترجمة لمقياس الكفاءة الاجتماعية المُعَّدل وهو ما يؤشر على دقة عملية ترجمة المقياس.

التحليلات الأولية لبيانات مقياس الكفاءة الاجتماعية المُعَّدل:

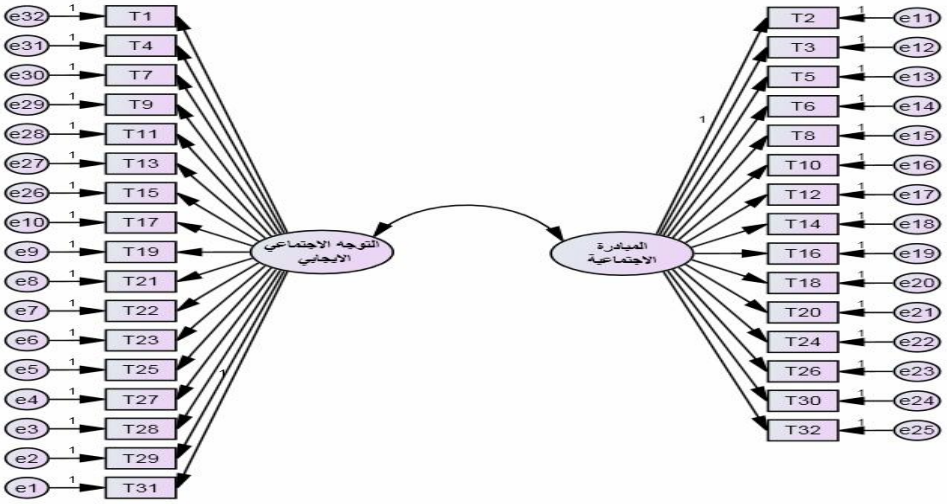
تم إجراء بعض التحليلات الأولية باستخدام برنامج SPSS 22.0 بهدف تعرف الخصائص الإحصائية لبيانات مقياس الكفاءة الاجتماعية المُعَّدل حيث إن اختيار الطريقة المناسبة لتحليل مصفوفة التباينات والتباينات المشتركة في التحليل العاملي التوكيدي يعتمد على الخصائص الإحصائية للبيانات (Byrne, 2010; Kline, 2010). وشملت التحليلات الأولية التحقق من وجود قيم متطرفة متعددة المتغيرات Multivariate Outliers ووجود حالة من التطابق والارتباط الشديد بين بيانات بعض العبارات Multicollinearity، وكذلك اعتدالية توزيع البيانات (Brown, 2006). وأظهرت نتائج التحليلات الأولية باستخدام اختبار 'Mahalanobis

Distance (MD) أنه لا توجد قيم متطرفة متعددة المتغيرات حيث كانت قيمة كا^{2} (χ^2) لكل قيمة من قيم البيانات غير دالة إحصائيا عند مستوي٠,٠٥ وكذلك بلغت قيمة عامل تضخم التباين Variance Inflation Factor وعامل التحمل Tolerance Factor ٢,٦٦ و٠,٨٥ على الترتيب وهو ما يشير إلى عدم وجود تطابق أو ارتباط شديد بين العبارات. فقيمة عامل تضخم التباين التي تقل عن ١٠ ,Hair 1995) Anderson, Tatham, & Black,) وقيمة عامل التحمل التي تزيد عن ٠,١٠ (Tabachnick & Fidell, 2001). تؤشر على عدم وجود تطابق أو ارتباط شديد بين المتغيرات. وتم حساب معاملي التواء وتفلطح البيانات على مستويين هما: أحادي المتغير Univariate ومتعدد المتغيرات Multivariate وذلك لأن بعض الباحثين يرون - ومنهم Kline (2010) وByrne (2010) - أن البيانات ربما تبدو غير ملتوية أو غير مفلطحة عند فحصها على مستوي أحادي المتغير بينما قد تُظهر نفس البيانات مؤشرات التواء أو تفلطح عند فحصها على المستوي متعدد المتغيرات. ويؤكد دي كارلو DeCarlo (1997) أن قيم الالتواء والتفلطح أحادي المتغير التي لا تزيد عن ٢ و٧ على الترتيب تشير إلى أن البيانات تحظي بتوزيع اعتدالي. وفي البيانات الحالية لم تتعد أي من قيم الالتواء والتفلطح أحادي المتغير القيم الإرشادية التي وصفها DeCarlo حيث إن قيم معاملي التواء والتفلطح للبيانات كانت قريبة من الصفر مما يؤكد على اعتدالية توزيع هذه البيانات. أما بالنسبة لقيم معاملي التواء وتفلطح البيانات في المستوي متعدد المتغيرات فيؤكد DeCarlo على أن هذه القيم ما لم تتعدي ٣ فهي تشير إلى اعتدالية توزيع البيانات ,Kline (2010). وقد بلغت قيمة مؤشر اعتدالية التوزيع متعدد المتغيرات ١,٨٥ مما يؤكد أن البيانات تحظي بتوزيع اعتدالي.

الصدق البنائي للنسخة العربية لمقياس الكفاءة الاجتماعية المُعَّدل:

تم في الدراسة الحالية استخدام التحليل العاملي التوكيدي من الدرجة الأولي بهدف فحص البنية العاملية للنسخة العربية لمقياس الكفاءة الاجتماعية المُعَّدل وذلك لأن البنية العاملية لهذا المقياس محددة ومعروفة مسبقا وذلك في ضـــوء الإطـــار النظري للكفاءة الاجتماعية الذي تبناه ريدل وهاجيكول وكذلك سبق وأن اســـتخدم ريدل وهاجيكول التحليل العاملي التوكيدي عند التحقق من البنية العاملية للمقياس

الأصلي (Rydell & Hagekull, 2012). ويوضح شكل (١) نموذج التحليل العاملي التوكيدي المفترض وهو يتكون من عاملين مرتبطين؛ العامل الأول هو التوجه الاجتماعي الإيجابي وتتشبع عليه ١٧ عبارة هي ١، ٤، ٧، ٩، ١١، ١٣، ١٥، ١٧، ١٩، ٢١، ٢٢، ٢٣، ٢٥، ٢٧، ٢٨، ٢٩، ٣١، أما العامل الثاني فهو المبادرة الاجتماعية وتتشبع عليه ١٥ عبارة هي ٢، ٣، ٥، ٦، ٨، ١٠، ١٢، ١٤، ١٦، ١٨، ٢٠، ٢٤، ٢٦، ٣٠، ٣٢.



شكل (١) نموذج افتراضي للبنية العاملية للنسخة العربية لمقياس الكفاءة الاجتماعية المُعَّدل

وفي ضوء نتائج التحليلات الأولية والتي أظهرت اعتدالية توزيع البيانات فقد تم إجراء التحليل العاملي التوكيدي باستخدام طريقة الاحتمالية القصوى Maximum Likelihood لتحليل مصفوفات التباينات والتباينات المشتركة للبيانات Variance-Covariance Matrix باستخدام برنامج AMOS 22.0 (Arbuckle, 2015). وقد تم الحكم على مدى مطابقة النموذج للبيانات فى ضوء عدد من المؤشرات الإحصائية المطلقة Absolute Fit Indices والنسبية (التزايدية) Incremental Fit Indices والتي تم اختيارها إما لأنها أقل اعتماداً على حجم العينة أو أقل تأثراً بتعقد النموذج أو كليهما معاً (Byrne, 2010). ويوضح الجدول (١) هذه المؤشرات وأسباب اختيارها وحدود قيمها المقبولة. والمؤشرات

الإحصائية المطلقة تعمل على التحقق من مطابقة النموذج المقترح لبيانات الدراسة دون مقارنته بأي نموذج أساسى، أما المؤشرات الإحصائية النسبية (التزايدية) فتعمل على التحقق من مطابقة النموذج المقترح لبيانات الدراسة من خلال مقارنته بالنموذج الأساسي وذلك عن طريق مقارنة قيمة مربع كاي (χ^2) للنموذج المقترح مع قيمة مربع كاي للنموذج الأساسى فى ضوء الفرق فى درجات الحرية بين النموذجين (Hancock & Mueller, 2006) . والنموذج الأساسي Baseline Model هو نموذج صفري Null Model لأن قيم التباينات والتباينات المشتركة بين المتغيرات الكامنة فى النموذج تساوي صفر، ومن ثم فالنموذج الأساسى أو الصفري يفترض أن الأبعاد المكونة للمتغير كامن غير مترابطة أو مستقلة عن بعضها البعض (Blunch, 2008; Kline,2010).

جدول (١)

المؤشرات الإحصائية المستخدمة في الحكم على مطابقة نموذج التحليل العاملي التوكيدي للنسخة العربية لمقياس الكفاءة الاجتماعية المُعَّدل للبيانات وأسباب اختيار كل مؤشر وحدود قيمه المقبولة.

مؤشرات المطابقة	سبب اختيار المؤشر [١]	حدود القيم المقبولة للمؤشر
أولا: مؤشرات المطابقة المطلقة Absolute Fit Indices		
١. اختبار مربع كاى النسبى Relative Chi-Squared (χ^2/df)	أقل اعتماداً على حجم العينة	بين ١ و٣
٢. الجذر التربيعى لمتوسط مربعات خطأ التقارب Root Mean Square Error of Approximation (RMSEA)	أقل اعتماداً على حجم العينة وأقل تأثراً بتعقد النموذج	RMSEA ≤ ٠,٠٦
٣. الجذر التربيعى لمتوسطات مربعات البواقى المعيارية Standardized Root Mean (SRMR) Square Residual	لا يتأثر بتعقد النموذج	SRMR ≤ ٠,٠٨
ثانيا: مؤشرات المطابقة النسبية (التزايدية) Incremental Fit Indices		
١. مؤشر المطابقة المقارن Comparative Fit Index (CFI)	لا يتأثر بحجم العينة	CFI ≤ ٠,٩٥
٢. مؤشر المطابقة غير المعيارية Nonnormed Fit Index (NNFI)	لا يتأثر بتعقد النموذج	NNFI ≤ ٠,٩٥

[١]ملحوظة: سبب الاختيار وحدود القيم المقبولة لكل مؤشر (انظر Byrne, 2010; Blunch 2008; Hancock & Mueller, 2006)

وأظهرت نتائج التحليل العاملي التوكيدي أن قيم المؤشرات الإحصائية لنموذج العاملين المرتبطين تقع في مدى القيم المقبولة

(χ^2 =502.40 , df =463 , χ^2/df = 1.08 , RMSEA =.04 , SRMR =.06, CFI = 0.97 , NNFI = 0.96)

والموضحة في جدول (١) حيث تدعم هذه النتائج البنية العاملية للنسخة العربية لمقياس الكفاءة الاجتماعية المُعّدل تتكون من عاملين مرتبطين. ويعرض جدول (٢) قيم تشعبات العبارات على عاملي التوجه الاجتماعي الإيجابي والمبادرة الاجتماعية حيث كان معامل الارتباط بين العاملين (٠,٦٢) وتشير القيمة الحرجة إلى إحصائية لها توزيع قيمة Z ومن ثم تكون دالة عند مستوي (٠,٠٥) إذا كانت قيمتها > ١,٩٦ وهي تساوي حاصل قسمة تشبع العبارة غير المعياري على العامل (بيتا غير المعيارية) مقسوما على الخطأ المعياري لتشبع العبارة على العامل (الخطأ المعياري لبيتا غير المعيارية)(Byrne, 2010, Kline, 2010) .

جدول (٢)

تشبع عبارات النسخة العربية لمقياس الكفاءة الاجتماعية المُعَّدل على عاملي التوجه الاجتماعي الإيجابي والمبادرة الاجتماعية والقيمة الحرجة لكل تشبع

العبارات	**قيم التشبعات**	**القيمة الحرجة**
العامل الأول: التوجه الاجتماعي الإيجابي		
١. يحاول أن يساعد أحد أقرانه عندما يبدو معتلا، أو ليس على ما يرام، أو أصيب بأذي	٠,٧٥	١٠,٢٥
٤. قادر على أن يتبين مشاعر طفل أخر إذا ما كان سعيدا أو غاضبا أو حزينا	٠,٦٨	٨,٤٥
٧. يجيد منع الصراعات	٠,٨٢	٧,٦٣
٩. قادر على الأخذ والعطاء في التفاعلات الاجتماعية	٠,٧٤	٩,٢٢
١١. غالبا ما يساعد أقرانه في التنظيف، البحث عن الأشياء المفقودة، أو إصلاح شيء مكسور	٠,٧٠	٤,٥٢
١٣. غالبا ما يكون قادرا على إيجاد الحلول أو التسويات عندما يتدخل في صراع	٠,٨٠	١٢,٣٣
١٥. يمتدح أقرانه (أفكارهم، ومظهرهم، وأفعالهم)	٠,٥٩	٦,٦٦
١٧. قادر على التعاطف مع أقرانه	٠,٧٧	٧,١٢
١٩. عادة ما يتشارك أو يُقرض متعلقاته	٠,٦٤	١١,٣٦
٢١. يحاول التدخل في مشاجرات وصراعات الأقران	٠,٥٧	٨,٤٢
٢٢. يدعو أقرانه الخجولين إلى المشاركة في اللعب	٠,٦٠	٧,٩٦
٢٣. يظهر كرما نحو أقرانه	٠,٧٣	٦,٥٣
٢٥. سهل التأثر ويشارك أقرانه سعادتهم ومزاجهم الجيد	٠,٨١	٧,٦٦
٢٧. يظهر الاستعداد للمساعدة الإيثار نحو الآخرين سواء كانوا أطفالا أم كبار	٠,٧٩	١٢,٤٥
٢٨. غالبا ما ينتقد الأقران*	٠,٥٩	٩,٤٠

العبارات	قيم التشبعات	القيمة الحرجة
٢٩. يقدم يد العون للكبار	٠,٦٥	٨,١١
٣١. يلعب ويتعاون على نحو جيد مع أقرانه	٠,٧٧	٥,٢٣
العامل الثاني: المبادرة الاجتماعية		
٢. غالبا ما يقترح أنشطة وألعاب للهو مع أقرانه	٠,٧٦	٨,٢٢
٣. انسحابي في مواجهة أقرانه *	٠,٦٤	٦,٤٣
٥. متردد في مواجهة أقرانه*	٠,٨٠	٧,٥٠
٦. في أغلب الأحيان يكون متفرجا أكثر منه مشاركا بينما يلعب الآخرون*	٠,٧٣	٩,١١
٨. خجول أو متردد مع الكبار الغرباء*	٠,٦٥	١٠,٣٣
١٠. يميل أن يهيمن عليه أقرانه *	٠,٦١	١٢,٣٦
١٢.عصبي وفظ مع الأطفال الغرباء*	٠,٥٨	٤,٥٢
١٤. غالبا ما يكون هو القائد في الألعاب أو الأنشطة	٠,٧٥	٦,٧٧
١٦. يمتنع عن اللعب مع غير أقرانه *	٠,٧٨	٨,٣٠
١٨. يعمل كموجه ومرشد لأقرانه أثناء اللعب	٠,٧١	٩,٢٤
٢٠. يعاني الارتباك في المواقف الجديدة*	٠,٦٣	٦,٦٦
٢٤. يعطي التعليمات والتوجهات لأقرانه	٠,٥٦	٧,٣٠
٢٦. يدخل في صراعات مع الأطفال الغرباء*	٠,٧١	٨,٤٥
٣٠. يتواصل بسهولة مع الأطفال الغرباء	٠,٦٩	٤,٣٣
٣٢. يحرص على أن يأخذ زمام المبادرة في كل تصرفاته	٠,٧٣	٦,٢٢

* عبارة سالبة

ثبات النسخة العربية لمقياس الكفاءة الاجتماعية المُعَّدل

تم حساب ثبات النسخة العربية لمقياس الكفاءة الاجتماعية المُعَّدل بطريقتين هما ثبات الاتساق الداخلي وثبات إعادة التطبيق. ففي طريقة ثبات الاتساق الداخلي بلغت قيم معامل الفا كرونباخ (٠,٨٩) لعامل التوجه الاجتماعي الإيجابي و(٠,٨٦) لعامل المبادرة الاجتماعية، و(٠,٩١) للمقياس ككل حيث تشير هذه القيم إلى أن عاملي التوجه الاجتماعي الإيجابي والمبادرة الاجتماعية وكذلك المقياس ككل يتمتعون بمستوي مرتفع من الاتساق الداخلي. وفي طريقة إعادة التطبيق، تم إعادة تطبيق النسخة العربية لمقياس الكفاءة الاجتماعية المُعَّدل مرة أخري على عينة فرعية من العينة الاستطلاعية (ن =٨٠) وبفاصل زمني قدره (٣) أسابيع عن التطبيق الأول حيث تم حساب معامل ارتباط بيرسون بين درجات الأطفال في التقييمين الأول والثاني فكانت قيم معامل الارتباط (معامل الاستقرار) (٠,٨٥) لعامل التوجه الاجتماعي الإيجابي و(٠,٨٨) لعامل المبادرة الاجتماعية، و(٠,٨٦) للمقياس ككل. وتشير هذه القيم إلى عاملي التوجه الاجتماعي الإيجابي والمبادرة

الاجتماعية والمقياس ككل يتمتعون بمستوي مرتفع من الثبات (الاستقرار) عبر الفترة الزمنية الفاصلة بين التطبيقين الأول والثاني.

العينة الأساسية:

شملت عينة الدراسة (٩٠) طفلا بالصف الثالث موزعين على فصلين في إحدى المدارس الحكومية للتعليم الأساسي (١-٤) في محافظة الباطنة جنوب في سلطنة عمان وهي من المحافظات ذات الكثافة الطلابية المرتفعة في الفصول المدرسية (وزارة التربية، ٢٠١٥). ففي الفصل الأول، قامت معلمة الفصل بتقييم الكفاءة الاجتماعية لدى (٤٥) طفلا (ذكور = ٢٧، إناث = ١٨، متوسط العمر = ٨,٦٠، انحراف معياري = ٠,٥٣) من أصل (٤٧) طفلا مسجلين في الفصل وذلك باستخدام النسخة العربية لمقياس الكفاءة الاجتماعية المُعَّدل، ومن ثم أعادت المعلمة المساعدة تقييم نفس الأطفال مرة أخرى على نحو مستقل مستخدمة نفس أداة القياس. وفي الفصل الثاني، قامت معلمة الفصل بتقييم الكفاءة الاجتماعية لدى (٤٥) طفلا (ذكور = ٢٤، إناث = ٢٣، متوسط العمر = ٨,٣٧، انحراف معياري = ٠,٤٢) من أصل (٤٦) طفلا مسجلين في الفصل باستخدام النسخة العربية لمقياس الكفاءة الاجتماعية المُعَّدل، ومن ثم أعادت المعلمة المساعدة تقييم نفس الأطفال مرة أخرى على نحو مستقل مستخدمة نفس أداة القياس. وأظهرت نتائج تحليل البيانات الديموغرافية أن متوسط عدد سنوات الخبرة التدريسية للمعلمة والمعلمة المساعدة في الفصل الأول كان (٧,٣) بانحراف معياري (٠,٢٢) وفي الفصل الثاني كان (٥,٧) بانحراف معياري (٠,١٧)، وأنه في كلا الفصلين كانت المعلمة حاصلة على درجة الماجستير بينما كانت المعلمة المساعدة حاصلة على درجة البكالوريوس، وأن اللغة العربية هي لغتهن الأم.

الإجـــــراءات:

تم التنسيق مع إداراتي مدرستين سيتم فيهما جمع بيانات العينة الاستطلاعية والعينة الأساسية في الدراسة الحالية وذلك بهدف عقد اجتماع منفصل في كل مدرسة شارك فيه جميع المعلمات والمعلمات المساعدات ممن يُدّرسْن بالصف الثالث الأساسي، حيث تم خلال الاجتماع شرح الهدف من الدراسة ودور كل من المعلمة والمعلمة المساعدة في عملية جمع البيانات، مع التأكيد على أن المشاركة ستكون طواعية، وأن البيانات التي سيتم جمعها ستبقي سرية ولن تستخدم في غير أغراض البحث العلمي.

فالنسبة للعينة الاستطلاعية فقد تمت عملية جمع البيانات على جولتين في إحدى المدارس الحكومية للتعليم الأساسي (١-٤) بمحافظة الداخلية في سلطنة

عمان. وكانت الجولة الأولي لجمع البيانات بعد قرابة شهر من بداية العام الدراسي ٢٠١٤/ ٢٠١٥ بهدف إتاحة الفرصة للمعلمة والأطفال في كل فصل أن يتعرفوا على بعضهم البعض وأن تتوثق العلاقات بينهم. وقد قامت معلمة الفصل بتقييم الكفاءة الاجتماعية لدى الأطفال باستخدام النسخة العربية لمقياس الكفاءة الاجتماعية المُعّدل أثناء حصص الدراسة العادية في المدرسة حيث استغرق تقييم كل طفل في المتوسط (١٠) دقائق واستمرت عملية جمع البيانات قرابة أسبوعين. وكانت الجولة الثانية لجمع البيانات بعد ثلاثة أسابيع من الجولة الأولى. وقد قامت معلمة الفصل بتقييم الكفاءة الاجتماعية لدى نفس الأطفال مستخدمة نفس الأداة من الجولة الأولي وذلك أثناء حصص الدراسة العادية في المدرسة حيث استغرق تقييم كل طفل في المتوسط (٨) دقائق واستمرت عملية جمع البيانات قرابة (١٠) أيام. وقد تم التنبيه في كل جولة من جولتي جمع البيانات على المعلمتين بعدم مناقشة تقييمات الطلاب مع بعضهما البعض حفاظا على مبدأ استقلالية التقييم.

والنسبة للعينة الأساسية فقد تمت عملية جمع البيانات على جولتين في فصلين بإحدى المدارس الحكومية للتعليم الأساسي (١-٤) بمحافظة الباطنة جنوب في سلطنة عمان. وكانت الجولة الأولي لجمع البيانات قرب نهاية الفصل الأول للعام الدراسي ٢٠١٤/ ٢٠١٥. وفي كل فصل مدرسي قامت معلمة الفصل بتقييم الكفاءة الاجتماعية لدى الأطفال باستخدام النسخة العربية لمقياس الكفاءة الاجتماعية المُعّدل وذلك أثناء حصص الدراسة العادية في المدرسة ومن ثم قامت المعلمة المساعدة بإعادة تقييم نفس الأطفال مرة أخرى على نحو مستقل مستخدمة نفس أداة القياس. وقد استغرق تقييم كل طفل من جانب معلمة الفصل أو المعلمة المساعدة في المتوسط (١٠) دقائق واستمرت عملية جمع البيانات قرابة أسبوعين. وكانت الجولة الثانية لجمع البيانات بعد شهر من نهاية الجولة الأولى- أي في بداية الشهر الثاني. وفي كل فصل مدرسي قامت معلمة الفصل بتقييم الكفاءة الاجتماعية لدى نفس الأطفال مستخدمة نفس الأداة من الجولة الأولي وذلك أثناء حصص الدراسة العادية في المدرسة ومن ثم قامت المعلمة المساعدة بإعادة تقييم نفس الأطفال مرة أخرى على نحو مستقل مستخدمة نفس الأداة من الجولة الأولي. وقد استغرق تقييم كل طفل من جانب معلمة الفصل أو المعلمة المساعدة في المتوسط (٧) دقائق واستمرت عملية جمع البيانات قرابة (١٠) أيام. وقد تم التنبيه في كل جولة من جولتي جمع البيانات على المعلمتين والمعلمتين المساعدتين بعدم مناقشة تقييمات الطلاب مع بعضهن البعض حفاظا على مبدأ استقلالية التقييم. وقد

تم في جميع جولات جميع البيانات في العينتين الاستطلاعية والأساسية جمع بيانات ديموغرافية خاصة بالمعلمة والمعلمة المساعدة منها الاسم، ونوع الوظيفة، وعدد سنوات الخبرة التدريسية، وأعلى مؤهل علمي حصلت عليه، واللغة الأم، وكذلك بيانات خاصة بالطفل منها الاسم والنوع الاجتماعي واللغة الأم.

تصميم تحليلات نظرية إمكانية التعميم في الدراسة الحالية:

تم في الدراسة الحالية إجراء دراسات إمكانية التعميم G-Studies بهدف تقدير حجم تباين الخطأ الذي يمكن أن يفسره كل من أوجه المُقيّمات (المعلمات والمعلمات المساعدات)، وعبارات المقياس، ووقت التطبيق في التباين الكلي لدرجات تقييمات معلمتين ومعلمتين مساعدتين للكفاءة الاجتماعية لدى الأطفال باستخدام النسخة العربية لمقياس الكفاءة الاجتماعية المُعَّدل. ويري كل من سميث (Smith, 1981) وبرينن (Brennan, 2001) أنه كلما زاد عدد الأوجه التي يمكنها تفسير تباين الخطأ كلما زادت إمكانية إجراء تصميمات معقدة في دراسات إمكانية التعميم، ومع ذلك فإن إجراء عدة دراسات منفردة لإمكانية التعميم يظل هو الخيار الأفضل والأسهل في مثل هذه الحالة إذ أن ذلك يمكننا من تقدير مكونات التباين الملاحظ على نحو دقيق ومستقر. وفي ضوء هذا التوجيه، فقد تم تحليل بيانات الدراسة الحالية من خلال تصميمات ثلاثية الأوجه تامة التقاطع Fully Crossed Three-Facet Designs (Brennan, 20001). وتوصف هذه التصميمات بأنها ثلاثية الأوجه لأنها تضم ثلاثة أوجه محتملة لتباين الخطأ هي المُقيّمات (المعلمات أو المعلمات المساعدات)، وعبارات المقياس، ووقت التطبيق، وهي تامة التقاطع لأنه يتم خلالها تقيم كل طفل من خلال كل معلمة ومعلمة مساعدة في كل جولة من جولتي جمع البيانات وعلى كل عبارة من عبارات النسخة العربية لمقياس الكفاءة الاجتماعية المُعَّدل. والتصميم الذي تم استخدامه في الدراسة الحالية يضم الفرد Person (p) (الطفل)، والمُقيّمة Rater (r) (المعلمة والمعلمة المساعدة في كل فصل)، ووقت التطبيق Occasion (O)، والعبارة Item (i)، ويشار إلىه اختصارا بـــ p×r×o×i مع مراعاة أن الفرد ليس مصدر (وجه) محتمل للخطأ، وذلك لأن الفرد هنا يشير إلى التباين الحقيقي الذي يمثل الفروق الفردية الفعلية بين الأطفال في الكفاءة الاجتماعية ويشار إليه بـــ Measurement Object في اطار نظرية إمكانية التعميم (Shavelson & Webb, 1991).

وقد تم تحليل البيانات باستخدام برنامج GENOVA (Crick & Brennan, 1982) حيث يقوم البرنامج بحساب حجم التباين الحقيقي (p)، وحجم تباين الخطأ الذي يفسره في التباين الكلي أوجه المُقيّمة (r) Rater، ووقت التطبيق Occasion

(O)، والعبارة (Items (i، وإمكانية التفاعل بين هذه أوجه الخطأ وبعضها البعض، وإمكانية التفاعل بين أوجه الخطأ والتباين الحقيقي وذلك كله من خلال دراسات إمكانية التعميم G-Studies، كما يجري البرنامج تحليل دراسات القرارات D-Studies لمعرفة أفضل الشروط (مثلا عدد المُقيّمات، عدد العبارات، عدد مرات التطبيق) التي يمكن من خلالها تحقيق أعلى مستوى ثبات (بيانات قابلة للتعميم) للنسخة العربية لمقياس الكفاءة الاجتماعية المُعَّدل. وقد تم إجراء التصميم p×r×o×i بصورة منفصلة لكل فصل من فصلي عينة الدراسة وبصورة مستقلة لكل من عامل التوجه الاجتماعي الإيجابي وعامل المبادرة الاجتماعية.

النتائـــــــج:

السؤال الأول:

ينص السؤال الأول على: ما حجم تباين الخطأ الذي يفسره كل من أوجه المُقيّمات (المعلمات والمعلمات المساعدات)، وعبارات المقياس، ووقت التطبيق في التباين الكلي لدرجات الأطفال على النسخة العربية لمقياس الكفاءة الاجتماعية المُعَّدل؟

وللإجابة عن السؤال الأول تم إجراء دراسات إمكانية التعميم G-Studies على عاملي التوجه الاجتماعي الإيجابي والمبادرة الاجتماعية في الفصلين الأول والثاني كل على حدة.

(١) **عامل التوجه الاجتماعي الإيجابي**. يتضح من نتائج تحليلات دراسات إمكانية التعميم لعامل التوجه الاجتماعي الإيجابي في جدول (٢) أن نسبة التباين الحقيقي (التباين المنتظم) الذي يصف الفروق الحقيقية بين الأطفال في التوجه الاجتماعي الإيجابي كانت (٤٦,٣%) من التباين الكلي في الفصل الأول ٤٢,٢(%) من التباين الكلي في الفصل الثاني. وتشير هذه النسب إلى أنه بالأخذ في الاعتبار متوسط المُقيّمات والعبارات ووقت التطبيق فإن هناك فروقا حقيقية بين الأطفال في التوجه الاجتماعي الإيجابي في كل فصل مدرسي (Brennan, 2001).

الفصل الأول: يوضح جدول (٢) أن المُقيّمات لم يفسرن أي نسبة (٠%) من التباين الكلي مما يشير إلى ثبات تقييمات المُقيّمات لسلوك التوجه الاجتماعي الإيجابي لدى الأطفال، في حين فسر وقت التطبيق ٩.% من التباين الكلي وهي نسبة ضئيلة تشير إلى أن سلوك التوجه الاجتماعي الإيجابي لدى الأطفال لم يختلف عبر مرات التطبيق، وفسرت العبارات ١,٦% من التباين الكلي وهي نسبة

ضئيلة أيضا تشير إلى أن العبارات لم تتباين في مستوى صعوبتها (أي لم تختلف العبارات في تأثيرها على المُقيّمات). ويوضح جدول ٢ أن هناك عدة تفاعلات مهمة منها التفاعل بين الفرد ووقت التطبيق والذي فسر ٧,٦% من التباين الكلي وهو ما يشير إلى تباين الوضع النسبي (الوضع الرتبي) للأطفال في عامل التوجه الاجتماعي الإيجابي عبر مرات التطبيق، والتفاعل بين الفرد والعبارات والذي فسر ١١,١% من التباين الكلي وهو ما يشير إلى تباين الوضع النسبي (الوضع الرتبي) للأطفال في عامل التوجه الاجتماعي الإيجابي عبر العبارات. أما التفاعل الرباعي بين الفرد والمُقيّمات والعبارات ووقت التطبيق فيمثل التباين غير المفسر Undifferentiated Error (Residual) وقد فسر ١٤,٣ % من التباين الكلي وهو ما يعني تباين الوضع النسبي (الوضع الرتبي) للأطفال في سلوك التوجه الاجتماعي الإيجابي عبر المُقيّمات والعبارات ووقت التطبيق و/أو أي أوجه خطأ أخرى لم يتم تضمينها في دراسة إمكانية التعميم لبيانات الفصل الأول.

الفصل الثاني: يتضح من جدول (٢) أن المُقيّمات فسرن ٠,٤% من التباين الكلي وهي نسبة ضئيلة للغاية تشير إلى ثبات تقييمات المُقيّمات لسلوك التوجه الاجتماعي الإيجابي لدى الأطفال، في حين فسر وقت التطبيق ٠,٦% من التباين الكلي وهي نسبة ضئيلة للغاية تشير إلى أن سلوك التوجه الاجتماعي الإيجابي لدى الأطفال لم يختلف عبر مرات التطبيق، وفسرت العبارات ٨,٥% من التباين الكلي وهي نسبة تشير إلى أن العبارات تباينت بقدر ما في مستوى صعوبتها (أي أن العبارات قد اختلفت في تأثيرها على المُقيّمات). ويلاحظ من جدول (٢) أن هناك عدة تفاعلات مهمة منها التفاعل بين الفرد والعبارات والذي فسر ٧,٢% من التباين الكلي وهو ما يشير إلى تباين الوضع النسبي (الوضع الرتبي) للأطفال في التوجه الاجتماعي الإيجابي عبر العبارات، والتفاعل بين المُقيّمات والعبارات والذي فسر ١٢,٢ % من التباين الكلي وهو ما يشير إلى تباين تقييمات المُقيّمات عبر العبارات، وتفاعل بين الفرد والمُقيمات والعبارات والذي فسر ١٤,٢% من التباين الكلي وهو ما يشير إلى تباين تقييمات المُقيمات لسلوك الأطفال على العبارات. أما التفاعل الرباعي بين الفرد والمُقيمات ووقت التطبيق والعبارات فيمثل التباين غير المفسر Undifferentiated Error (Residual) وفقد فسر ١٢,٢% من التباين الكلي وهو ما يعني تباين الوضع النسبي (الوضع الرتبي) للأطفال في سلوك التوجه الاجتماعي الإيجابي عبر المُقيّمات والعبارات ووقت التطبيق و/ أو أي أوجه خطأ أخرى لم يتم تضمينها في دراسة إمكانية التعميم لبيانات الفصل الثاني.

جدول(٣)

نتائج تحليلات دراسات إمكانية التعميم G- Studies

لعامل التوجه الاجتماعي الإيجابي

مصادر التباين	الفصل	درجات الحرية	متوسط المربعات	مكونات التباين المقدرة	نسبة التباين الكلي
الفرد	الأول	٤٥	١٤,٤٤	٠,٠٩٤٨	%٤٦,٣
	الثاني	٤٥	٣٥,٥٥	٠,٣١٥٥	%٤٢,٢
المُقيّمات	الأول	١	٠,٢٢	٠	٠
	الثاني	١	٠,٥١٠	٠,٠٠٢٥	%٠,٤
وقت التطبيق	الأول	١	٢,٠٢	٠,٠٠٠٧	%٠,٩
	الثاني	١	٥,١٠	٠,٠٠٦٦	%٠,٦
العبارة	الأول	١٧	٠,٧٢	٠,٠٤٥	%١,٦
	الثاني	١٧	٦,٤٩	٠,٠٩٠٠	%٨,٥
الفرد × المُقيّمات	الأول	٤٥	٠,٥٦	٠,٠٠٦٠	%٣,١
	الثاني	٤٥	٠,٨٠	٠,٠٠٣٣	%٠,٤
الفرد × وقت التطبيق	الأول	٤٥	١,٧٠	٠,٠٢٩٢	%٧,٦
	الثاني	٤٥	٠,٧٣	٠,٠٠٧٢	%٠,٦
الفرد × العبارة	الأول	٨٥٥	٠,٣٢	٠,٠٤١٠	%١١,١
	الثاني	٨١٠	٠,٧٥	٠,٠٧٢٣	%٧,٢
المُقيّمات × وقت التطبيق	الأول	١	٠,١٧	٠	٠
	الثاني	١	١,٢٩	٠,٠٠٣٩	%٠,١
المُقيّمات × العبارة	الأول	١٧	٠,٣٣	٠,٠٠٦٥	%٢,١
	الثاني	١٧	٢,٨٢	٠,١١١٩	%١٢,٢
وقت التطبيق × العبارة	الأول	١٧	٠,٣٣	٠,٠٠٥٢	%٢,٥
	الثاني	١٧	٠,٢٥	٠	٠
الفرد × المُقيّمات× وقت التطبيق	الأول	٤٥	٠,٣٦	٠,٠٠٦٦	%٢,٧
	الثاني	٤٥	٠,٣٦	٠,٠٠٧٥	%٠,٥
الفرد × المُقيّمات× العبارة	الأول	٨٥٥	٠,٢٠	٠,٠١٦٢	%٤,٢
	الثاني	٨١٠	٠,٥٥	٠,١٤٠٩	%١٤,٢
الفرد × وقت التطبيق× العبارة	الأول	٨٥٥	٠,١٤	٠,٠١٢٣	%٣,٥
	الثاني	٨١٠	٠,٢٠	٠,٠١٥٢	%٠,٢
المُقيّمات ×وقت التطبيق× العبارة	الأول	١٧	٠,١٦	٠,٠٠٠٦	%٠,١
	الثاني	١٧	٠,٢٨	٠,٠٠٨٥	%٠,٧
الفرد × المُقيّمات × وقت التطبيق× العبارة	الأول	٨٥٥	٠,١٦	٠,١٤٤٢	%١٤,٣
	الثاني	٨١٠	٠,١٨	٠,١٧٩٥	%١٢,٢

(٢) عامل المبادرة الاجتماعية. يتضح من نتائج تحليلات إمكانية التعميم لعامل المبادرة الاجتماعية في جدول (٣) أن نسبة التباين الحقيقي (التباين المنتظم) الذي يصف الفروق الحقيقية بين الأطفال في المبادرة الاجتماعية كانت ٤١,٧٢ % من التباين الكلي في الفصل الأول ٤٠,٥٢% من التباين الكلي في

الفصل الثاني. وتشير هذه النسب إلى أنه بالأخذ في الاعتبار متوسط المُقيّمين والعبارات ووقت التطبيق فإن هناك فروقا حقيقية بين الأطفال في المبادرة الاجتماعية في كل فصل مدرسي (Brennan, 2001).

الفصل الأول: يتضح من جدول (٣) أن المُقيّمات فسروا نسبة ٠,٢% من التباين الكلي وهي نسبة ضئيلة للغاية تشير إلى ثبات تقييمات المُقيّمات لسلوك المبادرة الاجتماعية لدى الأطفال، في حين لم يفسر وقت التطبيق أي نسبة (٠%) في التباين الكلي وهو ما يعني أن سلوك المبادرة الاجتماعية لدى الأطفال لم يختلف عبر مرات التطبيق، وفسرت العبارات ٢,٢% من التباين الكلي وهي نسبة ضئيلة إلى حد ما وتشير إلى أن العبارات ربما تكون تباينت في مستوى صعوبتها (أي أن العبارات قد اختلفت في تأثيرها على المُقيّمات). وكذلك يلاحظ من جدول ٣ أن هناك عدة تفاعلات مهمة منها التفاعل بين الفرد والمُقيّمات والذي فسر ٨,٥% من التباين الكلي وهو ما يشير إلى تباين تقييمات المُقيّمات لسلوك المبادرة الاجتماعية لدى الأطفال، والتفاعل بين الفرد والعبارات والذي فسر ١٢,٨ % من التباين الكلي وهو ما يشير إلى تباين الوضع النسبي (الوضع الرتبي) للأطفال في سلوك المبادرة الاجتماعية عبر العبارات، والتفاعل بين الفرد والمُقيّمات ووقت التطبيق والذي فسر ٥,٧% من التباين الكلي وهو ما يشير إلى تباين تقييمات المُقيّمات لسلوك المبادرة الاجتماعية لدى الأطفال عبر مرات التطبيق. أما التفاعل الرباعي بين الفرد والمُقيّمات والعبارات ووقت التطبيق فيمثل التباين غير المفسر (Residual) Undifferentiated Error وقد فسر ١٣,٦% من التباين الكلي وهو ما يعني تباين الوضع النسبي (الوضع الرتبي) للأطفال في سلوك المبادرة الاجتماعية عبر المُقيّمات والعبارات ووقت التطبيق و/أو أي أوجه خطأ أخرى لم يتم تضمينها في دراسة إمكانية التعميم لبيانات الفصل الأول.

الفصل الثاني: يتضح من جدول (٣) أن المُقيّمات فسروا ٠,٧% من التباين الكلي وهي نسبة ضئيلة للغاية تشير إلى ثبات تقييمات المُقيّمات لسلوك المبادرة الاجتماعية لدى الأطفال، في حين فسر وقت التطبيق ١,٢% من التباين الكلي وهي أيضا نسبة ضئيلة جدا تشير إلى أن سلوك المبادرة الاجتماعية لدى الأطفال لم يختلف عبر مرات التطبيق، وفسرت العبارات ٥,٨% من التباين الكلي وهي نسبة تشير إلى العبارات تباينت في مستوى صعوبتها (أي أن العبارات قد اختلفت في تأثيرها على المُقيّمات). وكذلك يلاحظ من جدول ٣ أن هناك عدة تفاعلات مهمة منها التفاعل بين الفرد والعبارات والذي فسر ١٤,٤% من التباين الكلي وهو ما يشير إلى تباين الوضع النسبي (الوضع الرتبي) للأطفال في سلوك المبادرة

الاجتماعية عبر العبارات، والتفاعل بين الفرد والمُقيّمات والعبارات والذي فسر ١١,٩% من التباين الكلي وهو ما يشير إلى تباين تقييمات المقيمين لسلوك المبادرة الاجتماعية لدى الأطفال عبر العبارات، والتفاعل بين الفرد والمُقيّمات ووقت التطبيق والذي فسر ٥,٣% من التباين الكلي وهو ما يشير إلى تباين تقييمات المُقيّمات لسلوك المبادرة الاجتماعية لدى الأطفال عبر مرات التطبيق. أما التفاعل الرباعي بين الفرد والمُقيّمات ووقت التطبيق والعبارات فيمثل التباين غير المفسر (Residual) Undifferentiated Error وقد فسر ١٠,٢% من التباين الكلي وهو ما يعني تباين الوضع النسبي (الوضع الرتبي) للأطفال في سلوك المبادرة الاجتماعية عبر المُقيّمات والعبارات ووقت التطبيق و/ أو أي أوجه خطأ أخرى لم يتم تضمينها في دراسة إمكانية التعميم لبيانات الفصل الثاني.

جدول (٤)

نتائج تحليلات دراسات إمكانية التعميم G- Studies لعامل المبادرة الاجتماعية

مصادر التباين	الفصل	درجات الحرية	متوسط المربعات	مكونات التباين المقدرة	نسبة التباين الكلي
الفرد	الأول	٤٥	٨,٤٢	٠,٥٩٦٠	%٤١,٧٢
	الثاني	٤٥	١٣,٤٧	٠,١٣٢٠	%٤٠,٥٢
المُقيّمات	الأول	١	٥,١٠	٠,٠٠٣٥	%٠,٢
	الثاني	١	٢,١٢	٠,٠٠٣٩	%٠,٧
وقت التطبيق	الأول	١	٠,١٢	٠	%٠
	الثاني	١	١,٤٠	٠,٠٠١٥	%١,٢
العبارة	الأول	١٥	٠,٩٥	٠,٠١٠٩	%٢,٢
	الثاني	١٥	٠,٧٥	٠,٠٣٥٤	%٥,٨
الفرد × المُقيّمات	الأول	٤٥	٢,١٠	٠,٠٣٢٠	%٨,٥
	الثاني	٤٥	٠,٨٤	٠	٠
الفرد × وقت التطبيق	الأول	٤٥	٠,٦٥	٠,٠٠١٠	%٢,٨٨
	الثاني	٤٥	٠,٧٢	٠,٠٠٠٧	%٣,١
الفرد × العبارة	الأول	٦٧٥	٠,٣٠	٠,٠٤٣٠	%١٢,٨
	الثاني	٦٧٥	٠,٣٤	٠,١٠٥٢	%١٤,٤
المُقيّمات × وقت التطبيق	الأول	١	١,٩٢	٠,٠٠٤٢	%٣,٥
	الثاني	١	٠,٤٠	٠	٠
المُقيّمات × العبارة	الأول	١٥	٠,٣٧	٠,٠٠٧٢	%٢,٨
	الثاني	١٥	٠,٤٢	٠,٠٠٨٢	%٢,٨٦
وقت التطبيق × العبارة	الأول	١٥	٠,٢٠	٠	%٠,٢
	الثاني	١٥	٠,٢٨	٠,٠٠٤٦	%٠,٩
الفرد × المُقيّمات× وقت التطبيق	الأول	٤٥	٠,٥٢	٠,٠١٥٤	%٥,٧
	الثاني	٤٥	٠,٧٢	٠,٠٢٤٣	%٥,٣
الفرد × المُقيّمات× العبارة	الأول	٦٧٥	٠,٢٢	٠,٠٠١٥	%٣,٣
	الثاني	٦٧٥	٠,٢٥	٠,٠٥٤٢	%١١,٩

مصادر التباين	الفصل	درجات الحرية	متوسط المربعات	مكونات التباين المقدرة	نسبة التباين الكلي
الفرد × وقت التطبيق× العبارة	الأول	٦٧٥	٠,١٦	٠	٠%
	الثاني	٦٧٥	٠,١٢	٠,٠٠٠٩	٣,١٢%
المُقيّمات ×وقت التطبيق× العبارة	الأول	١٥	٠,١٦	٠,٠٠٤٢	٢,٦%
	الثاني	١٥	٠,١٢	٠	٠
الفرد × المُقيّمات× وقت التطبيق× العبارة	الأول	٦٧٥	٠,١٢	٠,١٤٣٣	١٣,٦%
	الثاني	٦٧٥	٠,١٧	٠,١٤٧٠	١٠,٢%

السؤال الثاني:

ينص السؤال الثاني على: ما الشروط تحقيق النسخة العربية لمقياس الكفاءة الاجتماعية المُعَّدل لأفضل مستويات الثبات؟

وللإجابة عن السؤال الأول تم إجراء دراسات القرارات D-Studies على عاملي التوجه الاجتماعي الإيجابي والمبادرة الاجتماعية لكل على حدة في الفصلين الأول والثاني.

يلخص جدول (٤) قيم معاملات إمكانية التعميم لعينة الدراسة وذلك لعاملي التوجه الاجتماعي الإيجابي والمبادرة الاجتماعية في الفصلين الأول والثاني حيث يتضح أنه في كل الحالات فإن استخدام اثنتين من المُقيّمات وتطبيق النسخة العربية لمقياس الكفاءة الاجتماعية المُعَّدل مرتين يُعد إجراءً كافيا للحصول على أعلى قيم لمعاملات إمكانية التعميم (القيم مكتوبة بخط ثقيل في جدول ٤) وفي هذه الحالة فقد بلغ معامل إمكانية التعميم ٠,٨٤٦ بالنسبة لعامل التوجه الاجتماعي الإيجابي في الفصل الأول و٠,٩٥٤ في الفصل الثاني، أما معامل إمكانية التعميم لعامل المبادرة الاجتماعية فكان ٠,٧٨٨ للفصل الأول و٠,٩٢٦ للفصل الثاني.

ويلاحظ من جدول (٤) أنه في كل الحالات أن زيادة عدد المُقيّمات إلى ثلاث مُقيّمات وزيادة عدد مرات التطبيق إلى ثلاث مرات يزيد من قيم معامل إمكانية التعميم ولكن هذه الزيادة تبدو ضئيلة وغير مؤثرة، وأن قيم معاملات إمكانية التعميم تكون أقل ما يمكن عند استخدام مُقيّمه واحدة وتطبيق المقياس لمرة واحدة. ولو أراد الباحث الاختيار ما بين زيادة عدد المُقيّمات أو عدد مرات التطبيق بالنسبة لعاملي التوجه الاجتماعي الإيجابي أو المبادرة الاجتماعية سواء على مستوى الفصل الأول أو الفصل فإن النتائج في جدول (٤) تشير إلى أنه من المناسب زيادة عدد المُقيّمات إلى مُقيّمتين اثنتين وأن يتم تطبيق المقياس مرة واحدة فقط.

جدول (٥) نتائج دراسات القرارات D- Studies
للنسخة العربية لمقياس الكفاءة الاجتماعية المُعدل

الفصل المدرسي	العامل	الفرد (P)	المُقيم (R)	وقت التطبيق (O)	العبارات (I)	معامل التعميم (G C)
١	التوجه الاجتماعي الإيجابي	٤٥	١	١	١٧	٠,٦٧٠
١	**التوجه الاجتماعي الإيجابي**	**٤٥**	**٢**	**٢**	**١٧**	٠,٨٤٦
١	التوجه الاجتماعي الإيجابي	٤٥	٣	٣	١٧	**٠,٨٥٣**
١	التوجه الاجتماعي الإيجابي	٤٥	١	٢	١٧	**٠,٧٠٧**
١	التوجه الاجتماعي الإيجابي	٤٥	٢	١	١٧	**٠,٧٦٨**
٢	التوجه الاجتماعي الإيجابي	٤٥	١	١	١٧	**٠,٦٦٤**
٢	**التوجه الاجتماعي الإيجابي**	**٤٥**	**٢**	**٢**	**١٧**	٠,٩٥٤
٢	التوجه الاجتماعي الإيجابي	٤٥	٣	٣	١٧	**٠,٩٦٨**
٢	التوجه الاجتماعي الإيجابي	٤٥	١	٢	١٧	**٠,٧٩٣**
٢	التوجه الاجتماعي الإيجابي	٤٥	٢	١	١٧	**٠,٨٧٤**
١	المبادرة الاجتماعية	٤٥	١	١	١٥	**٠,٥١٤**
١	**المبادرة الاجتماعية**	**٤٥**	**٢**	**٢**	**١٥**	٠,٧٨٨
١	المبادرة الاجتماعية	٤٥	٣	٣	١٥	**٠,٧٩٨**
١	المبادرة الاجتماعية	٤٥	١	٢	١٥	**٠,٥٧٠**
١	المبادرة الاجتماعية	٤٥	٢	١	١٥	**٠,٧٢٣**
٢	المبادرة الاجتماعية	٤٥	١	١	١٥	**٠,٧٧٢**
٢	**المبادرة الاجتماعية**	**٤٥**	**٢**	**٢**	**١٥**	٠,٩٢٦
٢	المبادرة الاجتماعية	٤٥	٣	٣	١٥	**٠,٩٥٧**
٢	المبادرة الاجتماعية	٤٥	١	٢	١٥	**٠,٧٩٥**
٢	المبادرة الاجتماعية	٤٥	٢	١	١٥	**٠,٨٧٨**

المناقشـــــــــة:

يتمثل الهدف الأول للدراسة الحالية في الكشف عن حجم تباين الخطأ الذي يفسره كل من أوجه المُقيّمات (معلمات الفصل والمعلمات المساعدات)، وعبارات المقياس، ووقت التطبيق في التباين الكلي لدرجات الأطفال على النسخة العربية لمقياس الكفاءة الاجتماعية المُعَّدل حيث أظهرت نتائج دراسات إمكانية التعميم أن هذه الأوجه منفردة لم تفسر سوى نسبة ضئيلة من التباين الكلي في درجات الأطفال على المقياس. وتتفق هذه النتيجة مع نتائج دراسات أخرى أظهرت أن أوجه المُقيم ووقت التطبيق والعبارات لم تفسر نسبة كبيرة من التباين الكلي (Christ, 2012; Luka & Cheung, 2013; Lueng et al., 2010; Ribar, 2005) ولكن هذه النتيجة تختلف مع نتائج بعض الدراسات التي أظهرت أن أوجه المُقيم ووقت التطبيق والعبارات فسرت نسبة ملحوظة من التباين (Carminla, 2010; Broger, et al., 2012). وتشير هذه النتيجة إلى أن عبارات المقياس لم تختلف في

مستوى صعوبتها بالنسبة للمُقيّمات وهو ما يدل على جودة صياغة هذه العبارات ووضوحها، وسلاستها، ودقة تعبيرها عن المظاهر السلوكية الدالة على سمة الكفاءة الاجتماعية لدى الأطفال، وأن هذه العبارات يمكن أن تمثل نطاق أوسع وأشمل من العبارات (Brennan, 2000)، وكذلك تشير هذه النتيجة إلى أن المُقيّمات يتمتعن ببعض المزايا التي دعّمت ثبات تقييماتهن للكفاءة الاجتماعية عند الأطفال ومنها: (١) التنبيه على المُقيّمات بعدم الدخول في مناقشات حول تقييم الكفاءة الاجتماعية لدى الأطفال إعمالا لمبدأ استقلالية التقييم حيث قامت المعلمة والمعلمة المساعدة بتقييم الأطفال بصورة مستقلة عن بعضهن البعض، (٢) أمضت المُقيّمات وقتا كافيا مع الأطفال مما خلق نوع من الألفة بين الطرفين ومن ثم لم يلجأ الأطفال إلى تغيير سلوكياتهم في مواقف التقييم وهو أمر حرص الباحث في الدراسة الحالية على توفيره من خلال جمع البيانات قرب منتصف الفصل الأول من العام الدراسي ٢٠١٥/٢٠١٤ حتى تتاح الفرصة للمُقيّمات والأطفال أن يتعرفوا على بعضهم البعض وأن تتوثق العلاقات بينهم، (٣) أن المُقيّمات يتمتعن بالقدر الكاف من المؤهلات العلمية والتدريب والخبرة التدريسية بما ساعدهن على رصد وملاحظة سلوكيات الطفل التي تعكس كفاءة اجتماعية حقيقية وتمييزها عن السلوكيات التي تعكس خصائص عابرة، فالمعلمات في الفصلين الأول والثاني حصلتا على درجة الماجستير في التربية، في حين أن المعلمتين المساعدتين في نفس الفصلين حصلتا على درجة البكالوريوس في التربية، وكان متوسط عدد سنوات الخبرة التدريسية للمعلمة والمعلمة المساعدة في الفصل الأول ٧,٣ وفي الفصل الثاني كان ٥,٧. وفيما يتعلق بتأثير مرات التطبيق فإن النتيجة الحالية تشير إلى أن الفترة الزمنية الفاصلة بين التقييمين الأول والثاني وقدرها ثلاثة شهور كان لها تأثير ضئيلا على تغيير سمة الكفاءة الاجتماعية لدى الأطفال، وهو ما يشير إلى الثبات النسبي لمستوى لهذه السمة عبر هذه الفترة الزمنية وربما يرجع ذلك إلى عدم تعرض الأطفال طيلة هذه الفترة إلى مواقف يمكنها إحداث تغييرات جذرية في مهاراتهم وسلوكياتهم الاجتماعية، كما من الممكن أن يرجع ذلك إلى طبيعة سمة الكفاءة الاجتماعية كونها سمة ثابتة نسبيا (Semrud-Clikeman, 2007).

وبالمثل أظهرت نتائج دراسات التعميم أن التفاعل بين أوجه الخطأ وبعضها البعض أو التفاعل بين أوجه الخطأ وبين مصدر التباين الحقيقي (الفرد) لم يفسر سوى نسب ضئيلة من التباين الكلي، بل إن بعضا من هذه التفاعلات لم تفسر أي

نسبة في التباين الكلي على الإطلاق. وتتفق هذه النتيجة مع نتائج دراسة (Christ 2010) حيث لم يفسر التفاعل بين المُقيّمتين (المعلمتين) والعبارات أكثر من ٢% من التباين الكلي، كما تتفق هذه النتيجة مع نتائج دراسة بروجر وآخرون (Broger et al., 2012) والتي أظهرت أن التفاعل بين المُقيّمتين (المعلمتين) ووقت التطبيق كان ضعيفا بحيث إنه فسر فقط ١,٧% من التباين الكلي. ولكن هذه النتيجة تختلف مع نتائج دراسات أخري والتي رصدت أن التفاعلات بين بعض من هذه الأوجه تفسر نسب كبيرة من التباين الكلي ومنها دراسة (Carminla, 2010) والتي أظهرت نتائجها أن التفاعل بين المُقيّمة (المعلمة) ووقت التطبيق فسر ٢٠,٢% من التباين الكلي، ودراسة لوكا وشينج (Luka & Cheung, 2013) والتي أظهرت نتائجها أن التفاعل بين العبارات والمُقيّميين (المعلمين) فسر ٢٧% من التباين الكلي.

وقد كان لافتا للانتباه أن الفرد باعتباره مصدر التباين الحقيقي حاضرا وبقوة كطرف أصيل في كل هذه التفاعلات على اعتبار أنه يمثل الفروق الحقيقة في سمة الكفاءة الاجتماعية وهو ما تسميه نظرية إمكانية التعميم بـــ Object of Measurement. فتفاعل الفرد مع عبارات المقياس يعكس تباين الوضع النسبي (الوضع الرتبي) للأطفال في السمة المقاسة عبر هذه العبارات، وتفاعل الفرد مع المُقيّمات يعكس تباين تقييمات المُقيّمات لسلوك الكفاءة الاجتماعية لدى الأطفال، وتفاعل الفرد مع مرات التطبيق يعكس تباين الوضع النسبي (الوضع الرتبي) للأطفال في السمة المقاسة عبر مرات التطبيق (Brennan, 2001; Cardinet, Tourneur, & Allal, 1981). وبعبارة أوضح فأنه يمكن القول بأن الفروق الحقيقية بين الأطفال في سمة الكفاءة الاجتماعية كانت قادرة على أن تصنع تباينا في الوضع النسبي للأفراد عبر كل من عبارات المقياس، ومرات التطبيق، وتقييمات المُقيّمات.

وبالإضافة إلى هذه النتائج فإن نسبة الخطأ غير المفسر Undifferentiated Error (التفاعل الرباعي) في أي من تحليلات دراسات إمكانية التعميم في الدراسة الحالية لم تتعد ١٥% (التفاعل الرباعي بين الفرد والمُقيّمات ووقت التطبيق والعبارات) من التباين الكلي وهي نسبة مقبولة إلى حد بعيد بالأخذ في الاعتبار تعقد التصميم ثلاثي الأوجه تام التقاطع Fully Crossed Three-Facet Designs الذي استخدمته

الدراسة الحالية وكذلك تعدد أوجه الخطأ المحتملة (المُقيّمات، وعبارات المقياس، ووقت التطبيق) والتي تم تضمينها في هذا التصميم. وبوجه عام يشير خطأ القياس غير المفسر إلى أي مصدر (وجه) لخطأ القياس لم يتم تضمينه في تصميم الدراسة ولكن هذه الأوجه يمكنها تفسير قدر محدد من التباين الكلي (Brennan, 2000). وقد أظهرت نتائج بعض الدراسات نسب مرتفعة من الخطأ غير المفسر حيث كانت النسبة في دراسة ليان وهارمند (Lian & Harmand, 2015) ٣٠% وهي نفس النسبة التي سجلتها دراسة مابي (Mabe, 2014).

وفي المقابل فإن الفرد باعتباره مصدر التباين الحقيقي True Variance قد فسر نسب مرتفعة من التباين الكلي وهو ما يعني أنه بالأخذ في الاعتبار متوسط المُقيّمات والعبارات ووقت التطبيق فإن هناك فروقا حقيقية بين الأطفال في الكفاءة الاجتماعية كما تقيسها النسخة العربية لمقياس الكفاءة الاجتماعية المُعَّدل حيث إن نسبة ما فسرته الفروق الفردية الحقيقة بين الأطفال لم تقل عن ٤٠% في أي من تحليلات دراسات إمكانية التعميم. وتتفق هذه النتيجة مع نتائج عدة دراسات والتي أظهرت أن نسبة التباين الحقيقي الذي يصف الفروق الحقيقية بين المفحوصين في المهارات الاجتماعية كانت مرتفعة ومنها دراسة ليان وهارمند (Lian & Harmand, 2015) حيث كانت نسبة التباين الحقيقي ٣٥%، ودراسة لوكا وشينج (Luka & Cheung, 2013) حيث كانت نسبة التباين الحقيقي ٤١%، ودراسة مابي (Mabe, 2014) حيث كانت نسبة التباين الحقيقي ٤٣%.

وبوجه عام، تشير هذه النتائج إلى انخفاض حجم التباين الكلي الذي تفسره أوجه المُقيّمات، وعبارات المقياس، ووقت التطبيق في الدرجات الملاحظة التي تصف الكفاءة الاجتماعية لدى الأطفال سواء كانت هذه الأوجه بصورة منفردة أو من خلال التفاعل بين بعضها البعض، أو التفاعل بينها وبين مصدر التباين الحقيقي (الفرد)، في حين ترتفع نسبة التباين الحقيقي الذي يصف فروقا حقيقية بين الأطفال في الكفاءة الاجتماعية وهو ما يعني أن المقياس يتمتع بمستويات مقبولة وكافية من الثبات. ويمكن في ضوء هذه النتائج وفي إطار نظرية إمكانية التعميم توصيف ثبات المقياس بأنه دقة التعميم من الدرجة الملاحظة Observed Score للأطفال على المقياس إلى متوسط الدرجة التي سيحصل عليها الأطفال تحت كل الشروط الممكنة مثل كل نسخ المقياس الممكنة، أو كل مرات تطبيق المقياس الممكنة، أو كل أسئلة المقياس الممكنة أو كل المُقيّمات الممكنة. ويطلق على

متوسط الدرجة هذه الدرجة الشاملة Universe Score وهي تشبه الدرجة الحقيقة True Score في النظرية الكلاسيكية للقياس (Kieffer,1998). فنظرية إمكانية التعميم تعتبر الدرجات الملاحظة ثابتة إذا ما أمكن من خلال هذه الدرجات الوصول إلى استنتاجات دقيقة عن نطاق الملاحظات المقبولة Universe of Admissible Observation التي تمثلها (Cronbach et al., 1972; Eason, 1989).

ويتمثل الهدف الثاني في الدراسة الحالية في الكشف عن الشروط التي يمكن من خلالها أن تحقق الصورة العربية لمقياس الكفاءة الاجتماعية المُعَّدل أفضل مستويات الثبات حيث يتضح من خلال نتائج دراسات اتخاذ القرارات D Studies وما نتج عنها من معاملات إمكانية التعميم أنه في جميع الحالات فإن استخدام اثنتين من المُقيّمات وتطبيق المقياس مرتين يُعد إجراءً كافيا للحصول على أعلى قيم لمعاملات إمكانية التعميم. ولكن بالتدقيق في جدول (٣) فإنه يلاحظ أن زيادة عدد المُقيّمات إلى ثلاث مُقيّمات وزيادة عدد مرات التطبيق إلى ثلاث مرات يزيد من قيم معاملات إمكانية التعميم للمقياس، ولكن هذه الزيادات المتوقعة في معاملات إمكانية التعميم تبدو ضئيلة وغير مؤثرة مقارنة بالوقت والجهد المطلوبان لجمع مزيدا من هذه البيانات. وتتفق هذه النتيجة مع نتائج دراسة قام بها فريق من الباحثين من مركز تطوير التعلم والمهارات الذاتية في جامعة أوساكا في اليابان (Lueng et al., 2010) حيث أظهرت نتائج دراسات القرارات D-Studies أن استخدام ثلاثة مُقيّمين وتطبيق المقياس مرتين يُعد إجراءً كافيا للحصول على قيم مرتفعة لمعاملات إمكانية التعميم وذلك على الرغم من أن استخدام ثلاثة مُقيّمين وتطبيق المقياس ثلاث مرات يمكن أن يزيد من قيم معاملات إمكانية التعميم بمقدار ٢% إلا أن الباحثين في الدراسة اليابانية اعتبروا أن هذه النسبة ضئيلة وغير مؤثرة ولا تستحق زيادة عدد مرات التطبيق من مرتين إلى ثلاث مرات.

وتعليقا على هذه النتيجة، يؤكد بعض الباحثين (Brennan, 2001; Kieffer, 1998; O'brien, 1995) على أن مقدار الزيادة في معاملات إمكانية التعميم المتوقعة من معالجة مصادر الخطأ يجب أن تزيد بمقدار ١٠% على الأقل عن قيمة معاملات إمكانية التعميم الأصلية خاصة إذا ما كانت معاملات إمكانية التعميم الأصلية تشير إلى مستويات كافية من الثبات. وبمراجعة جدول (٣) في الدراسة الحالية يتضح أن أعلى زيادة متوقعة في معامل إمكانية التعميم كانت في عامل

المبادرة الاجتماعية في الفصل الثاني حيث من المتوقع أن يزيد معامل إمكانية التعميم من ٩٢٦، (اثنتين من المقيمات وتطبيق الأداة مرتين) ليصبح ٩٥٧، (ثلاث مقيمات وتطبيق الأداة ثلاث مرات) أي بزيادة قدرها ٠٣١،% وهو فرق ضئيل مقارنة بالقيمة الاسترشادية ١٠%.

وتشير نتائج دراسات اتخاذ القرارات في الدراسة الحالية إلى أنه لو أراد الباحث الاختيار ما بين زيادة عدد المُقيّمات أو عدد مرات التطبيق لتحسين ثبات المقياس فإنه من المناسب زيادة عدد المُقيّمات إلى مُقيّمتين اثنتين وأن يتم تطبيق المقياس مرة واحدة فقط، وهذه نتيجة متوقعة إذ أنها تتفق مع طبيعة بيئة جمع البيانات في مدارس التعليم الأساسي (١-٤) في سلطنة عمان حيث يتوفر اثنتان من المُقيّمات هما معلمة الفصل والمعلمة المساعدة. وأخيرا، أظهرت نتائج دراسات القرارات أن قيم معاملات إمكانية التعميم للمقياس تكون أقل ما يمكن عند استخدام مُقيّمة واحدة وتطبيق المقياس مرة واحدة، وهي نتيجة منطقية ومتوقعة إذا أن مثل هذه الممارسات سترفع من مقدار نسبة تباين الخطأ إلى التباين الكلي (O'brien, 1995). وتتفق هذه النتيجة مع نتائج عدة دراسات أخرى خاصة فيما يتعلق بانخفاض معاملات إمكانية التعميم عند تطبيق المقياس مرة واحدة (Broger et al., 2012; Carminla, 2010; Christ, 2012; Lueng et al., 2010).

وختاما، وفي ضوء ما تم مناقشته من نتائج الدراسة الحالية فإنه يمكن القول إن النسخة العربية لمقياس الكفاءة الاجتماعية المُعَّدل تتمتع بمستويات جيدة من الثبات (معاملات إمكانية التعميم) ومن ثم يمكن استخدامها من جانب المعلمين في تقييم الكفاءة الاجتماعية لدى الأطفال من سن ٧ إلى ١٠ سنوات حيث توصي النتائج الحالية بأن يتم التقييم من خلال اثنتين من المُقيّمات وتطبيق الأداة مرتين الحصول على أعلى معاملات إمكانية التعميم.

بعض التضمينات المهمة في مجال القياس النفسي والتربوي:

في ضوء نتائج الدراسة الحالية، يمكن تحديد بعض التضمينات المهمة فيما يتعلق بتقييم السمات النفسية والتربوية في إطار نظرية إمكانية التعميم:

(١) ضرورة التدقيق في اختيار المُقيّم بحيث نراعي فيه المؤهل العلمي، والخبرة السابقة بالسمة موضوع التقييم.

(٢) التأكد من معرفة المُقيّم بأداة التقييم والهدف منها وكيفية تطبيقها وتصحيحها مع ضرورة تدريبه على الأدوات الجديدة.

(٣) تخصيص وتحديد أعداد المفحوصين بما يتناسب مع قدرات وإمكانيات المُقيّم.
(٤) التحقق من توافر عامل الألفة بين المُقيّم والمفحوصين حتى لا يغير المفحوصون أنماط سلوكياتهم.
(٥) توجيه وتوعية المُقيّم إلى ضرورة التدقيق والتركيز أثناء عملية التقييم بحيث يميز بين السلوكيات التي تعكس سمات نفسية وتربوية متأصلة لدى المفحوص وتلك التي تمثل مجرد سلوكيات عابرة.
(٦) توجيه وتوعية المُقيّم إلى ضرورة عدم الانسياق وراء أخطاء البنية الشخصية مثل التساهل أو الليونةLeniency أو أثر الهالة Halo Effect وذلك عند إجراء عملية التقييم.
(٧) تفعيل عامل استقلالية التقييم بحيث لا يتأثر تقييم أحد المُقيّمين بآراء وتوجهات مُقيّم أخر ويتم ذلك من خلال توجيه المُقيّمين بعدم مناقشة تقييماتهم للمفحوصين مع بعضهم البعض.
(٨) اللجوء إلى جمع البيانات من أكثر من مُقيّم وفي أكثر من مناسبة بحيث يمكن تقليل أخطأ القياس المتوقعة من الاعتماد على مُقيّم واحد وجمع البيانات مرة واحدة.

المراجـــــع

أولاً- المراجع العربية:

الحسيني، زينب وعبد السلام شاكر (٢٠١٤). *الشراكة الأسرية- المدرسية: رؤية نقدية معاصرة*. مكتبة العهد الجديد، بعلبك: لبنان.

عبيدات، ذوقان وعدس، عبد الرحمن؛ وعبد الحق، كايد (٢٠٠٥). *البحث العلمي: مفهومه، أدواته، أساليبه*. الرياض: دار أسامة للنشر والتوزيع.

علام، صلاح الدين محمود (٢٠٠٢). *القياس والتقويم والتربوى والنفسى. أساسياته وتطبيقاته وتوجهاته المعاصرة*. دار الفكر العربي للطبع والنشر: القاهرة.

ثانيًا- المراجع الأجنبية:

Anastasi, A. (1976). *Psychological testing*. New York: MacMillan Publishing Company Inc.

Arbuckle, J. L. (2015). AMOS (Version 22.0) [Computer program]. Chicago, IL: SPSS.

Bandura, A. (2008). Toward an agentic theory of the self. In H.W. Marsh, R.G. Craven, & D.M. McInerey (Eds.). *Self-processes, learning, and enabling human potential* (pp. 15-49). International advances in self-research. Greenwich, CT.: Information Age Publishing.

Blumberg, S. J., Carle, A. C., O'Connor, K. S., Moore, K. A., & Lippman, L. H. (2008). Social competence: Development of an indicator for children and adolescents. *Child Indicators Research, 1(2),* 176-197.

Blunch (2008). *Introduction to structural equation modelling using SPSS and AMOS.* New York: Sage Publications

Bornstein, M. H., Hahn, C., & Haynes, O. M. (2010). Social competence, externalizing, and internalizing behavioral adjustment from early childhood through early adolescence: Developmental cascades. *Development and Psychopathology, 22(4),* 717-735.

Brennan, R. L. (1997). A perspective on the history of generalizability theory. *Educational Measurement: Issues*

and Practice, 16, 14-20.

Brennan, R. L. (2000). (Mis)conceptions about generalizability theory. *Educational Measurement: Issues and Practice, 19*, 5-10.

Brennan, R. L. (2001). *Generalizability Theory*. New York: Springer

Broger, T., Saintmeen,C., & Keleven, W. (2012). The generalizability of systematic rating scales across time: A preliminary analysis using the generalizability theory. *Journal of Educational Psychology*,3, 75-92.

Brown, T.A. (2006). Confirmatory factor analysis for applied research. New York: Guilford Press.

Brown, W. H., Odom, S. L., & McConnell, S. R. (2008). Social competence of young

children: Risk, disability, and intervention. New York: Paul H. Brookes Publications

Burt, K. B., Obradovic, J., Long, J. D., & Masten, A. S. (2008). The interplay of social competence and psychopathology over 20 years: Testing transactional and cascade models. *Child Development, 79(2),* 359-374.

Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2 ed.). New York, NY: Routledge.

Caldarella, P., & Merrell, K. W. (1997). Common dimensions of social skills of children and adolescents: A taxonomy of positive behaviors. *School Psychology Review, 26*, 264-278.

Cardinet, J., Tourneur, Y., & Allal, L. (1981). Extensions of generalizability theory and its applications in educational measurement. *Journal of Educational Measurement, 18*,183-204.

Carminla, S. (2010). Social function assessment tools for children within the framework of the generalizability theory.

Clinical Psychology Review, 4, 52-65.

Cassidy, M. (2005). 'They do it anyway': A study of Primary 1 teachers' perceptionsof children's transition into primary education. *Early Years, 25*, 143-153.

Cavell, T. (1990). Social adjustment, social performance, and social skills: A tri-component model of social competence. *Journal of Clinical Child Psychology*, 19,111–122.

Christ, G. (2012). An application of the generalizability theory to the Social Skills Improvement System –Rating Scales. *North American Journal of Psychology*, 3, 25-37.

Cohen, J.(1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 2, 37-46.

Cone, J. D. (1977). The relevance of reliability and validity for behavioral assessment. *Behavior Therapy*, 8, 411-426.

Conduct Problems Prevention Research Group (1999). Initial impact of the fast track prevention trial for conduct problems: The high-risk sample. *Journal of Consulting and Clinical Psychology, 67*, 631-647.

Crick, J. E., & Brennan, R. L. (1982). GENOVA: A generalized analysis of variance system (FORTRAN IV computer program and manual). Dorchester, MA: Computer Facilities, University of Massachusetts at Boston.

Crocker, L., & Algina, J. (2006). *Introduction to classical and modern test theory*. Belmont, CA: Wadsworth Publishing Company

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability scores and profiles*. New York: Wiley.

Cronbach, L. J., Rajaratnam, N., & Gleser, G.C. (1963). Theory of generalizability: A liberization of reliability theory. *British Journal of Statistical Psychology, 16*, 137-163.

DeCarlo, L. (1997). On the meaning and use of kurtosis.

Psychological Methods, 2, 292-307.

Dennis, L.R. & Stockall, N, (2015). Using play to build the social competence of young children with language delays: Practical guidelines for teachers. *Early Childhood Education Journal, 43,* 1-7

Dockett, S., & Perry, B. (2007). Children's transition to school: Changing expectations. In A. Dunlop & H. Fabian (Eds.), *Informing transitions in the early years: Research, policy and practice* (pp. 26-42). Maidenhead, England: Open University Press.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students‘ social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development, 82(1),* 405-432

Durlak, J. A., & Wells, A. M. (1997). Primary prevention mental health programs for children and adolescents: A meta-analytic review. *American Journal of Community Psychology, 25,* 115-152.

Eason, S. H. (1989, November). *Why generalizability theory yields better results than classical test theory.* Paper presented at the Annual Meeting of the Mid-South Educational Research Association, Little Rock, Arizona.

Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (2006). Prosocial development. In N. Eisenberg (Ed.), *Handbook of child psychology*: Vol. 3. Social, emotional, and

personality development (6^{th} ed., pp. 646–718). Hoboken, NJ: John Wiley & Sons.

Foster, S. L., Inderbitzen, H. M., & Nangle, D. W. (1993). Assessing acceptance and social skills with peers in childhood: Current issues. *Behavior Modification,17,* 255-286.

Hair, J. F. Jr., Anderson, R. E., Tatham, R. L. & Black, W. C. (1995). Multivariate Data Analysis (3rd ed). New York: Macmillan.

Gresham, F. M. (1986). Conceptual and definitional issues in the assessment of children's social skills: Implications for classification and training. *Journal of Clinical Child Psychology, 15,* 3-15.

Gresham, F. M., Carey, M. P. (1988). Research methodology and measurement. In Witt, J., Elliott, S., & F. Gresham (Eds). *Handbook of behavior therapy in education* (pp. 37-65). New York, NY, US: Plenum Press.

Gresham, F. M. & Elliott, S. N. (1990). *The Social Skills Rating System*. Circle Pines MN: American Guidance Service.

Gresham, F. M., & Elliott, S. N. (2008). *The Social Skills Improvement System: Rating Scales*. Bloomington, MN: Pearson Assess

Gresham, F. M., Elliott, S. N., Cook, C. R., Vance, M.J., & Kettler, R. J. (2010). Cross informant agreement for social and problem behavior ratings: An investigation of the Social Skills Improvement System Rating Scales. *Psychological Assessment, 22,* 157-166.

Gresham, F. M. & Reschly, D. J. (1987). Dimensions of social competence: Method factors in the assessment of adaptive behavior, social skills, and peer acceptance. *Journal of School Psychology, 25,* 367-381.

Hancock, G.R. & Mueller, R.O. (2006*). Structural equation modeling: A second course*. Greenwich, CT: Information Age Publishing.

Hand, D. J. (2010). *Measurement theory and practice: The world through quantification*. New York: Wiley

Kieffer, K. M. (1998, April). *Why generalizability theory is essential and classical test theory is often inadequate.* Paper presented at the Annual Meeting of the Southwestern

Psychological Association, New Orleans, LA.

Kline, R. (2010). Principles and practice of structural equation modeling (3rd ed.). New York: Guilford Press.

Ladd, G. W. (1999). Peer relationships and social competence during early and middle childhood. *Annual Review of Psychology, 50,* 333-359.

Landis J. R, Koch G. G. (1977) The measurement of observer agreement forcategorical data. *Biometrics*, 33:159-74.

Lian, Z. & Harmand, D. (2015). Social competence in the school setting: An evaluation of the generalizability theory. *Australian Journal of Psychology,* 2, 22-31.

Luka, T. & Cheung,S. (2013). The generalizability of social skills in a competition context across time and rater. *Journal of Applied Social Science*.4, 74-84.

Lueng, T. Lui, R. & Yim, H. (2010). Advances in the assessment of social competence: Finding from a preliminary dependability investigation in a Japanese sample. *Japanese Journal of Educational Psychology,* 2, 56-79. Mabe, M. (2014). *Validation of a social skills construct using multitrat-multimethod and generalizability approaches*. (Unpublished doctoral dissertation). University of Rhode Island, Rhode Island, USA.

Maneesriwongul, W., & Dixon, J. K. (2004). Instrument translation process: A methods review. *Journal of Advanced Nursing Research, 48*, 175-186.

Merrell, K. W. (1993). Using behavior rating scales to assess social skills and antisocial behavior in school settings: Development of the School Social Behavior Scales. *School Psychology Review*, 22, 115-133

Merrell, K. W. (2001). Assessment of children's social skills: Recent developments,

best practices, and new directions. *Exceptionality, 9,* 3-18.

Milfort, R., & Greenfield, D. B. (2002). Teacher and observer ratings of head start children's social skills. *Early Childhood Research Quarterly, 17*, 581-595.

Missal, K. N., & Hojnoski, R. L. (2008). The critical nature of young children's emerging peer-related social competence for transition to school. In W. H. Brown, S. L. Odom & S. R. McConnell (Eds.), *Social competence of young children: Risk, disability, and intervention* (pp. 117-137). Baltimore, MD: Paul H. Brookes Publishing Co.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

O'brien, R. M. (1995). Generalizability coefficients are reliability coefficients. *Quality and Quantity, 29 (4),* 421-428

Ribar, R. J. (2005).Generalizability theory and Social Competence Performance Checklist (Unpublished doctoral dissertation). The University of Wisconsin Milwaukee, Wisconsin, USA.

Renk, K., & Phares, V. (2004). Cross-informant ratings of social competence in children and adolescents. *Clinical Psychology Review, 24 (2),* 239–254

Reynolds, C. R. & Kamphaus, R. W. (2004). *Behavior Assessment System for Children* (2nd ed.). Circle Pines MN: American Guidance Service.

Rydell, A. M., Hagekull, B., & Bohlin, G. (1997). Measurement of two social competence aspects in middle childhood. *Developmental Psychology, 33,* 824-833.

Rydell, A. M. & Hagekull, B. (2012). The development of Social Competence Inventory-Revised: Testing for validity and reliability indices. *Advances in Social Psychology, 2 (1), 39-54.*

Semrud-Clikeman, M. (2007). *Social competence in children*. New York: Springer.

Shah, F., & Morgan, S. B. (1996). Teachers' ratings of social competence of children with high versus low levels of depressive symptoms. *Journal of School Psychology, 34*, 337-349.

Shavelson, R.J., & Webb, N.M. (1981). Generalizability theory: 1973 – 1980. *British Journal of Mathematical and Statistical Psychology, 34*, 133- 166.

Shavelson, R. J., & Webb, N. M. (1991). *Generalizability theory*. Newbury Park, CA: Sage

Smith, P. L. (1981). Gaining accuracy in generalizability theory: Using multiple designs. *Journal of Educational Measurement, 18*, 147-154.

Stoiber, K. C. (2004). Social competence performance checklist. Functional Assessment and Intervention System. San Antonio TX: The Psychological Corporation.

Stoiber, K. C. (2011). Translating knowledge of social emotional learning and evidence based practice into responsive school innovations. *Journal of Educational and Psychological Consultation, 21*, 46-55.

Tabachnick, B. G., & Fidell, L. S. (2001). Using Multivariate Statistics (4th ed.). Boston, MA: Allyn and Bacon

Webster-Stratton, C., & Lindsay, D. W. (1999). Social competence and early-onset conduct problems: Issues in assessment. *Journal of Child Clinical Psychology, 28*, 25-93.

Welsh, M., Parke, R., Widaman, K., & O'Neil, R. (2001). Linkages between children's social and academic competence: A longitudinal analysis. *Journal of School Psychology, 39*, 463-481.