RESEARCH ARTICLE

Directions for Professional Development: Increasing Knowledge of Early Childhood Measurement

Rebecca Gokiert

University of Alberta

Lisa Noble

Glenrose Rehabilitation Hospital

Lori Baugh Littlejohns

School of Public Health, University of Alberta

Research has indicated that the education and experience of early childhood (EC) professionals in measurement is often inadequate to meet the requirements for best practice, and very few studies have investigated this issue. A survey was used to explore the measurement knowledge and use of 159 EC professionals to help set directions for professional development. Results indicated that respondents had a very high interest in professional development and for the most part had received little to no training for the tools and procedures they were currently using in practice. Furthermore, moderate to large effects were found between an individual's measurement knowledge and competency and how important they believed it was for their practice across all measurement domains (screening, assessment, and evaluation). The majority of respondents (93.9%) reported that they modify measurement procedures, which can have implications for the reliability and validity of the tool. Respondents highlighted various measurement issues and needs in terms of professional development, such as the need for measurement standards across programs, how to appropriately modify measurement tools/procedures to meet individual program and professional needs, and how to link measurement information to programming.

Research has consistently confirmed the importance of early childhood development in setting the foundation for learning, behaviour, and health outcomes throughout the lifespan (Nelson, 2000; Shonkoff & Phillips, 2000; Greenspan & Meisels, 1996; Campbell & Ramey, 1994; Hertzman, 1999; McCain, Mustard, & Shanker, 2007; VanLandeghem, Curgins, & Abrams, 2002). Given that not all environments and experiences in early childhood are optimal, it is estimated that one quarter of Canada's children between the ages of zero to six are experiencing difficulties associated with learning and/or behaviour (McCain, et al., 2007). Given this high estimate, many programs and services are aimed at supporting the positive development of all young children with targeted intervention for those experiencing difficulties (Anderson, et al.,

2003; Dawson & Osterling, 1997; Nelson, 2000; Ramey & Ramey, 2004; Schweinhart, Barnes, & Weikart, 1993). For this reason, early childhood (EC) professionals are playing an increasing role in the measurement of young children in childcare, preschool, head start, intervention, and school settings. Measurement processes can be used to monitor development, plan and implement effective early childhood intervention and services, and evaluate program effectiveness (Meisels & Atkins-Burnett, 2000; Administration for Children and Families, 2002). In early childhood literature and the field, three complementary and often overlapping forms of measurement (screening, assessment and evaluation) are described (Appl, 2000). Each plays a critical role in enhancing the understanding of a child's development, the environments in which they develop, and the effectiveness of interventions and programs for supporting development.

   Assessment is defined as a process of gathering information and observations and making decisions about an individual based on the information (Bredekamp & Copple, 1997). The goal of assessment with respect to early childhood development is to acquire information and understanding of a child's developmental strengths and needs. If a child is experiencing difficulties, this information can guide referral or the development of programs and services to meet the needs of the child (Meisels & Atkins-Burnett, 2000). The term *screening* is often incorrectly used synonymously with *assessment* however screening is commonly differentiated to be a *purpose* of assessment. Screening can be defined as the use of a brief procedure or measurement tool designed to identify, from within a large population of children, those who may benefit from further assessment to verify developmental and/or health risks (Martella, 2004). The term *assessment* is also associated with *evaluation*. Martella (2004) defined evaluation in early childhood development as the measurement, comparison, and judgment of the value, quality or worth of children's work and/or of the environments in which they learn**.** EC professionals want to know if the supports they provide are helping children develop, parents want to know if their children are making gains, and funders want to know if the programs they are financially supporting are producing positive outcomes (McConnell, 2000; Horton & Bowman, 2002).

   Significant issues related to the measurement practices of EC professionals have been raised in the literature (Gredler, 1997; Hirsh-Pasek, Kochanoff, Newcombe, & Villiers, 2005; Neisworth & Bagnato, 2000; Greenwood, Luze, & Carter, 2002). Over the past 15 years standards have been developed to guide best practices in early childhood assessment and programming (Administration for Children and Families, 2003; AERA, APA, & NCME, 1999; NAEYC, 2003, 2009; National Association of School Psychologists (NASP), 2005; National Education Goals Panel, 1998). These guidelines are incorporated into many program performance standards and policies (e.g., Head Start see U.S. Department of Health and Human Services, 2007), making explicit the importance of using research-based assessment methods and tools that possess adequate psychometric properties, and are accurate for the child being assessed. However, it is not clear if EC professionals are following these guidelines when selecting and appropriately using screening and assessment tools in their practice (Allen, 2007; Appl, 2000; Bredekamp & Copple, 1997; Brown & Rolfe, 2005; Pretti-Frontczak, Kowalksi, & Brown, 2002; Sandall, McLean, & Smith, 2000). Shepard, Kagan, & Wurtz (1998) found, due to a lack of knowledge and formal training, that EC professionals are using tests to address multiple purposes which compromises the validity of the inferences that can be made from results. Many assessment tools have a distinct purpose and therefore using them interchangeably for alternative purposes is not recommended. Furthermore, if there is not a clear purpose that will result in positive benefits for the child, an assessment should not occur (McConnell, 2000; Shepard,

1994). The use of tools for multiple purposes is likely influenced by increasing pressure for programs to adhere to accountability frameworks and produce outcomes for funding agencies (Hirsh-Pasek et al., 2005).

Few studies have investigated the assessment practices and beliefs of EC professionals. Sheperd (1997) highlighted the need for knowledge of developmental indicators and test administration as key factors in valid and reliable assessment. However, the majority of individuals responsible for this are lacking in both. It has also been found that the education and experience of EC professionals is often inadequate to meet the requirements for best practice in assessment (NAEYC, 2009). Pretti-Frontczak et al. (2002) conducted a large-scale study of teachers in Ohio (n =586) to examine the degree to which best practices in assessment were being implemented. The relationship between teacher education, program type, years of experience, and the use of assessment was examined. Results indicated that teachers who had higher levels of education reported using a larger range of assessments than those with less education. There were varied responses with respect to whether particular assessments were directly linked to their intended purpose. It was also reported that there seemed to be a widespread use of self-developed or program-developed assessment tools/tests and caution was expressed that these may lack important psychometric properties. Others have noted that those conducting assessments are often frustrated with the inadequacy of available measurement tools in meeting needs and may modify existing tools without a clear understanding of the validity and reliability implications (Allen, 2007; Appl, 2000; Bordignon & Lam, 2004). The misuse and modification of standardized tools, is likely exacerbated by the vast number of assessment tools currently available and the multiple domains of child development measured (Niemeyer & Scott-Little, 2001; Berry, Bridges, & Zaslow, 2004).

Questions arise as to what assessment tools EC professionals are using. The Centre of Excellence for Children and Adolescents with Special Needs (Kineapple, Lyon, McSorely, Morse, & Smith, 2005) in Canada developed a list of the most commonly used assessment tools for children with special needs. This was based on a review of the literature and a survey of agencies and individuals working with young children with special needs and their families. The report outlined numerous tools for children aged birth to six years that are available, however, it did not provide information regarding the extent to which certain tools are used, the reasons that specific tools are selected, and whether these tools are used appropriately.

Brown and Rolfe (2005) explored whether EC professionals (n=10) in Australia were using formal (standardized instruments) or informal (non-standardized instruments) assessments. They also asked students (n=10) in the final year of their bachelor of early childhood program about their intentions to use assessment in their future practice. The majority of practitioners (90%) and all students reported that they use or plan to use informal assessments. Seventy percent of students planned to use formal assessments in their future practice compared to only 10% of professionals. EC professionals rated "ease of use" as their primary rationale for choosing assessments and students rated "accuracy of instrument" as the most important characteristic for choosing an assessment instrument. Students noted that "identifying children at-risk" was the most important reason to engage in assessment while "providing a good overall picture of development" was cited as the most important reason for choosing assessments by EC professionals. Given the small sample size, it is difficult to generalize from these results; however the study provides some insights into pre- and current EC professional's perceptions of assessment.

There is a call for research to investigate assessment tools that are used by EC professionals and whether usage is aligned with the intended purpose (Hirsh-Pasek et al., 2005). Furthermore, there is scarce literature with respect to Canadian EC professionals' assessment knowledge and practices. Therefore, the purpose of this research was to explore (a) the current measurement knowledge of EC professionals in a mid-sized metropolitan area in Canada, (b) the measurement tools that EC professionals are using, and (c) the extent to which the tools are meeting their needs and being used for the right purpose.

RESEARCH METHODS

In order to explore the measurement knowledge, use, and needs of EC professionals, a comprehensive survey was developed and implemented for this research. The survey was used as an efficient means to gather information from a large group of EC professionals.

Survey Development

The survey instrument was developed in two stages. First, a review of the literature was conducted to identify key definitions, concepts, and issues in EC measurement, and from this information a preliminary survey was developed. This step was considered important to build upon past research and strengthen the content validity of the survey instrument. Second, the preliminary survey was pre-tested through a focus group with ten individuals working within the EC field. To enhance the relevancy of the tool to the local context individuals were selected to represent front-line service delivery (i.e., early intervention, childcare, head start, Aboriginal and immigrant serving agencies, social and health services sectors), and funding agencies. Focus group participants reviewed and evaluated the survey instructions and each question with respect to clarity, relevance and purpose, and response formats. Based on recommendations and feedback from the focus group members, instructions and questions were modified to refine the clarity and purpose of the survey instrument.

The final survey instrument was comprised of 68 close- and open-formed questions, likert scales, and fill in the blank questions across four sections: (1) background information, (2) measurement knowledge and competency, (3) measurement issues and needs, and (4) measurement tool use. Definitions for assessment, screening, and evaluation were provided for participants to refer to while completing the survey. For a copy of the survey please contact the first author.

The first section, *background information*, asked for participant's highest level of education, years worked in early childhood, current occupation, services regularly provided, ages of clientele, and primary population served across six questions.

The second section, *measurement knowledge and competency*, was the lengthiest component of the survey instrument. Nine questions were used to gauge if respondents modified measurement procedures and the main purposes for screening, assessment and evaluation in their practice. Participants were then asked to rate their knowledge/competency and degree of importance they placed with respect to several areas of measurement using a likert scale across 25 questions.

The 30 questions that comprised the third section, *measurement issues and needs*, centered on what participants believed to be the major measurement issues and training needs among EC professionals. Respondents were asked specific questions about the importance of tool selection and use, perceptions of testing procedures, and areas where they perceive they require more training. Participants were also asked to respond to two open-ended questions about issues related to measurement, and the measurement services and resources they would deem beneficial to their practice.

The final section, *measurement tool use,* asked participants for the name of measurement tools used, the amount of training they had received with respect to the tool, the purpose of the tool, a description of the tool, and how the information gathered from the tool is used in their practice. These five questions were repeated if respondents use more than one tool in their practice.

It was determined that the survey be made available in two formats to increase participation. A paper-based survey was developed in order to make it available to the sample population without access to a computer and the Internet. An online survey was also constructed to provide another means of participation. The paper-based survey was formatted onto transoptic® sheets that could be easily scanned and compiled in a database. A web-based survey was developed utilizing Survey Monkey (see www.surveymonkey.com), a web–based survey software. The survey took approximately 10-15 minutes to complete.

Sampling

The study utilized a convenience sample as the field of early childhood is very broad in scope and there is not one single contact point for distributing information. The research team relied heavily on partnerships with community agencies, school boards, individuals that participated in the focus group for survey development, and interagency networks to distribute the survey in order to reach the spectrum of EC professionals. There are obvious drawbacks to a convenience sample, such as a lack of representation of the population and therefore, limiting the ability to generalize findings. Despite the obvious drawbacks to a convenience sample, this research provides information to enhance understanding of the measurement knowledge, use, and needs of EC professionals.

The target distribution sample included 530 invitations to early childhood development professionals with potential knowledge and a role in the assessment of children aged birth to six years. The survey was distributed over a six-month period through list serves, formal networks, school administrators, community organizations, and mail out to local childcare sites. EC professionals included psychologists, speech and language pathologists, teachers, directors/managers, community daycare and family day home providers, and early intervention specialists (i.e., individuals working in head start programs and family resource centres).

Of the target 530, the paper-based questionnaire was mailed to 430 individuals or organizations and included a letter of information with instructions and dates for return, a consent form, and a prepaid return envelope. A significant number of the paper-based surveys were mailed directly to childcare centres and family day home agencies. An email invitation was sent to 100 individuals, organizations, and interagency networks, and described the purpose of the study through an information letter and included the link to the web-based survey instrument. Individuals consented to participate by reading the information letter and indicating that they

understood the purpose of the research and wanted to continue with the survey. The total response rate was 29.4% (n=159), with n=54 (34%) responding via paper-based survey and n=105 (66%) responding via the web-based survey. While the overall response rate is lower than would be expected, the sample of respondents as described in the background information portion of the results section is very indicative of individuals from the larger sample that would have a direct connection to measurement within their work. The survey was sent to a large proportion of frontline early childcare staff (e.g., daycare) and while their response rate was low this is not surprising as they are not typically responsible for aspects of measurement (screening, assessment and/or evaluation) within their practice.

Data Management and Analysis

The paper-based surveys used transoptic® sheets that allowed for machine scanning and later compilation of an SPSS database and thereby eliminating the need for manual data entry and the possibility of errors. The online survey data also reduced data entry error, in that the web-based program takes all entered questions by participants and compiles a downloadable Excel spreadsheet. The paper-based and online survey data was merged into one database using SPSS. Most of the data analysis procedures were descriptive in nature including frequencies, means, and standard deviations in order to summarize findings. In addition, t-test analyses were conducted utilizing an alpha of .05 to assess significant differences between groups. Effect sizes were also calculated to determine the magnitude of the difference if the t-test was found to be statistically significant. The following guidelines (Cohen, 1988) were followed when interpreting effect sizes: $d = 0.2$, small; $d = 0.5$, moderate; and $d = 0.8$, large.

RESULTS

Of the 159 surveys completed, participants did not respond to every question and this resulted in missing data across the survey. An overall examination of missing data did not reveal strong associations between certain types of questions, content, or format. Most data is reported as a percentage of total respondents that answered the question with the exclusion of missing data. In the following section, key results from each of the survey sections are presented in the order that they appear on the survey: background information, knowledge and competency, issues and needs, and measurement tool use. Where appropriate, internal consistency is reported for each section of the survey with coefficients ranging from 0.80 to 0.95.

Background Information

Background information of survey respondents is reported in Table 1. Over half of the respondents stated that the highest level of education obtained was an undergraduate degree (56.8%) followed by almost a quarter reporting that they had a college diploma. It is of note that 21% of respondents had a Master's degree. Overall the respondents were well educated and also had substantial experience in the early childhood development field. There were very few

respondents with less that one year of experience (4.2%) and 70.2% had over 6 years experience. It is notable that 21.5% had over 21 years of experience.

Respondents were asked to identify their current occupation in terms of a 'best fit' from a list. Teacher was the most frequent occupation reported (42.8%), Directors (26.9%) the second largest occupation, and 10.3% of respondents indicated they were early childhood educators. The most frequently provided service was *educational programming for children* (75.5%), and two additional services that fall under the umbrella of educational programming [*individualized programming* (67.3%), and *group/classroom programming* (65.4%)]. A large number of respondents also reported that they provide early childhood development support, assessment, and screening (over 50%). Majority of EC professionals reported that they work with five-(81.8%) and four-year-olds (79.2%). Respondents were asked to indicate all the populations of children that they serve and at least half indicated they work with *typically developing children* (73.6%), *children with English as a second language* (54%), *children with diverse cultural backgrounds* (53.5%), *children with disabilities* (48.4%), and *children exposed to at-risk conditions* (42.8%).

TABLE 1
Background Information of Respondents and Clientele

Variable	Descriptive	
Highest Education (n=146)	*Frequency*	*Percent*
High school	1	0.7
Some college/university	7	4.8
College diploma	36	24.7
Undergraduate degree	83	56.8
Master's degree	19	13
Years worked in Early Childhood (n=144)		
less than 1	6	4.2
1-5	37	25.7
6-10	26	18.1
11-15	27	18.8
16-20	17	11.8
21+	31	21.5
Occupation (n=145)		
Teacher	62	42.8
Support service provider	4	2.8
Educational assistant	2	1.4
Early childhood educator	15	10.3
Outreach worker/home visitor	9	6.2
Director	39	26.9
Consultant	4	2.8
Speech language pathologist	8	5.5
Psychologist	1	0.7
Nurse	1	0.7

(Continue)

Services Provided Regularly (n=159)		
Educational programming for children	120	75.5
Educational programming for families	61	38.4
Individualized programming	107	67.3
Group/classroom programming	104	65.4
Early childhood development	109	68.6
Childcare	52	32.7
Screening	72	45.3
Assessment	91	57.2
Research	8	5
In-service to other organizations	24	15.1
Consultation	46	28.9
Home-based intervention	28	17.6
Outpatient Therapy	7	4.4
Ages of Clientele		
Prenatal	12	7.5
Birth to 1 year	32	20.1
1 year	46	28.9
2 years	62	39
3 years	83	52.2
4 years	126	79.2
5 years	130	81.8
6-8 years	59	37.1
Greater than 8	21	13.2
Primary Populations Served (n=159)		
Typically developing children	117	73.6
Children with English as a second language	86	54
Children with diverse cultural backgrounds	85	53.5

Measurement Knowledge/Competency

Level and importance of knowledge/competency.   Respondents were asked to indicate their level of knowledge/competency in several measurement domains and also to rate the importance of having knowledge/competency for their position. Respondents rated their level of knowledge/competency on a likert scale ranging from one to five (1=*not at all knowledgeable,* 2=*not very knowledgeable*, 3=*somewhat knowledgeable*, 4=*knowledgeable*, and 5=*extremely knowledgeable*) and importance from one to five (1=*not at all important*, 2=*not very important*, 3=*somewhat important,* 4=*important*, and 5=*extremely important*). The internal consistency for the 19 knowledge/competency items was $\alpha=0.93$, and the 19 importance items was $\alpha=0.95$.

T-tests with an alpha level of .05 were used to compare the differences between the ratings with respect to the level of importance of having knowledge/competency to determine if there were significant differences. If the average value of knowledge/competency is larger than the average value of importance (and the p-value is $< .05$) this indicates that respondents rated the level of importance as low when they thought they were more knowledgeable or competent.

The opposite is also true, if the importance average value is larger than the knowledge/competency average value (and the p-value is < .05) this suggests that respondents considered an item more important when they were less knowledgeable. To determine the magnitude of the differences found, Cohen's d effect sizes were calculated and interpreted as follows: 0.2= small, 0.5= moderate, 0.8= large.

There were statistically significant differences found between 15 of the 19 items across general measurement, screening, assessment, and evaluation as shown in Tables 2, 3, 4, and 5. However, the degree of difference when considering effect size calculations indicates that five of the differences were small (items 2, 4, 8, 10 and 12), four were moderate (items 1, 9, 11 and 13), and six were large (items 3, 5, 7, 15, 17 and 19) across all domains of measurement. With respect to some questions, respondents stated that they were *knowledgeable* (4 point response) in a particular measurement area and they considered it *somewhat important* (3 point response) that they were knowledgeable/competent for their position. In other questions, respondents stated that they were *somewhat knowledgeable* (3 point response) in a particular area and they believed that it was *important* (4 point response) that they were knowledgeable/competent for their position.

For example, the results of items 1, 3, 5, and 7 in Table 2 are notable because they were statistically significant and possessed moderate to large effect sizes. In these questions the difference between level of knowledge and importance suggests that the less knowledgeable/competent respondents considered themselves to be in general measurement, the more important they thought it was for their position. Conversely, the results of items 2 and 4 on Table 2, while statistically significant demonstrate small differences with respect to the more knowledgeable/competent they were, the less important they thought it was for their position. Table 5 illustrates the level of knowledge/competency and importance with respect to evaluation, and the three items (15, 17 and 19) that were found to be statistically significant also possessed large differences suggesting that they feel less knowledgeable about evaluation but view it as very important. These results provide some insights into potential areas of training for EC professionals. Some of the questions that EC professionals indicated as important but with lower levels of competence/knowledge were around areas that might be out of their scope of practice (e.g., cross-cultural considerations, connecting tools to purpose, and utilizing a variety of screening and assessment tools). On the other hand, the results may be indicative of areas that the participants feel relatively confident in because it is part of their scope of practice, there are clear organizational policies and procedures in place, and they are already receiving ongoing training and mentorship.

TABLE 2
Knowledge/competency and importance of knowledge/competency in the specific area
for your position re General Measurement

Question: Please indicate your knowledge and/or competency of the following questions and then rate
the importance of having knowledge/competency in the specific area for your position.

	Knowledge/ Competency		Importance		
	N	Mean	N	Mean	p
1. Gathering information across multiple areas of development (cognitive, behavioral, physical)	130	4.17	133	3.80	*0.00
2. Selecting Measurement tools based upon purpose of gathering information	131	3.66	132	4.04	*0.00
3. Selecting Measurement tools based upon quality of the tool (i.e. reliability and validity)	128	4.01	131	3.34	*0.00
4. Selecting, adapting, and administering Measurement tools and procedures for children with special/different needs.	131	3.62	131	3.85	*0.04
5. Using Assessment to identify children's special needs	128	4.16	132	3.35	*0.00
6. Understanding of cultural differences when using measurement tools	131	3.59	131	3.77	0.10
7. Understanding of screening, assessment, diagnosis, evaluation and the differences between them	128	4.14	130	3.20	*0.00

Notes: 1: Not at all Knowledgeable -- 5: Extremely knowledgeable/competent; 1: Not at all important – 5:
Extremely important
* Significant at the .05 level

TABLE 3
Knowledge/competency and importance of knowledge/competency in the specific area
for your position re Screening

Question: Please indicate your knowledge and/or competency of the following questions and then rate
the importance of having knowledge/competency in the specific area for your position

	Knowledge/Competency		Importance		
	N	Mean	N	Mean	p
8. Administering a variety of screening tools	129	3.60	130	3.89	*0.01
9. Scoring a variety of screening tools	128	3.91	132	3.30	*0.00
10. Interpreting a variety of screening tools	129	3.58	130	3.85	*0.05

Notes: 1: Not at all Knowledgeable -- 5: Extremely knowledgeable/competent
          1: Not at all important – 5: Extremely important
* Significant at the 0.05 level

TABLE 4
Knowledge/competency and importance of knowledge/competency in the specific area
for your position re Assessment

Question: Please indicate your knowledge and/or competency of the following questions and then rate
the importance of having knowledge/competency in the specific area for your position

	Knowledge/Competency		Importance		
	N	Mean	N	Mean	p
11. Administering a variety of assessment tools	127	4.01	130	3.38	*0.00
12. Scoring a variety of assessment tools	129	3.57	130	3.97	*0.00
13. Interpreting a variety of assessment tools	128	4.05	131	3.50	*0.00

Notes: 1: Not at all Knowledgeable -- 5: Extremely knowledgeable/competent
        1: Not at all important – 5: Extremely important
* Significant at the 0.05 level

Table 5
Knowledge/competency and importance of knowledge/competency in the specific area
for your position re Evaluation

Question: Please indicate your knowledge and/or competency of the following questions and then rate
the importance of having knowledge/competency in the specific area for your position

	Knowledge/ Competency		Importance		
	N	Mean	N	Mean	p
14. Administering a variety of evaluation tools and procedures	130	3.78	130	3.80	0.93
15. Scoring a variety of evaluation tools and procedures	129	4.20	130	3.20	*0.00
16. Interpreting a variety of evaluation tools and procedures	131	3.76	129	3.64	0.24
17. Monitoring and evaluation of program effectiveness in meeting the needs of children and families.	129	4.36	130	3.38	*0.00
18. Evaluating program effectiveness related to children's outcomes.	131	3.78	128	3.87	0.43
19. Using measurement tools to monitor and evaluate programming	128	4.13	131	3.20	*0.00

Notes: 1: Not at all Knowledgeable -- 5: Extremely knowledgeable/competent
        1: Not at all important – 5: Extremely important
* Significant at the 0.05 level

Purpose for measurement and modification of procedures.   Three questions asked respondents to indicate the main purposes for engaging in screening, assessment, and evaluation. The same response options were available for each area of measurement and Table 6 presents the three most frequently reported purposes. As respondents could select all responses that applied to them, internal consistency was not calculated. In the area of screening, *identify special*

needs/disability, identify children at-risk, and *support learning* were the top three purposes listed. *Support learning* was also identified in the top three with respect to assessment and evaluation. *Communicate progress* was indicated as the main purpose for both assessment and evaluation. Additionally, respondents listed *access to funding* as a main purpose of assessment, and *program evaluation* as a main purpose of evaluation.

TABLE 6
Three most frequently reported purposes for screening, assessment, and evaluation

	Main purposes	N	Percentage
	Screening		
1	Identify special needs/disability	107	67.3%
2	Identify children at-risk	101	63.5%
3	Support learning	85	53.5%
	Assessment		
1	Support learning	97	61.0%
2	Communicate progress	92	57.9%
3	Access funding	90	56.6%
	Evaluation		
1	Program evaluation	87	54.7%
2	Support learning	82	51.6%
3	Communicate progress	74	46.5%

Respondents were also asked whether or not they modify measurement procedures to accommodate children with disabilities or from diverse cultural backgrounds. On a 4-point likert scale (never, sometimes, often, and almost always), the majority (93.9%) of respondents reported that they modify measurement procedures (see Table 7). Only 6.1% reported that they never make modifications.

TABLE 7
Modification of measurement procedures to accommodate children with disabilities and/or children from diverse cultural backgrounds.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Never	8	5.0	6.1	6.1
	Sometimes	43	27.0	32.8	38.9
	Often	38	23.9	29.0	67.9
	Almost always	42	26.4	32.1	100.0
	Total	131	82.4	100.0	
	Missing data	28	17.6		
Total		159	100.0		

Measurement Issues and Needs of EC Professionals

The third section of the questionnaire consisted of 29 items across three areas that asked about important factors in selecting measurement tools, level of agreement with statements about measurement, and interest in learning more about measurement tools. The overall internal consistency for the 29 items was α=.80. Two additional open-ended questions were presented to explore other issues or challenge that individuals face when engaging in early childhood measurement and what services and resources would be beneficial.

Choosing measurement tools. Respondents were asked to rate the level of importance of 12 factors in their choice of measurement tools from *not at all important* to *extremely important*. Between 56% to 70% of respondents rated the following as extremely important: *reliability, validity, suitability for children with a variety of needs, provides insights into children's strengths and needs, assists in identifying children at-risk,* and *easy to explain to parents.* For the most part, respondents considered all factors important or extremely important with the exception of *input/requirements from funders.* Only a small percentage of respondents reported that the following factors were either *not at all* or *not very important* in their selection of a tool; *ease of use, cost, personal knowledge of the tool, input/requirements from funders,* and *administration procedures are flexible.*

Training. Eighty percent of respondents (n=127) answered questions related to whether or not they agreed with statements that staff are properly trained and prepared to conduct and interpret results from screening, assessment or evaluation tools. The response options ranged from *strongly disagree* to *strongly agree* with a *neutral* option. The frequency of results was aggregated and the average value across each measurement area was as follows: 3.33 for screening, 3.22 for assessment, and 3.15 for evaluation. These results suggest that respondents neither agreed nor disagreed as they selected the *neutral* response most often.

Participants were also asked to rate their level of interest in learning more about the use and administration of measurement tools, how to interpret results and share information with parents, how to link results to programming for children, how to use tools with diverse cultural groups, and properties of the tests such as reliability and validity. Approximately 80% of respondents indicated that they were either *interested* or *extremely interested* in learning more about all the topics. In particular, respondents were *extremely interested* in learning more about how to link results to programming and instruction, how to interpret results and share the information with parents, and how to use tools with culturally diverse populations.

Other issues and needed resources/services. Two open-ended questions were provided to determine if respondents wanted to highlight any additional measurement issues that they are dealing with in their practice as well as recommend measurement resources or services that would benefit their practice. Two researchers reviewed the data to identify recurring phrases and concepts, and to subsequently group these recurring phrases and concepts into categories (Mayan, 2009). From this process, two themes emerged: (1) a need for measurement standards across all programs that are delivering services for young children; and (2) a need for Program Directors to be more knowledgeable in the area of measurement and evaluation, and more specifically knowledgeable about identifying children with special needs or at-risk and linking

this information to programs and services. Some participants commented on a lack of resources within their organization and within the metropolitan area for addressing issues related to early childhood measurement. For example one respondent stated that there is "not enough support staff to assist teachers in following through with the results of assessments." Other suggested resources and services were easy access to specific measurement tools accompanied by training to provide knowledge on how to identify delays, information on evaluation and how it relates to curriculum development, information on the difference between screening and assessment, and interpretation of screening tools, and finally information on appropriate referrals and follow-up procedures.

Measurement Tools Used in Practice

Respondents had five open ended questions to specify the names of screening, assessment, and evaluation tools they use within their practice with children aged birth to six years. This was followed by four closed questions, which asked about the amount of training participants had on the tool, the best description of the tools, the purposes for using the measurement tool, and finally the uses of the information gathered.

Measurement tools used. The results revealed a large diversity in tools used by respondents. Across all respondents, 52 unique tools were identified and the 11 most frequent are listed in Table 8.

TABLE 8
Most frequently cited measurement tools used

Measurement tool	Rank	Frequency
Nipissing District Developmental Screen (NDDS)	1	36
The Brigance Inventory of Early Development (Brigance)	2	24
Diagnostic Inventory for Screening Children (DISC)/DPS	3	17
Ages and Stages Questionnaires (ASQ)	4	8
Clinical Evaluation of Language Fundamentals – Preschool second edition (CELF-P2)	5	8
Preschool Language Scales – Fourth Edition (PLS-4)	6	9
Developmental Indicators for the Assessment of Learning (DIAL)	7	5

The data was further refined and tools were recoded into four categories: screening, assessment, observation, and curriculum-based. Forty percent of participants were using screening tools, 31% used observational tools, 22% used assessment tools, and 7% were using curriculum-based tools. It appeared more common for participants to report the use of one tool in their practice (20%) than two (13.8%), three (12.5%), four (6.9%) or more than five (8.1%).

Amount of training. Participants were asked to report the level of training they received for each tool that they listed and the response options ranged from *no training* to *university courses.* The research team received over 240 responses to this question because it

was repeated for each unique tool a respondent reported. Across the 240 responses, 61% (n=146) of participants reported *no training* or *1 day/in-house training*. Due to small sample sizes within each cell, significance tests could not be conducted to determine if there was a significant difference between tool type and level of training.

DISCUSSION

The purpose of this research was to describe EC professional's knowledge and competency related to early childhood measurement, measurement issues and needs, and current measurement practices. Teachers and Program Directors participated in this research to a higher degree than others working in the early childhood development field. Perhaps this is an indication, at least in the metropolitan area where the study was conducted, of who has a greater comfort level in reporting on their knowledge and competency in early childhood measurement. Over half of the respondents reported that they serve populations of children that are typically developing, from diverse cultural backgrounds (i.e., immigrant, refugee, and Aboriginal) with English as a second language, children with disabilities, and children exposed to at-risk conditions (e.g., poverty). Although the sample of survey participants (n=159) was one of convenience due to the method of distribution and response rate, given the breadth of populations they serve the information from this study has relevance to many EC professionals such as head start, childcare and family day home providers, early intervention, and early educators. The limited participation specifically from childcare professionals is not clear but could be as a result of the lower involvement of frontline childcare professionals in screening, assessment, and evaluation. To understand this more fully, perhaps more targeted strategies need to be employed in the future to engage this population of professionals such as a series of focus groups or one-on-one interviews.

The majority of respondents were educated at the post secondary level and had many years of experience in the early childhood development field. One might argue that the professional development needs of these respondents might be less than for others working in this field. However, the findings revealed that respondents had a high interest in learning more about screening, assessment and evaluation and for the most part had received little to no training for the current tools and procedures that they were using.

It was interesting to find that respondents who indicated that they were *knowledgeable* in the area of measurement generally rated it as only *somewhat important* that they were knowledgeable/competent for their position. This seems contradictory; when someone views their knowledge and competency in a specific area as high, one might also expect them to rate importance of having knowledge and competency as high. While this finding may seem counterintuitive to some, it may suggest that professional development efforts should be planned and implemented in close collaboration with EC professionals to address areas considered to be of high importance. For example, some of the areas indicated as important (e.g., cross cultural knowledge, utilizing a variety of measurement tools, and selecting tools based on purpose) might currently be out of the scope of practice of the EC professionals in this study or inconsistent with the needs or context of the populations of children they serve. However, there is a clear indication that these are areas that individuals may want to increase both their knowledge and competence. The findings highlight a few specific areas of importance to support existing

knowledge and build measurement literacy among EC professionals. The following areas could be used as starting points in planning training opportunities:

- Increasing understanding of the difference between screening, assessment, and evaluation and the purposes of each measurement strategy;
- The importance of gathering information across multiple areas of development (cognitive, behavioural, and physical domains);
- How to select measurement tools based upon the quality of the tool (e.g., reliability and validity);
- Using assessment tools and procedures to identify children's special needs;
- Ways to link assessment results to programming and instruction;
- Methods for interpreting results and sharing the information with parents; and
- Using measurement tools with culturally diverse populations.

A little under half of respondents indicated they provide screening as a regular service and the most frequently reported purpose of screening was to identify special needs/disability and this is not considered a best practice (Rydz et al., 2005). This finding suggests that respondents lack knowledge about the purposes of screening and the limitations of screening. Also, over half of respondents indicated that they provide assessment, and a*ccess to funding* was indicated as a main purpose for assessment. The outcome of an assessment might result in funding to provide supports or services for children; however, this is not a purpose of assessment that reflects best practice (AERA, APA, & NCME, 1999). With accountability frameworks in place, assessment outcome data appears to fulfill an important role in the provision of funding and programming for children (Administration for Children and Families, 2002). This finding supports the need for professional development to assist EC professionals with appropriate methods for presenting relevant data to existing or potential funders. Evaluation was broadly defined on the questionnaire and therefore it is not surprising that over half of respondents identified the main purpose was for program evaluation. Respondents were not asked to stipulate whether evaluation was an activity regularly conducted therefore little information is available with respect to knowledge and competency in this area. Program evaluation may be a beneficial area for professional development, as it would support curriculum and programming enhancement and provide evidence for external funding proposals.

Another area that particularly stands out in the findings of this study is the common practice among respondents to modify measurement procedures. The study did not address the type of modifications made; however, it is well known that existing standardized tests can be problematic for EC professionals. For example, the content may not reflect what is valued and relevant to the culture or context of the individuals being assessed, as tests are often normed on Caucasian, middle-class children (Fuchs, Fuchs, Benowitz, & Barringer, 1987; Meisels & Atkins-Burnett, 2000). Lyman, Njoroge, & Willis (2007) found that a number of standardized assessments do not reflect the diversity of Canada's population in terms of exceptionalities and culture. Other issues with respect to standardized tests are that some are inappropriate for special populations resulting, for example, in bias against individuals with low socio-economic status and that there is little linkage to programs, services and interventions (Fuchs, et al., 1987; Meisels & Atkins-Burnett, 2000). This leaves EC professionals with two options - they can use existing tools that may not reflect the individual or group being assessed or they can modify the testing procedure to reflect the individual or group (Appl, 2000). Either option is fraught with limitations, as modifications may provide a better picture of how a child is performing there is no

way of knowing what impact the modifications or accommodations have on the integrity of the tool and therefore the validity of the results obtained. A third option is to complement conventional assessment with authentic assessment which relies heavily on context and evidence that is gathered on an ongoing basis through natural settings, and can be used to align with early learning standards (Bagnato, 2005; Grisham-Brown, Hallam, & Brookshire, 2006; Meisels, Wen, & Beachy-Quick, 2010). This finding regarding the modification of measurement tools and lack of knowledge of alternatives points to an important area for training and development of EC professionals. In addition to training, it is recommended that EC professionals seek the expertise or advice of researchers or measurement experts at local universities when considering modification or selection of measurement tools to meet their needs. These connections have the potential to lead to research partnerships that engage the expertise of academic researchers with the expertise of EC professionals in examining the relevancy and appropriateness of tools for use with certain populations (e.g. Gokiert, Georgis, Chow, & Chui, 2012).

Respondents identified that the most important factors in choosing a measurement tool were ease of use, reliability and validity, and these are consistent with findings from previous studies (Brown & Rolfe, 2005; Johnson & Beauchamp, 1987; Pretti-Frontczak et al., 2002). The top three measurement tools utilized were the Nipissing District Developmental Screen (NDDS), The Brigance Inventory of Early Development (Brigance, 2004), and Observation Survey of Early Literacy Achievement (Clay, 2002, 2006). These findings show that participants utilize screening, assessment, and observational tools that appear to fit with their level of training. The screening tools are intended to be user-friendly, for instance, the NDDS is completed by a parent and can easily be scored and compared to cut-off points to determine if a child is experiencing difficulty in a particular developmental domain. These types of early developmental screening tools require less training to administer, score, and interpret. Measurement tools listed as being less frequently used by participants, such as the Developmental Indicators for the Assessment of Learning (DIAL-3; Mardell & Goldenberg, 1998) and the Preschool Language Scale – Fourth Edition (PLS-4; Zimmerman, Steiner, & Pond, 2002), are standardized assessment tools. Not only are these tools more specific to certain developmental domains, they also require specialized training.

The use of observational tools is consistent with previous findings that suggest early childhood professionals use checklists because they are quick and simple to use and usually only require selecting behaviours or skills as they occur (Brown & Rolfe, 2005). It is not surprising that the Observation Survey of Early Literacy Achievement was reported as a vast number of respondents were teachers and this aligns with current early literacy initiatives within the metropolitan area that the study took place. Furthermore, the use of curriculum-based measures is not surprising given that the majority of participants were teachers and they typically need to link skills and outcomes to curricular objectives. On average, 42% of the participants reported using more than one type of tool in their practice, which is consistent with previous findings (Johnson & Beauchamp, 1987; Pretti-Frontczak et al., 2002) and in line with best practices (NAEYC, 2003).

Some participants commented on the need for measurement standards across all programs that are delivering services for young children. Presently, in most provinces across Canada there are no set standards for the types of early childhood screening, assessment, and evaluation developmental tools that should be used, how often they should be used, and for what purposes. There are some provincial standards that support the use of certain standardized assessment tools as they relate to coding and funding. The majority of respondents in this study

had post secondary education and many years of experience but well over half reported that they are administering measurement tools with little to no training. For this reason, there is clearly a need for ongoing and targeted professional development opportunities to support best practice for all who work in the early childhood field. training.

REFERENCES

Administration for Children and Families (2003). *Resources for measuring services and outcomes in Head Start programs serving infants and toddlers.* Washington, DC: U.S. Department of Health and Human Services.

AERA, APA, & NCME. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Anderson, L., Shinn, C., Fullilove, M., Scrimshaw, S., Fielding, J., Normans, J., & Carande-Kulis, V. (2003). The effectiveness of early childhood development programs: a systemic review. *American Journal of Preventive Medicine,* 24 (suppl 3), 32–46.

Allen, S. (2007). Assessing the Development of Young Children in Child Care: A survey of Formal Assessment Practices in One State. *Early Childhood Education Journal,* 34(6), 455-465.

Appl, D. (2000). Clarifying the preschool assessment process: Traditional practices and alternative approaches. *Early Childhood Education Journal*, 27, 219-225.

Bagnato, S.J. (2005). The authentic alternative for assessment in early intervention: An emerging evidence-based practice. *Journal of Early Intervention, 28*(1), 17-22.

Berry, D. J., Bridges, L. J., & Zaslow, M. J. (2004). *Early childhood measures profiles*. Washington, DC: Child Trends.

Bordignon, C. M., & Lam, T. C. M. (2004). The early assessment conundrum: Lessons from the past, implications for the future. *Psychology in the Schools, 41*(7), 737-749.

Bredekamp, S., & Copple, C. (Eds.). (1997). *Developmentally appropriate practice in early childhood programs* (Rev. Ed.). National Association for the Education of Young Children: Washington, DC.

Brigance, A. H. (2004). *Brigance Diagnostic Inventory of Early Development-Second Edition (IED-II).* N. Billerica, MA: Curriculum Associates.

Brown, J., & Rolfe, S. (2005). Use of child development assessment in early education: Early childhood practitioner and student attitudes toward formal and informal testing. *Early Child Development and Care*, 175, 193-202.

Campbell, F. A., Ramey, C. T. (1994). Effects of early intervention on intellectual and academic achievement: A follow-up study of children from low-income families. *Child Development,* 65, 684-698.

Clay, M. M. (2002, 2006). *An Observation Survey of Early Literacy Achievement*. Portsmouth, NH: Heinemann.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Dawson, G. & Osterling, J. (1997). Early intervention in autism: Effectiveness and common elements of current approaches. In Guralnick (Ed.) *The effectiveness of early intervention: Second generation research*. (pp. 307-326). Baltimore: Brookes.

Fuchs, D., Fuchs, L. S., Benowitz, S., & Barringer, K. (1987). Norm-referenced tests: Are they valid for use with handicapped students? *Exceptional Children*, 54 (31), 263-271.

Gokiert, R.J., Georgis, R., Chow, W., Chui, Y. (2012). Early childhood developmental screening: Does culture play a role? *Health Research Transfer in Alberta: Knowledge Translation Casebook, 3, 22-25.*

Gredler, G. (2000). Early childhood education – assessment and intervention: What the future holds. *Psychology in the Schools*, 37, 73-79.

Greenspan, S. I., & Meisels, S. J. (1996). Toward a new vision for the developmental assessment of infants and young children. In S. J. Meisels & E. Fenichel (Eds.), *New visions for the developmental assessment of infants and young children* (pp. 11-26). Washington, DC: Zero To Three National Center for Infants, Toddlers, and Families.

Greenwood, C. R., Luze, G. J., & Carta, J. J. (2002). Assessment of intervention results with infants and toddlers. In A. Thomas, & J. Grimes (Eds.), *Best practices in school psychology IV* (Vol.   2,   pp.1219-1230). Washington DC.

Grisham-Brown, J., Hallam, R., & Brookshire, R. (2006). Using authentic assessment to evidence children's progress toward early learning standards. *Early Childhood Education Journal, 34*(1), 45-51.

Hertzman, C. (1999). The Biological Embedding of Early Experience and Its Effects on Health   in   Adulthood. *Annals of the New York Academy of Sciences,* 896 (1), 85–95.

Hirsh-Pasek, K. Kochanoff, A., Newcombe, N., & de Villiers, J. (2005). Using scientific knowledge to inform preschool assessment: Making the case for "empirical validity". *Social Policy Report*, *19* (1), Society for Research in Child Development.

Horton, C., & Bowman, B. T. (2002). *Child assessment at the preprimary level: Expert opinion and state trends* (Occasional Paper No. 3). Chicago: Erikson Institute.

Johnson, L. & Beauchamp, K. (1987). Preschool assessment measures: What are teachers using? *Journal for the Division of Early Childhood*, 12 (1), 70-76.

Kienapple, K., Lyon, M., McSorely, A., Morse, P., & Smith, M. (2005). *Commonly used assessment tools for children with special needs ages birth to six and their families.* Report produced by The Early Intervention Task Force, Mount Saint Vincent University for *Centre for Excellence for Children and Adolescents with Special Needs. http://www.coespecialneeds.ca/PDF/assessmenttools.pdf.*

Lyman, D.R., Njoroge, W.F.M., & Willis, D.W. (2007). Early childhood psychosocial screening in culturally diverse populations: A survey of clinical experience with the Ages and Stages Questionnaires: Social-Emotional (ASQ:SE). *Zero to Three,* 27(5), 46-54.

Mardell, C. & Goldenberg, D.S. (1998). *Developmental Indicators for the Assessment of Learning - Third Edition (DIAL-3) Technical Information.* Pearson Assessments, Bloomington, MN.

Martella, J. (2004). The Words We Use: A Glossary of Terms for Early Childhood Education Standards and Assessments. www.ccsso.org/eceaglossary.

McCain, M.N., Mustard, F., & Shanker, S. (2007). *Early Years Study 2: Putting Science into Action.* Council for Early Childhood Development. Toronto, Canada.

McConnell, S. R. (2000). Assessment in early intervention and early childhood special education: Building on the past to project into our future. *Topics in Early Childhood Special Education*, 20, 43- 48.

Meisels, S. J., & Atkins-Burnett, S. (2000). The elements of early childhood assessment. In J. P. Shonkoff & S. J. Meisels (Eds.), *Handbook of early childhood intervention* (2nd ed. 231-257). New York: Cambridge University Press.

Meisels, S., & Fenichel, E. (Eds.). (1996). *New visions for the developmental assessment of infants and young children*. Zero to Three National Center for Infants, Toddlers, and Families: Washington, DC.

Meisels, S., Wen, X., & Beachy-Quick, K. (2010). Authentic assessment for infants and toddlers: exploring the reliability and validity of the ounce scale. *Applied Developmental Science, 14*(2), 55-71.

National Association for the Education of Young Children (NAEYC), & National Association of Early Childhood Specialists in State Departments of Education (NAECS/SDE) (2003). *Position Statement: Early childhood curriculum, assessment, and program evaluation: Building an effective, accountable system in programs for children birth through age 8*. Washington DC.

National Association for the Education of Young Children (NAEYC). (2009). *Standard for Early Childhood Professional Preparation Programs.* Retrieved January 28, 2011 from http://www.naeyc.org/positionstatements

National Association of School Psychologists. (2005). *Position Statement on Early Childhood Assessment*. Retrieved September 15, 2006 from http://www.nasponline.org/information/pospaper_eca.html

National Education Goals Panel. (1998). *Principles and recommendations for early childhood assessments*. Washington, DC: U.S. Government Printing Office.

Neisworth, J T., & Bagnato, S. J. (2000). Recommended practices in assessment. In S. Sandall, M. E. McLean, & B. J. Smith (Eds.), *DEC recommended practices in early intervention/early childhood special education* (pp. 17-27).

Nelson, C.A. (2000). The neurobiological bases of early intervention. In J. P. Shonkoff & S.J. Meisels (Eds.), *Handbook of early childhood intervention*, second edition (pp. 204-227). Cambridge University Press: Cambridge, MA.

Niemeyer, J., & C. Scott-Little. (2001). Assessing kindergarten children: A compendium of assessment instruments. Tallahassee, FL: SERVE. Retrieved January 14, 2011 www.serve.org/publications/rdakcc.pdf.

Pretti-Frontczak, K., Kowalksi, K., & Brown, R. D. (2002). Preschool teachers' use of assessment and curricula: A statewide examination. *Exceptional Children*, *69*,109-123.

Ramey, C.T., & Ramey, S.L. (2004). Early learning and school readiness: Can early intervention make a difference? *Merrill-Palmer Quarterly,* 50, 471-491.

Rydz, D., Shevell, M.I., Majnemer, A., & Oskoui, M. (2005). Developmental screening. *Journal of Child Neurology, 20*, 4-21.

Sandall, S., McLean, M. E., & Smith, B. J. (2000). DEC recommended practices in early intervention/early childhood special education. Longmont, CO: Sopris West.

Schweinhart, L.J., Barnes, H.V., & Weikart, D.P. (1993). *Significant benefits: The high/ scope perry preschool study through age 27*. Ypsilanti, MI: High/Scope Press.

Shepard, L. (1994). The challenge of assessing young children appropriately. *Phi Delta Kappan*, 76, 206-211.

Shepard, L.A. (1997). Children not ready to learn? The invalidity of school readiness testing. *Psychology in the Schools*, 34, 85-97.

Shepard, L., Kagan, S., & Wurtz, E. (1998) *Principles and recommendations for early childhood assessments.* Washington DC: National Education Goals Panel. September 3, 2010 http://govinfo.library.unt.edu/negp/Reports/prinrec.pdf

Shonkoff, J., & Phillips, D. (2000). *From neurons to neighborhoods: The science of early childhood development.* Washington, DC: National Academy Press.

U.S. Department of Health and Human Services (2007). Head Start Act.

VanLandeghem, K., Curgins, D., & Abrams, M. (2002). *Reasons and strategies for strengthening childhood development services in the healthcare system.* Portland, ME: National Academy for State Health Policy.

Zimmerman, I.L., Steiner, V.G., & Pond, R.E. (2002). *Preschool Language Scale, Fourth Edition (PLS-4).* Pearson Assessment, Bloomington, MN.