Cognitive interviews with children as a research tool for instrument validation in adapted physical activity

Nancy L. I. Spencer[1]*, Marcel Bouffard[1] and E. Jane Watkinson[1]

Abstract: The development and use of instruments to assess individuals with impairments in various domains is common practice in the field of Adapted Physical Activity. Test developers and users often question the validity of their instruments and use different conceptual and/or empirical strategies for validation purposes. One validation strategy, still rarely used in the sport sciences, is cognitive interviews with participants. This study is an attempt to show the utility of cognitive interviews for instrument development with children with relevance to the field of adapted physical activity. Specifically, we investigated the question-and-answer processes of children with impairments when responding to the Athletic Competence Domain Subscale from Harter's (1985) Self-Perception Profile for Children. Eight children with different diagnoses (ages 8-13 years) took part in cognitive interviews. The study revealed sources of validity and invalidity in the instrument. Differences and concerns in the children's comprehension of questions and interpretation of words, leading to potentially limited and varied sources of information for response production, emerged. While judgment generally appeared to be unproblematic, topic sensitivity and limited response options were a common constraint. The data obtained from this study could be used to further refine the instrument.

Keywords: perceived competence; questionnaire research; validity

Introduction

The development and use of instruments to assess individuals in different domains such as motor skill and development, behavior, social competence, attitudes, and self-concept, among others (Horvat, Block, & Kelly, 2007), is common practice in the field of Adapted Physical Activity (APA). As an area of kinesiology, APA researchers rely on parent disciplines to realize their knowledge generation goals (Bouffard & Spencer-Cavaliere, 2016), which often includes performing different types of assessments. For APA researchers who engage questions concentrated on the self, there is a heavy reliance on the parent discipline of psychology and its various branches that are dedicated to understanding this area. The growth of interest in the area of the self is evidenced in the development of self-report instruments designed to measure constructs such as self-esteem, self-worth, and self-efficacy (Harter, 1999). Of these instruments designed to measure individual self-processes, several have also been developed and modified for use with children with and without impairments.

The use of self-report instruments to gain information about children's perceptions, thoughts, and feelings is becoming increasingly more common. Within the development of these types of instruments, a shared theme is producing and reporting validity evidence. Building on the work of Dietrich and Ehrlenspiel (2010), in this paper we contribute to the argument, that among the

evidences provided, a key source of validation, namely the investigation of cognitive processes in item response, is often neglected. In doing so, we highlight the relevance of the question-and-answer model to instrument development, and offer an example of how cognitive interviews may be used to contribute to the validation of a self-report instrument designed to measure perceived athletic competence with children with impairments.

Cognitive Processes in Test Response

Understanding the response production process of respondents is essential to the instrument validation process. At the present time, the development of theories of response production processes are in their infancy. We do, however, have models of response production such as the one advanced by Tourangeau (1984)(Tourangeau, Rips, & Rasinski, 2000) that could be used to gain a better understanding of the response production process. Models about response processes have been available for a long time but are rarely used in the sport sciences (Dietrich & Ehrlenspiel, 2010) and by extension, the field of APA. These models have the potential to offer validity evidence for the development of instruments and their adaptation for use with diverse populations (e.g., children with impairments).

The questionnaire as a method of data collection is typically administered with a standardized set of questions, in a fixed order, with fixed response options (Groves et al., 2009). Underlying the traditional standardization of testing instruments are assumptions that all respondents understand the questions and understand them in the same way, consistent with the intentions of the researcher (Beatty & Willis, 2007; Collins, 2015). According to Tourangeau's (1984) question-and-answer model developed in cognitive psychology, respondents complete four actions, comprehension, retrieval, judgment, and response, in order to answer a question. As described by Tourangeau et al., (2000), comprehension involves attending to and understanding the question. It may be compromised of such things as grammar and word meanings that are complex or vague and by past experiences. Retrieval relies heavily on the respondent's memory and ability to access the relevant information to answer the question. Judgment involves formulating an answer and depends on the ability of the respondent to integrate the retrieved information and to draw inferences based on this retrieval. Finally, response requires mapping of the judgment by formatting and editing the answer onto a response category (Tourangeau et al., 2000). Despite efforts to standardize tests and procedures used to collect information, errors in measurement continue to occur (Collins, 2015). Understanding what occurs between the moment a question is asked and the time it is answered is essential for validity. It can inform whether or not questions are understood in line with the researcher's intention and can provide information about the appropriate use of instruments in various disciplines, including the field of APA.

Cognitive Interviewing

Cognitive interviewing applied to questionnaire development is a verbal report method founded in cognitive psychology and has been used to obtain information about peoples' cognitive processes in response to test items and to enhance understanding of what those responses mean (Beatty & Willis, 2007; Knafl et al., 2007; Willis, 2005, 2015). It is increasingly used as a pretest for survey instruments to examine the processes employed by respondents, as reported by them, when answering questions (Beatty & Willis, 2007). Cognitive interviews help to identify issues underlying the cognitive processes used by respondents in generating responses to questionnaire items (Willis, 2015). The results of cognitive interviews may provide information about what a response means, expose errors in item interpretation and response, as well as providing direction for future item development and revision (Collins, 2015; Willis, 2015).

In general, cognitive interviewing may involve eliciting verbal responses to questions about past events, information retrieval, and question response using different methods. These methods typically consist of a think-aloud procedure where-in respondents are asked to verbalize their thoughts as they

respond to questions and or verbal probing which involves asking questions designed to elicit information about the cognitive processes involved in response (Beatty & Willis, 2007; Willis, 2005,2015). Cognitive interviews are a critical step in the instrument development process (Willis, 2015) and align well with Tourangeau's (1984) question-and-answer model given the nature of the questions posed and the four actions associated with the model (i.e., comprehension, retrieval, judgment and response). Cognitive interviews, however, are rarely considered in the accumulation of validity evidence in psychological testing (Castillo-Díaz & Padilla, 2013). This criticism extends to the development of measurement instruments within the sport sciences (Dietrich & Ehrlenspiel, 2010; for examples see Neilson, Ullman, Robson, Friedenreich, & Csizmadi, 2013; Ullrich-French, Cox, Cole, Roades Cooper, & Gotch, 2017) and to the field of APA.

Purpose

The purpose of this paper was to explore the cognitive response processes of children with impairments through the use of cognitive interviews. We focused on the Athletic Competence Domain Subscale (ACDS) of Harter's (1985) Self-Perception Profile for Children (SPPC), a self-report instrument designed to measure children's perceptions of competence. Along with its variants, this instrument has been used in the field of APA (e.g., Shapiro, Moffett, Lieberman, & Dummer, 2005; Yun & Ulrich, 1997). Our hope was to further contribute to the development of this instrument for use with children with impairments and in doing so, highlight the relevance of the question-and-answer model.

Harter's Theory of Competence Motivation

Numerous instruments have been developed to gain a better understanding of self-concept (for examples see Harter, 1990), which is an individual's assessment of themselves and is comprised of "evaluative judgments of attributes within discrete domains" (Harter, 1999, p.5). The SPPC (Harter, 1985) was created for use with children, is in keeping with a multidimensional approach to self-concept, and includes the domains of athletic competence, behavioral conduct, scholastic competence, physical appearance, and social acceptance as well as a global self-worth subscale. It emerged out of Harter's cognitive developmental theory of competence motivation (1978, 1981a). According to Harter's (1978, 1981a, 1999) theory, motivation is associated with children's perceptions of competence which are influenced by three primary factors: (a) experiences of success or failure, (b) perceived control over outcomes, and (c) feedback from socializing agents and the affective responses generated by them. As children develop, the different sources of competence information they attend to changes. At around the age of eight years old, children begin to integrate both positive and negative representations of the self, meaning they can perceive themselves to be both nice and mean or smart and dumb, whereas in early childhood, these opposites do not coexist (Harter, 1999). In the athletic domain, high perceptions of competence, for example, in wheelchair basketball are predicted to lead to greater motivation to participate in that sport, whereas low perceptions of competence are more likely to result in withdrawal.

A number of scholars have examined the perceptions of children with different impairments with regard to their athletic (i.e., physical and motor) competence in various contexts. The terms perceptions of athletic competence, motor competence, and physical competence used here refer to the same concept. Shapiro and Ulrich (2002) did a comparative study examining expectancies, values, and perceptions of physical competence of children with and without learning disabilities across the contexts of physical education, recess, and home. Perceived physical competence has also been examined in children with visual impairments in physical activity (Brian, Haegele, & Bostic, 2016), sports camp (Brian, Taunton, Haibach-Beach, & Liberman, 2018), and within the concept of physical literacy in physical education (Brian et al., 2019). Likewise, researchers have explored perceived motor competence of children with cerebral palsy in school (Schuengel et al., 2006), perceptions of physical competence of children with movement difficulties in physical education (Causgrove Dunn & Dunn,

2006), as well as with children with varied physical impairments across different motivational climates (Tsalavoutas & Reid, 2006). While there are some mixed results, a premise and common rationale for this type of work is that children with impairments are more likely to have lower perceptions of athletic competence which can have negative implications for their motivation to take part in physical activity.   For children with impairments, who are less likely to have positive sport and activity experiences and more likely to experience exclusion in activity settings (Spencer-Cavaliere & Watkinson, 2010), developing a strong sense of perceived athletic competence is essential to continued participation and opportunities for skill development. Therefore, a strong understanding of their perceived competence in physical activities settings is critical. Harter's theory, the SPPC-ACDS, and its various iterations (e.g., for use with young children and adolescents) have been used in the field of APA to understand motivation and specifically perceptions of athletic competence for children and youth with different impairments. However, despite a rigorous approach to instrument development as described below, evidence of the use of cognitive interviews for instrument development is lacking.

Materials and Methods

The approach used in our study was based on Tourangeau's (1984) question-and-answer model and relied on cognitive interviews. Although this approach is qualitative in nature and produces qualitative data, this method does not fit within traditional qualitative research, in that for example, there are no stated research assumptions underlying the specific model or cognitive interviews. As researchers, however, we articulate our own ontological and epistemological world views which align with a realist ontology and critical realism, respectively. We view reality as independent of perception but that our access to it and therefore knowledge of it is conceptually mediated (see Danermark, Ekström, Jakobsen, & Karsson, 2002 for further details).

Participants

A purposeful sampling strategy was used to identify children who would be information-rich cases about the concerns of key importance to the purposes of this study (Beatty & Willis, 2007; Patton, 2002). Children were recruited from two nonprofit, volunteer based, specialized sport and physical activity organizations for children with impairments. Within these programs children took part in a range of adapted physical activities such as tae kwon do, swimming, soccer, and sledge hockey. The eight children in this study, 7 boys and 1 girl, between the ages of 8 and 13 years old (one child turned 13 five days prior to the interview), had a range of diagnoses including: cerebral palsy, obstetrical brachial plexus injury, and developmental coordination disorder. None of the children had an intellectual impairment. Although the sample size was small, as Willis (2015) notes, small sample sizes are not uncommon in cognitive interviewing. Furthermore, the diversity of our sample is representative of the heterogeneity often found within APA settings (Spencer-Cavaliere, Thai, & Kingsley, 2017). Approval for this study was granted by a University Research Ethics Board and only children for whom informed consent was provided by parents, participated. Children also provided assent prior to taking part in the study.

Data Collection

Self-Perception Profile for Children – Athletic Competence Domain Subscale

The SPPC (Harter, 1985) is a revision of the Perceived Competence Scale for Children (PCSC; Harter, 1982). For more details on the development of the PCSC as foundational to the SPPC please see Harter (1981a, 1982). The ACDS is one of the subscales within the SPPC and was designed to elicit children's perceptions of their athletic ability with items tapping relevant sports and outdoor games content. It consists of six items with statements corresponding to high and low competence. Reported internal consistency reliability for the ACDS ranges from .80 to .86 and factorial validity (i.e., factor

patterns through oblique rotation) was demonstrated with the average loading of items ranging from .41 to .81 (Harter, 1985).

The SPPC uses a structured alternative format which was created to address issues arising from social desirability. In this format, children can identify with descriptions of other children who are most like them (e.g., good or bad at something) and whether or not the description is sort of true or really true for them. Identifying with other like children is intended to legitimize the child's selection thereby reducing social desirability errors. Items are scored on a 4-point scale with 1 indicating low and 4 indicating high perceived competence (Harter, 1981a, 1982). The effectiveness of this format was demonstrated through confirmatory verbal elaborations by children (Harter, 1981b), low correlations with a social desirability scale, and evidence that children used the full range of scores (see Harter, 1982 for more details).

Harter (1985) cautioned against the use of the SPPC with populations for whom the instrument was not designed or tested. For example, she indicated that the SPPC is not appropriate for use with children below third grade or eight years of age. Harter (1985) also suggested the SPPC may or may not be appropriate for special populations (i.e., children with impairments). Silon and Harter (1985) examined the appropriateness of the PCSC (predecessor to the SPPC) for use with children with intellectual impairments. In addition to examining psychometric properties, these authors also performed brief interviews to learn about the sources of comparison used by participants in forming self-evaluative judgments. Similarly, Renick and Harter (1989) incorporated brief interviews with the children with learning impairments in their study, in order to determine which social comparison group children used when responding to items. There is also reported evidence of interviews with children in the initial development of the PCSC and the structured alternative format (Harter, 1981a, 1981b, 1982). Despite the attention paid to the development of the SPPC (which includes the PCSC) evidence is lacking to support a detailed examination of children's response production processes (i.e., cognitive interviews). The asking of follow-up questions about how individuals understand and respond to items is critical to measurement. It can provide key information about whether or not respondents understand questions in line with the intentions of the test developer (Collins, 2015), as a pretesting approach to instrument development (Willis, 2005, 2015), and in exploring their use with other populations.

Interviews

Cognitive interviews, as previously described, were used to access children's response processes. There is evidence to support the relevance and usefulness of cognitive interviews with children in the development of self-report instruments (e.g., Leary, Ice & Cottrell, 2012; Woolley, Bowen, & Bowen, 2004). Cognitive interviews were used along with the SPPC-ACDS (Harter, 1985). Interviews took place in a private room at the University or in the homes of participants in keeping with parents' requests. A parent was present for two of the interviews. Only the child and the interviewer (first author) were present for the other six interviews. At the beginning of the interview each child was instructed how to answer the items of the SPPC-ACDS according to Harter's (1985) recommendations and completed a practice question. An interview guide (see Appendix) consisting of verbal probes to elicit information about comprehension, retrieval, judgment and response was developed and paired with items from the SPPC-ACDS to guide semi-structured interviews using cognitive interviewing techniques. Verbal probing, as recommended for use with children (Willis, 2005), was used concurrently as children filled out each questionnaire item. Interview strategies and questions were confirmed with another researcher and two pilot studies took place to ensure the probing questions were clear and the interviewing techniques were properly implemented. Based on the two pilot studies, one with a child with an impairment and one without an impairment, which found that children tired with many probes for the same item, the number of verbal probes was reduced to a maximum of two per statement and systematically varied across the items. Interviews took between twenty and forty-five minutes and were audio taped.

Field Notes

During the interviews, the interviewer/first author took brief field notes using shorthand and symbols to denote particular areas of interest. This was done in order to maintain the flow of the interviews and to keep the children engaged. Following each interview, she then elaborated on these brief notes and documented her impressions of the children, their responses to the instrument questions and probes, and initial thoughts on the question-and-answer model categories.

Analysis

The data for analysis were comprised of the interviews, which were transcribed verbatim and supplemented with the interviewer's field notes. These sources of data were integrated to confirm, expand, and challenge the analysis in order to be as comprehensive as possible (Willis, 2015). While the numerical data from the SPPC-ACDS were not quantitatively analyzed, the children's item selections were indicated in the author's field notes, as well as the interview transcripts and were therefore accounted for within the qualitative data analysis. A primarily top-down hybrid of the cognitive coding analysis model and the question feature coding model were used at the individual level within each transcript and accompanying interviewer field notes (Willis, 2015). According to Willis (2015) cognitive coding analysis models are deductive in nature and strongly reflect the Tourangeau question-and-answer model (1984) with each component of the model serving as a coding category. Categories therefore included comprehension, retrieval, judgment, and response, which focused on the cognitive processes of the respondents (Willis, 2015). A priori categories may be developed in advance, or as Willis (2015) suggests, subcategories within the Tourangeau model might be developed to increase precision and capture issues related to the instrument itself. The first author performed a formal analysis of the data by assigning coding categories on the basis of triggers to different segments of each interview (Willis, 2005). Within each of the coding categories, subcategories were developed to further specify question feature and processing difficulties. The subcategories were not determined a priori but rather emerged through constant comparison during the across case analysis following the analysis of individual cases (Willis, 2015). A second coder performed an independent analysis which confirmed the findings of the first author (Patton, 2002).

Study Quality

Given cognitive interviewing is a form of qualitative inquiry applied to instrument development, which underscores quantitative features (Willis, 2015), the quality of this study was sought in several ways. First, the interviewer engaged in pilot testing of the interview guide and probes to ensure their usefulness and understandability by children. This also afforded the opportunity to further refine her cognitive interviewing skills. Second, two sources of data (i.e., interviews and field notes) were used (Patton, 2002). Third, an independent coder confirmed the analysis of the data, an approach commonly used in cognitive interviewing (Willis, 2015). Finally given the important role of the researcher/interviewer in cognitive interviewing (Willis, 2015), working as a research team through the entire study assisted the first author/interviewer to question her assumptions and potential biases.

Results

In order to maintain the richness of the data, in recognition of our small sample size (Willis, 2015), and the systematically varied use of probes, we present the results using interview quotes to illustrate the categories and subcategories. Pseudonyms have been used to maintain the confidentiality of participants.

Comprehension

To ascertain if there was a match between respondent comprehension and researcher intention, children were asked to repeat questionnaire statements in their own words, to describe what they were

being asked, and to define different terms. Two subcategories emerged within the main category of comprehension: (a) understanding and (b) varying interpretations.

Understanding

In general children demonstrated an understanding of questionnaire statements that was in keeping with the researcher's intention. While children at times interpreted statements in slightly different ways, for the most part these interpretations confirmed that children understood the meaning of the statements. This is evidenced in the following dialogue with Ryan:

> Interviewer: Some kids wish they could be a lot better at sports, but other kids feel they are good enough at sports. So what is the difference between those two kids?

> Ryan: Umm, one kid's like bad at sports and another kid feels he's really good.

While not a direct interpretation, Ryan's paraphrasing demonstrated his understanding of the difference between the children in the statement. In the first part, he reported the child's actual competence, which he likely assumed was predictive of perceived competence. In the second part of his response, he indicated his understanding that it is not actual competence but perceived competence that the statement refers to. Likewise, Ethan responded to a comprehension probe about the statement 'some kids feel that they are better than others their age at sports, but other kids don't feel that they can play as well' by pointing to the second half of the statement and saying "these kinds of kids don't feel that they can do good."

Children were also asked to explain the meaning of the word 'sports'. Children provided examples of different kinds of sports and or skills involved in sports performance. Megan said, "sports involve kicking, running, maybe doing very high jumps" and cited soccer as an example. Tristan offered examples of sports that included: skiing, baseball, soccer, and snowboarding, while Emmett included examples of hockey and curling. While the children provided different sports examples and descriptions (e.g., skills versus sports), these were consistent with the athletic domain.

Varying interpretations

Despite examples and paraphrasing of statements that appeared to match with the intention of the researcher for the domain of athletic competence, subtle differences in the children's interpretations were also apparent. This is demonstrated in the following dialogue with Ethan:

> Interviewer: Okay, so you said like a sprint race or jumping thing, can you give me some more examples of other kinds of sports?

> Ethan: Umm, friendly sports where you are just playing for fun with your friend, you are just playing for fun and you go easy on each other and you're not too competitive you can still be competitive a little bit, but not too competitive, like 'err, you're going down, you're gonna eat my dirt.'

In this example, Ethan also incorporated the ideas of goal structures and motivational climates that have social and emotional implications. While his descriptions fell within an athletic domain, this quote also highlights how his expanded interpretation of the question might lead to a varied interpretation. A similar example is illustrated in the following quotes from children when they were asked to paraphrase statements:

> Interviewer: Now can you tell me in your own words what I was asking you about in that question [about some kids do very well at all kinds of sports, but other kids don't feel that they are very good when it comes to sports]?

> Billy: Am I comfortable doing sports.

Interviewer: In games and sports some kids usually watch instead of play, but other kids usually play rather than watch. [What is that question asking you?]

Megan: This question was asking me if I like to take part in sports, and I do.

Nicholas provided insight into his understanding of the children in the statement 'some kids think they could do well at just about any new sport activity they haven't tried before.' He said, "umm basically, this group has to uh, try to have more fun then, because umm, if you don't try you won't succeed at any goal you set for yourself." While within the athletic domain, children's comprehension of questions differed from one another in subtle ways, reflected values, and were tied to emotions. Differences also emerged in definitions of sports. Megan, when asked if dancing was a sport, responded that it wasn't "because dancing is more of a fun-time activity." She included in her definition of sports the idea that "lots of people have to try their best and this is what sports are all about." Billy said that sports meant "excited" and Emmett reported amid his examples of what sports were, that dancing and yoga were not sports.

It was evident from the children's responses to the comprehension probes, that they were making their interpretations within the athletic domain. Most children appeared to understand what was being asked of them (e.g., to identify with other children on the basis of how competent they felt in sports and games) and articulated an understanding of sports and games consistent with the athletic domain. At the same time, some of the children's interpretations diverged in interesting ways. Some children described enjoyment, goals, and the need for fun as they paraphrased the meaning of different statements. Likewise, certain physical activities, such as dance, were not considered to be part of sports or games. The degree to which the variations discussed are relevant when it comes to determining the quality of the instrument, then becomes of question of acceptable difference. These differences may have implications for the interpretation of test scores which emerge from questions assumed to tap the same domain in the same way for all participants.

Retrieval

The interviewer used verbal probes to investigate the kinds of information children recalled in item selection. Two subcategories emerged from the main category of retrieval: (a) skills and performance, and (b) past events and engagement.

Skills and performance

In recalling information about how they answered questions, children frequently referred to sports skills and performance. For example, Tristan indicated he felt good enough at sports because in skiing "I do sharp edges and I go fast." Megan selected that she was more like other kids who don't feel that they are very good at sports because as she said, "I'm not very good at soccer, I can't kick very hard or stop it." Use of skill performance knowledge for athletic competence decision making is also exemplified in the following quote from Ethan:

Because, umm, swimming, I can actually swim and really well actually…and umm, and I can do all sorts of stuff and soccer I kick really well. My secret kick is where I run back and then run up really fast and boot the ball.

Past events and engagement

Children also used examples of past experiences and activity engagement when answering questions about what they were thinking when responding to questionnaire items. "Like when I play sports, like when I get hit with the baseball when I play baseball, I chickened out," responded Emmett to a retrieval probe. Billy indicated he was more like kids who think they could do well at just about any new sports activity they haven't tried before because he had tried new things before and felt he did pretty well at them. Michael decided he was more like the kids who do very well at all kinds of

sports because when he played hockey and soccer he was selected to play in the more competitive group.

Skills and performance, and past events and engagement were often referred to in tandem. In a subsequent response Emmett said, "like when I go try something like baseball, it's like, it's like, I don't do well, like at first I keep hitting and hitting and missing and missing." When asked how he made his decision that he was more like the kids who do very well at all kinds of sports Ryan said, "cause I play a lot of sports, like I do running club, hockey, like I play hockey outside with my friends…ya and like I play different sports, last year my soccer team won silver." At times children also described their decision making based on personal characteristics such as exerting effort, persistence, and having fun. The information recalled to make decisions about the questionnaire items appeared consistent with the athletic domain, although children used different sources of retrieval information (e.g., playing lots of sports, making a team, winning) to make their item selections.

Judgment

Judgment involves the process of formulating an answer to a questionnaire item. Only one subcategory emerged in this category, that of certainty.

Certainty

When asked how certain they were about the answers they provided, all eight children indicated that they were either "sure" or "really sure" about their answers. Children indicated a high and unwavering degree of certainty about their answers.

Response

In formatting responses, the participants in this study had to select one of four boxes to indicate their perception of athletic competence for each item. Three subcategories emerged within the main category of response: (a) easy, (b) vulnerability, and (c) limited options.

Easy

When probed about the difficulty of answering questions, children most often responded that the questions were "easy" to answer because they knew their answers right away.

Vulnerability

At the same time, some children found it difficult to make questionnaire selections and respond to verbal probes when addressing feelings of low perceived competence. This is evidenced in the following conversation with Emmett:

> Interviewer: Is that a hard question for you to answer or an easy question for you to answer?
>
> Emmett: A hard one.
>
> Interviewer: What kinds of things make it hard to answer that kind of question for you?
>
> Emmett: Like do I want to tell her [the interviewer] the truth or should I just make it up or stuff like that?

In answer to a subsequent response probe, Emmett also shared, "if you don't feel you're good at it, it's hard [to answer]." While Nicholas said he found the questions pretty easy to answer, when asked how he felt about answering them he said, "umm, they are good for me cause then I get to uh, use examples and I really express how I feel." More often than not Nicholas' item selections suggested low perceptions of competence. Although he expressed feeling vulnerable (e.g., not feeling good at sports and games), he also perceived the interview as an opportunity to share his feelings. In this way, the

cognitive interviews elicited information about how children responded but also provided emotional support for Nicholas as he completed the ACDS.

Limited options

When children did express that a question was hard to answer it was most often because their answers did not fit into any of the four boxes provided. On at least one item, six of the eight children indicated they felt they were either in the middle or could be both of the statements on each side of the questionnaire. For example, Emmet and Ryan responded to the statement 'some kids feel they are better than other kids their age at sports but other kids don't feel they can play as well' that they could be both. Tristan hesitated in answering this same question and then said "actually, I do the same." On the next item, Tristan indicated that he plays and watches when it comes to games and sports. When asked if he was more like the kids who don't do well at new outdoor games or more like the ones who are good at new outdoor games right away Ethan responded:

> I think a little bit of both because it might take me some time to get used to the game, but like the first class I might be really, really good because other people have first classes too so they might be as good as me or I might be better than them. But when it takes some time I can be way better and they can be way better I can be really, really good, so I think it is like both.

In response to this same item, Megan indicated that both options were true for her because as she said, "I just think I don't do well at some outdoor games and I think I am good at other outdoor games right way." Response categories options that did not appear to align with the ways in which the children thought about themselves, is a critical finding.

Discussion

The comprehension results of this study for the most part revealed that children's paraphrasing and understanding of questions were in keeping with the meaning of the statements and the athletic domain. At times, however, despite general consistency within the domain, children's interpretations were different (e.g., focus on the need to have more fun versus emphasis on not feeling good at sports). Variation in comprehension is a source of invalidity when the comprehension does not overlap with the construct of interest as understood by the instrument developer. Determining an appropriate bandwidth for comprehension differences for construct validity purposes is a challenge in instrument development. Asking children additional questions about such things as past sport performances, to support understanding of responses, although time consuming, can offer additional insight that can also influence physical activity inclusion practices. Given the importance of perceived competence to feelings of inclusion (Spencer-Cavaliere & Watkinson, 2010), questions that also probe into more diverse physical activity contexts are important. Adding or modifying questions to also include adapted activities may be a way to increase question comprehension for children with impairments.

In response to retrieval probes, the children provided examples of sports skills and performance, and past events and engagement that informed their decision making in answering questions. Similar to the results of the comprehension probes, children provided examples that were different from each other. However, the retrieval examples provided were consistent with the athletic domain and as Harter (1981a, 1982) also found, the athletic domain was relevant and meaningful to participants. It was expected that different sources of information within the athletic domain would be used for retrieval, given that individuals have different experiences and unique perspectives to draw on in making decisions. Furthermore, it is these experiences and perspectives that inform children's perceptions of competence. Identifying the specific sources and types of experiences children with impairments attend to within the sport and games context can offer insight into how teachers and coaches may better support their participation in physical activity. For example, Tsalavoutas and Reid (2006) cautioned on the use of performance climates with children with physical impairments where

comparisons and performance disparities are prominent and Brian and colleagues (2018) emphasized the importance of accessible activities. Similar to the recommendation above, ensuring the concepts of sports and games in questionnaire items also reflect adapted activities could enhance relevance.

In reporting the certainty of their judgments, children unanimously responded that they were sure or really sure about their answers. This could mean that children were easily able to formulate their answers because they understood the questions, the questions were applicable, and the information required to answer them was accessible (Collins, 2015). At the same time it is important to consider that in questions about subjective phenomena such as an individual's self-evaluation of competence which may differ for children with impairments in particular, judgment is at the mercy of the subjective nature of the phenomena and the respondent's perceived certainty in question response (Schaeffer & Presser, 2003).

According to children's answers to the verbal response probes, questionnaire items were easy to answer. Instances where this was not the case occurred when expressions of low perceived competence led to vulnerability and when the questionnaire format was incongruent with children's self-perceptions. The structured alternative format of the SPPC-ACDS was designed to limit socially desirable answers by having respondents identify with other children who are like them. This format may have reduced the number of answers influenced by social desirability as suggested by Harter (1982) (although this is not known based on the results of this study), but the format itself did not alleviate the anxiety of discussing low perceptions of competence for at least two children. Recognizing the child as an important source of information brings with it ethical and methodological considerations about consent, vulnerability, children's competence and reliability as informants, and the best methods to elicit information from children (Dockett, Einarsdottir, & Perry, 2009; Oberg & Ellis, 2006). Vulnerability is also a key consideration given findings that children with different impairments often experience lower levels of perceived motor competence than their peers (Brian et al., 2018; Causgrove Dunn & Dunn, 2006) and exclusion from physical activity is not uncommon (Spencer-Cavaliere & Watkinson, 2010). A possible way to address vulnerability is in consideration of how the questionnaire is administered and the specific messaging children receive about the nature of the questions and their responses. For example, emphasizing that children's answers are their own and offering reassurance around feeling more like other children who are 'not good at sports and games' could be helpful.

In this study, several children indicated that on at least one question, the options provided on the questionnaire did not represent their self-perceptions. This underscores what has long been a criticism of closed-answer questions, that closed questions, or set response formats may shape the answers respondents provide (Schwarz, 1999). Stated differently, participants may answer in ways they otherwise would not if provided different or unlimited choices. The importance of response options that align with children's thinking was also a finding in the cognitive interviewing study conducted by Woolley and colleagues (2004). Moving forward, a recommendation for this particular scale would be to compare scores from a typical Likert scale to one with the structured alternative format. Also, given the children in this study had experiences in both integrated and segregated settings, questions that deliberately tap into these contexts are also important.

Our results with regard to response options also highlight what could be a discrepancy between the theory of competence motivation (Harter, 1978, 1981a) and the format of the SPPC-ACDS (Harter, 1985). Specifically, the structured alternative format appears to contradict Harter's theory that it is around the age of 8 years old that children begin to integrate positive and negative views of the self, meaning they can integrate seemingly opposing self-evaluations (Harter, 1999). In the athletic domain this suggests that a child could perceive herself to be good at sports, but at the same time wish she was better at sports, a response not supported by the questionnaire format. The alternative response format also does not account for children who legitimately perceive themselves to be between or similar to both types of children described in the statements. A mismatch in format and respondent thinking could lead children to select a response option that does not represent their

thinking, to select two options, or make no selection at all. The first possibility could result in inappropriate inferences about a child's self-evaluation, one that does not match their actual thinking. The second and third possibilities may be recorded as response errors or missed cases, when in fact the thinking behind the responses reveals something significant about self-perception. Further investigation of the structured alternative format is recommended. While this comment is relevant to all children, for children with impairments, who have fewer accessible physical activity choices, and may rely on adapted activities as a locale where they experience higher perceived competence (Wynnyk & Spencer-Cavaliere, 2013) response options that capture a broader range of possibilities is important.

The extent to which varying understanding of questions, retrieval of information, and response option limitations impact test scores and their appropriate interpretations is a challenging issue. Based on responses to the verbal probes in this study, the SPPC-ACDS appeared to tap children's self-evaluation in the athletic domain. However, the information revealed through the cognitive interviews highlights the value of this approach in gaining a deeper understanding of these perceptions as well as revealing some potential questionnaire pitfalls. These deeper understandings are particularly critical for children with impairments as they may lead to better and more ethical physical activity practices. We also contend that the identification of sources of invalidity is a crucial step in the development of instruments and in the modification of instruments for use with diverse populations. At the same time, we recognize that the development of Harter's scale took place before cognitive interview practices were more emergent and acknowledge the quality of the processes that were undertaken in the development of the SPPC.

The results of this study also highlighted issues related to individual differences in comprehension, retrieval, and response to questionnaire items which are challenging for interpretation and inference making. This problem was recognized by Messick (1989) who stated "that a test's construct interpretation might need to vary from one type of person to another (or from one setting or occasion to another) is a major current conundrum in educational and psychological measurement" (p. 55). Within a field such as APA, heterogeneity is expected (Spencer-Cavaliere et al., 2017) and therefore an important consideration in construct interpretation. Demonstrating that an instrument measures what is intended is essential to measurement based research that seeks to understand how people think and behave. Statistical techniques commonly used to infer validity for the interpretation of test scores do not address the issue of ontology. It is the causal relationship between construct and test scores as demonstrated through the processes that bring about this effect that provides validity evidence for test interpretation (Boorsboom, 2005).

> If one attempts to sidestep the most important part of test behaviour, which is what happens between item administration and item response, then one will find no clarity in tables of correlation coefficients. No amount of empirical data can fill a theoretical gap. (Borsboom, 2005, p.167)

Investigating the question-and-answer process may bring clarity to the issue of whether a test measures what it purports to measure, providing validity evidence and lending confidence to the interpretation of test scores particularly when questionnaires are adapted for diverse populations. According to Messick (1989), the validity of an inference is a matter of degree, an inference is neither absolutely valid nor invalid. Over time, validity or invalidity evidence continue to grow. In this study, cognitive interviews using Tourangeau's (1984) question-and-answer model to probe the thinking of children with impairments' in responding to the SPPC-ACDS provided both support for the validity of the scale and revealed areas for further inquiry.

When information is to be elicited from children about their own perspectives, it is critical that children understand the questions being asked of them and that researchers understand the responses given by them (Woolley et al., 2004). The methods used in this study elicited information about children's question interpretations that were primarily consistent with the domain of athletic

competence, but at times varied. When self-report questionnaires are administered there is an assumption that responses are based on questions perceived synonymously with other respondents and with the intention of the researcher (Collins, 2015). If this is not the case, then a potential source of invalidity has been identified therefore, the variations in interpretations of the ACDS items may need further exploration.

Limitations

In the present study we used the model advanced by Tourangeau (1984) to guide our efforts. Although this model has been a useful tool, more comprehensive theories of response production processes are needed. Furthermore, the interview context, the context within which the respondent exists, and the interviewer all make contributions to issues with questionnaire development and use (Willis, 2015). These are important considerations when interpreting data. Despite the acceptance of studying cognitive processes in questionnaire development (Collins, 2015; Willis, 2015) and the suggestion that verbal reports can be critical and reliable sources of information, the trustworthiness and usefulness of verbal reports has and continues to be debated (Ericsson & Simon, 1993).

Conclusions

Despite current limitations, verbal reports and in particular cognitive interviews, may allow for a better understanding of the question response process (Knafl et al., 2007) and provide validity, or invalidity, evidence not revealed through the use of other methods. Gaining access to children's thought processes is critical to ensuring a higher degree of congruence not only between researcher intention and participant understanding but also between the researcher's interpretation of the information provided by the child, but may pose specific methodological challenges (Branch, 2006). This study highlighted areas of strength and discord in children's understanding of the SPPC-ACDS that offer future consideration for how this questionnaire might be used in ways most useful and relevant to knowledge generation both within APA research and beyond it.

Perspectives

The findings of this study are particularly relevant to the field of APA. First, they demonstrate how the use of cognitive interviews can support researchers in both the development and adaptation of instruments for use with diverse populations and specifically, children with impairments (Collins, 2015). This work also offers a way in which researchers can attempt to gain deeper understanding of children's perceptions. Understanding children's perceptions is critical to the field of APA and has significant bearing on such things as inclusion, motivation, and competence in activity settings (Harter, 1999; Shapiro et al., 2005; Spencer-Cavaliere & Watkinson, 2010, Wynnyk & Spencer-Cavaliere, 2013). By extension, these interviews can corroborate quantitative findings and offer greater insight into APA ethical practice that importantly considers the voices of participants (Bredahl, 2008).

Author affiliations:
[1]   Faculty of Kinesiology, Sport, and Recreation, University of Alberta; ncavalie@ualberta.ca, marcel.bouffard@ualberta.ca, mailto:jane.watkinson@ualberta.ca
[*]   Correspondence: ncavalie@ualberta.ca; Tel.: +11-780-492-9615

Appendix

Interview Guide Probes

Comprehension

1. Can you tell me what the word (insert word) means?
2. Can you tell me the question in your own words?
3. Can you tell me what the question is asking you?

Retrieval

1. How did you know that?
2. What were you thinking when you were answering that question?
3. How did you decide to answer that question?

Judgement

1. How sure are you about your answer?

Response

1. How did you feel about answering that question?
2. Was that hard or easy to answer? Can you tell me why?

References

Beatty, P. C., & Willis, G. B. (2007). Research synthesis: The practice of cognitive interviewing. *Public Opinion Quarterly, 71,* 287-311. doi:10.1093/poq/nfm006

Borsboom, D. (2005). *Measuring the mind: Conceptual issues in contemporary psychometrics.* Cambridge, MA: Cambridge University Press.

Bouffard, M., & Spencer-Cavaliere, N. (2016). Interdisciplinarity in Adapted Physical Activity. *Quest, 68,* 4-14. doi:10.1080/00336297.2015.1117002

Branch, J. L. (2006). Using think alouds, think afters, and think togethers to research adolescents' inquiry experiences. *Alberta Journal of Educational Research, 52,* p. 148-159. https://journalhosting.ucalgary.ca/index.php/ajer/article/view/55153

Bredhal, A.-M. (2008). Ethical aspects in research in adapted physical activity. *Sport, Ethics, and Philosophy, 2,* 257-270. doi:10.1080/17511320802223881

Brian, A., De Meester, A., Klavina, A., Irwin, J. M., Taunton, S., Pennell, A., & Lieberman, L. (2019). Exploring children/adolescents with visual impairments' physical literacy: A preliminary investigation of autonomous motivation. *Journal of Teaching in Physical Education, 38,* 155-161. doi: 10.1123/jtpe.2018-0194

Brian, A., Haegele, J.A., & Bostick, L. (2016). Perceived motor competence of children with visual impairments: A preliminary investigation. *British Journal of Visual Impairment, 34,* 151–155. doi:10.1177/ 0264619616628575

Brian, A., Taunton, S., Haibach-Beach, P., & Lieberman, L.J. (2018). Influence of a sports camp on perceived physical competence of children who are visually impaired. *Journal of Visual Impairment & Blindness, 112,* 505–509. doi: 10.1177/0145482X1811200508

Causgrove Dunn, J., & Dunn, J. G. H. (2006). Perceptions of the motivational climate, perceived competence, and participation behaviors of children with movement difficulties in physical education. *Adapted Physical Activity Quarterly, 23,* 293-309. doi: 10.1123/apaq.23.3.293

Castillo-Díaz, M., & Padilla, J.-L. (2013). How cognitive interviewing can provide validity evidence of the response processes to scale items. *Social Indicators Research, 114,* 963-975. doi: 10.1007/s11205-012-0184-8

Collins, D. (2015). Cognitive interviewing: Origin, purpose and limitations. In, Collins, D. (Ed.), *Cognitive interviewing practice* (2nd ed., pp. 3-27). Thousand Oaks, CA: Sage.

Danermark, B., Ekström, M., Jakobsen, L., & Karlsson, J.C. (2002). *Explaining society: Critical realism in the social sciences.* New York: Routledge.

Dietrich, H., & Ehrlenspiel, F. (2010). Cognitive interviewing: A qualitative tool for improving questionnaires in sport science. *Measurement in Physical Education and Exercise Science, 14,* 51-60. doi:10.1080/10913670903455025

Dockett, S., Einarsdottir, J., & Perry, B. (2009). Researching with children: Ethical tensions. *Early Childhood Research, 7,* 283-298. doi: 10.1177/1476718X09336971

Ericsson, K. A. & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data.* Cambridge, MA: MIT Press.

Groves, R. M., Fowler, F. J., Jr., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey methodology* (2nd ed.). New Jersey: John Wiley & Sons Inc.

Harter, S. (1978). Effectance motivation reconsidered: Toward a developmental model. *Human Development, 21,* 34-64. doi:10.1159/000271574

Harter, S. (1981a). A model of mastery motivation in children: Individual differences and developmental change. In W. A. Collins (Ed.), *Minnesota Symposium on Child Psychology,* (Vol. 14, pp. 215-255). Hillsdale, NJ: Erlbaum.

Harter, S. (1981b). A new self-report scale of intrinsic versus extrinsic orientation in the classroom: Motivational and informational components. *Developmental Psychology, 17,* 300-312. doi:10.1037/0012-1649.17.3.300

Harter, S. (1982). The perceived competence scale for children. *Child Development, 53,* 87-97. doi: 10.2307/1129640

Harter, S. (1985). *Manual for the Self-Perception Profile for Children.* University of Denver.

Harter, S. (1990). Issues in the assessment of the self-concept of children and adolescents. In A. M. La Greca (Ed.), *Through the eyes of the child: Obtaining self-reports from children* (pp. 292-325). Boston: Allyn and Bacon:

Harter, S. (1999). *The construction of the self: A developmental perspective.* New York, NY: The Guilford Press.

Horvat, M., Block, M. E., & Kelly, L. E. (2007). *Developmental and adapted physical activity assessment.* Champaign, IL: Human Kinetics.

Knafl, K., Deatrick, J., Gallo, A., Holcombe, G., Bakitas, M., Dixon, J., & Grey, M. (2007). The analysis and interpretation of cognitive interviews for instrument development. *Research in Nursing & Health, 30,* 224-234. doi: 10.1002/nur.20195

Leary, J. M., Ice, C., & Cottrell, L. (2012). Adaptation and cognitive testing of physical activity measures for use with young, school-aged children and their parents. *Quality of Life Research, 21,* 1815-1828. doi: 10.1007/s11136-011-0095-1

Messick, S. (1989). Validity. In R. L. Linn, *Educational measurement* (3rd ed.). NewYork: Macmillan Publishing Company.

Neilson, H. K., Ullman, R., Robson, P. J., Friedenreich, C. M., & Csizmadi, I. (2013). Cognitive testing of the STAR-Q: Insights in activity and sedentary time reporting. *Journal of Physical Activity and Health, 10,* 379-389. doi: 10.1123/jpah.10.3.379

Oberg, D., & Ellis, J. (2006). Exploring the new paradigm for researching with children and youth. *Alberta Journal of Educational Research, 52*(3), 107-110. https://journalhosting.ucalgary.ca/index.php/ajer/article/view/55150

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Renick, M. J., & Harter, S. (1989). Impact of social comparisons on the developing self perceptions of learning disabled students. *Journal of Educational Psychology, 81,* 631-638. doi:10.1037/0022-0663.81.4.631

Schaeffer, N. C., & Presser, S. (2003). The science of asking questions. *Annual Review of Sociology, 29,* 65-88. doi: 10.1146/annurev.soc.29.110702.110112

Schuengel, C., Voorman, J., Stolk, J., Dallmeijer, A., Vermeer, A., & Becher, J. (2006). Self-worth, perceived competence, and behaviour problems in children with cerebral palsy. *Disability and Rehabilitation, 28,* 1251-1258. doi:10.1080/09638280600554652

Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist, 54,* 93-105. doi:10.1037/0003-066X.54.2.93

Shapiro, D. R., Moffett, A., Lieberman, L., & Dummer, G. M. (2005). Perceived competence of children with visual impairments. *Journal of Visual Impairment & Blindness, 99,* 15-25. doi: 10.1177/0145482X0509900103

Shapiro, D. R., & Ulrich, D. A. (2002). Expectancies, values, and perceptions of physical competence of children with and without learning disabilities. *Adapted Physical Activity Quarterly, 19,* 318-333. doi: 10.1123/apaq.19.3.318

Silon, E. L., & Harter, S. (1985). Assessment of perceived competence, motivational orientation, and anxiety in segregated and mainstreamed educable mentally retarded children. *Journal of Educational Psychology, 77,* 217-230. doi:10.1037/0022-0663.77.2.217

Spencer-Cavaliere, N., Thai, J., & Kingsley, B. (2017). A part of and apart from sport: Practitioner's experiences coaching in segregated youth sport. *Social Inclusion, 5,* 120-129. doi:10.17645/si.v5i2.889

Spencer-Cavaliere, N., & Watkinson, E. J. (2010). Inclusion understood from the perspectives of children with disability. *Adapted Physical Activity Quarterly, 27,* 275-293. doi: 10.1123/apaq.27.4.275

Tourangeau, R. (1984). Cognitive science and survey methods. In T. Jabine, M. Straf, J. Tanur, & R. Tourangeau (Eds.), *Cognitive aspects of survey design: Building a bridge between disciplines* (pp. 73-100). Washington, DC: National Academy Press.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response.* Cambridge, UK: Cambridge University Press.

Tsalavoutas, I., & Reid, G. (2006). Competence satisfaction: Risk taking and achievement. *Adapted Physical Activity Quarterly, 23,* 410-423. doi: 10.1123/apaq.23.4.410

Ullrich-French, S., Cox, A., Cole, A., Roades Cooper, B., & Gotch, C. (2017). Initial validity evidence for the state mindfulness scale for physical activity with youth. *Measurement in Physical Education and Exercise Science, 21,* 177-189. doi: 10.1080/1091367X.2017.1321543

Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design.* Thousand Oaks, CA: Sage Publications.

Willis, G. B. (2015). *Analysis of the cognitive interview in questionnaire design.* New York, NY: Oxford University Press.

Woolley, M. E., Bowen, G. L., & Bowen, N. K. (2004). Cognitive pretesting and the developmental validity of child self-report instruments: Theory and applications. *Research on Social Work Practice, 14,* 191-200. doi:10.1177/1049731503257882

Wynnyk, K., & Spencer-Cavaliere, N. (2013). Children's social relationships and motivation in sledge hockey. *Adapted Physical Activity Quarterly, 30,* 299-316. doi: 10.1123/apaq.30.4.299

Yun, J, & Ulrich, D. A. (1997). Perceived and actual physical competence in children with mild mental retardation. *Adapted Physical Activity Quarterly, 14,* 285-297. doi: 10.1123/apaq.14.4.285