



Development and validation of a questionnaire to measure primary school children's images of and attitudes towards curiosity (the CIAC questionnaire)

Tim Post^{1,2} · Juliette H. Walma van der Molen¹

Published online: 5 October 2018
© The Author(s) 2018

Abstract

This paper presents a validation study of a questionnaire to measure primary children's images of and attitudes towards curiosity (the CIAC questionnaire). Policy documents and scientific studies on twenty-first century learning increasingly promote the value of stimulating children's curiosity in primary school. However, no well-established measurement instruments yet exist to assess children's curiosity within educational settings. To fill this void, we focused on the measurement of children's *perceptions* of curiosity, as important precursors to children's potential curiosity-driven behavior. Based on attitude and curiosity theory, we developed seven components of children's images of and attitudes towards curiosity. We translated these components into corresponding measurement scales, which comprise the CIAC. Results of a validation study among 737 children (ages 8–13), using factor analyses, largely confirmed the factor structures of the image and attitude scales and indicated good convergent and discriminant validity. In addition, we provide evidence for the predictive power of children's images and attitudes on their motivation to be curious.

Keywords Curiosity · Attitude · Survey development · Instrument validation · Primary education

Introduction

This paper presents a validation study of a questionnaire to measure primary children's images of and attitudes towards curiosity (the CIAC questionnaire). Curiosity may be defined as a desire to seek and acquire new information (Berlyne 1954; Kashdan 2004; Litman 2008; Loewenstein 1994). Historically, Berlyne (1954) and Piaget (1952) were among the first to propose that curiosity may motivate complex exploratory learning behavior. Subsequent research by, for example, Loewenstein (1994), Litman et al. (2005), and Kashdan et al. (2007) has added to the work of Berlyne and Piaget by further defining the dimensionality, determinants, and measures of curiosity.

Recently, research has shifted focus to the investigation of curiosity within educational settings (for reviews on the topic, please see Grossnickle 2016; Jirout and Klahr 2012). Policy documents on twenty-first century learning increasingly promote school curricula that aim to engage children in the scientific process of knowledge development (Lucas et al. 2013; OECD 2015; Pellegrino and Hilton 2012). In school settings, children's curiosity is linked to wonderment (e.g., Opdal 2001), question-asking (e.g., Jirout 2011), and explanation-seeking behavior (e.g., Litman et al. 2005) and predominantly understood in terms of *epistemic* curiosity: the desire to seek and acquire new intellectual information (Litman and Spielberger 2003; Loewenstein 1994; Piotrowski et al. 2014). Epistemic curiosity is believed to improve children's undertaking of complex inquiry activities (e.g., Von Stumm et al. 2011), their persistence with learning (e.g., Metz 2008), and their memorization of new information (e.g., Jepma et al. 2012).

The clear educational value of epistemic curiosity has led many researchers and education policy-makers to advocate the implementation of curiosity-focused pedagogy in primary schools (Claxton and Carr 2004; Engel 2006, 2011; Jirout and Klahr 2012; OECD 2015; Osborne and Dillon

✉ Tim Post
t.post@utwente.nl

¹ University of Twente, Enschede, The Netherlands

² Centre of Science Education and Talent Development, Institute for Teacher Education, Science Communication & School Practices, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands

2008; Lucas et al. 2013). Unfortunately, however, no well-established guidelines or instruments yet exist to promote or assess children's epistemic curiosity in school settings (for a review on this topic, please see Jirout and Klahr 2012). Curiosity research has been mostly limited to measuring children's curiosity behavior in response to curiosity-eliciting stimuli (e.g., toys, games) in laboratory settings, rather than in everyday classroom practice. In addition, proposed curiosity definitions often confound the concept of curiosity with the concepts of interest, intelligence, or motivation, which makes it unclear what is measured or what mechanisms may underlie children's curiosity behavior (Grossnickle 2016; Silvia and Sanders 2010).

Studies concerning children's curiosity in primary school settings suggest that the educational content and pedagogy offer children little encouragement to be curious, even as part of inquiry-oriented educational activities (Engel 2006, 2011; Engel and Randall 2009; Fortus 2014). Lesson activities usually emphasize to children the notion that there is just one correct answer to or solution for questions and assignments, and that diverse question-asking and explanation-seeking is disruptive to teachers' pre-scripted instruction (Post and Walma van der Molen 2018; Van Booven 2015).

It is our belief that children's epistemic curiosity in school can only flourish in a positive classroom climate in which children are taught the epistemic value of being curious-minded, in which they derive pleasure from expressing questions and ideas, and feel that such questions and ideas are appreciated by their teachers and peers. Such perceptions of curiosity can be understood in terms of *attitudes*. Decades of social psychological research show that attitudinal beliefs and affects are important precursors to behavior (e.g., Ajzen 2001; Ajzen and Fishbein 1980; Eagly and Chaiken 1993). In addition, recent research on achievement motivation shows that students' motivational beliefs determine the type of learning strategies they employ (Muis et al. 2015a, b). According to the control-value theory as proposed by Pekrun (2006), the types of emotions students may experience in school settings depend on their perceptions of control and value evaluations, which can be understood as the degree to which students subjectively *attribute* importance to achievement-related activities (Pekrun et al. 2011). Thus, on the basis of attitude and motivation theory, we expect that children's images of 'epistemic curiosity' and their perceptions of the value of being curious in school would precede their curiosity behavior in the classroom. Rather than focusing on children's curiosity *behavior, skills, traits, or states*, in this paper, we therefore aimed to investigate children's *perceptions* of curiosity: their images of and attitudes towards curiosity. Shifting the focus towards the investigation of children's attitudes towards curiosity also has some practical advantages. First, attitudes can be measured (Blalock et al. 2008; Reid 2006). Second, while attitude is considered a

relatively stable psychological construct, attitudes can be improved over time (Vogel and Wänke 2016).

Present study

The current validation study occurred in the context of a larger investigation of the effects of an attitude-focused approach to fostering children's epistemic curiosity in primary schools. For this investigation, we conducted a large-scale intervention study in The Netherlands in which six primary schools participated in a school-wide, curiosity-focused teacher-training program. In addition to measuring changes in teachers' everyday practice over time, the study investigates children's images of and attitudes towards curiosity from the 4th to the 6th grade. To measure changes over time, we developed the Children's Images of and Attitudes towards Curiosity (CIAC) questionnaire. The CIAC provides a comprehensive ensemble of image and attitude components that, according to attitude theory (in particular, the Theory of Planned Behavior (TPB); Ajzen 2001), may precede children's curious behavior in the classroom.

In this paper, we describe the development of the CIAC and present the results of our validation study. Our study included qualitative as well as extensive quantitative methods to determine the convergent and discriminant validity of the developed measurement scales. In the next section, we first outline the theoretical framework that underlies the scales of the CIAC.

Theoretical framework for the CIAC questionnaire

While many different definitions of attitude exist in the literature, the concept of attitude is traditionally described as the psychological 'tendency' of a person to evaluate a particular 'attitude object' in terms of favorable or unfavorable perceptions (Ajzen 2001; Ajzen and Fishbein 1980; Eagly and Chaiken 1993). In this paper, we investigated *epistemic curiosity* as an object of children's attitude in school settings. One of the most well-known models of attitude is the TPB (Ajzen 1991, 2001; Armitage and Conner 2001; Godin and Kok 1996; Hausenblas et al. 1997).

The TPB generally distinguishes three dimensions of attitude. The first dimension, *Perceptions of Behavioral Attributes*, represents the beliefs and feelings that a person attributes to the attitude object (e.g., the belief that epistemic curiosity fosters one's learning or the pleasurable feeling of posing epistemic questions or ideas in class). The second dimension, *Perceptions of the Social Norm*, describes a person's perception of the social acceptability of the behavior. This may include both a person's negative judgment of others and a person's fear of other people's judgments (e.g.,

children may fear negative judgments about their own curiosity behavior or they may judge others for the same behavior). The third dimension, *Perceptions of Behavioral Control*, represents the perceived level of control that a person experiences when performing a certain behavior with respect to the attitude object. In attitude literature, self-efficacy is often regarded as a central component of the *Perceptions of Behavioral Control* dimension of attitude (Ajzen 1991; Eagly and Chaiken 1993; Olson and Zanna 1993). Self-efficacy is a person's perceived capability to perform a behavior when opportunity is provided to do so (e.g., a child's feeling of self-efficacy to pose epistemic questions) (Bandura 1997).

Together, cognitive, affective, normative, and perceived control perceptions may determine the formulation of a behavioral *intention* to perform or not perform related behavior (e.g., to pose epistemic questions, seek alternative explanations). According to the Expectancy-Value Model (Fishbein and Ajzen 1974), the strength of this behavioral intention depends on the strength of a person's attitudinal perceptions. In the present study, we asked what cognitive, affective, normative, and control perceptions may constitute important components of children's attitudes towards epistemic curiosity in school settings. Based on a review of research on attitude (e.g., Ajzen 2001; Eagly and Chaiken 1993), curiosity (e.g., Engel 2006, 2011; Grossnickle 2016; Jirout and Klahr 2012), scientific literacy (e.g., Osborne and Dillon 2008; National Research Council 2012), and lifelong learning and creativity (e.g., Claxton and Carr 2004; Lucas et al. 2013), we developed a theoretical framework that describes the components of children's attitudes towards epistemic curiosity that seem most relevant for primary school children's curiosity-related behavior.

In addition, we added a separate dimension to the framework that represents children's *images* of curiosity: their mental representation of the term 'curiosity' (Eagly and Chaiken 1993). These images constitute the 'object' of children's attitudinal evaluation and, therefore, function as an essential determinant of their attitude towards curiosity (Ajzen 2001; Eagly and Chaiken 1993; Osborne et al. 2003). Previous research has shown that many primary school children predominantly associate curiosity with social behavior, such as spying on or prying about others (see Litman and Pezzo 2007), as opposed to academic learning, such as exploring new subject matter or considering alternative explanations to intellectual problems (Grossnickle 2016; Post and Walma van der Molen 2018). Such a narrow image of curiosity may prevent children from perceiving the epistemic value of curiosity for academic learning. To measure children's images of curiosity and to examine the ways in which children's images are related to their attitudes towards epistemic curiosity, we included children's social and epistemic images of curiosity as two separate image components in the framework.

Children's images of curiosity

Epistemic image of curiosity

Epistemic images of curiosity portray behaviors of seeking or obtaining new intellectual information (Litman 2008; Litman et al. 2010; Piotrowski et al. 2014). Examples of such behavior include wanting to know how the human body works or how computers work. In addition, epistemic curiosity could also involve the desire to learn about the epistemology of certain ideas or inventions (e.g., wanting to know how computers were invented).

Social images of curiosity

Social images of curiosity represent behaviors of seeking or obtaining new information about social experiences (Litman and Pezzo 2007; Litman et al. 2016). Examples of such behavior include spying on or prying about other people. Although children's social curiosity does not directly serve their academic learning, social curiosity is considered to play an important role in children's social development (Grossnickle 2016), because information about others may help to form friendships (Rosnow 2001) or to avoid negative social confrontations (Galen and Underwood 1997).

Children's attitudes towards epistemic curiosity

Personal relevance

We derived the *Personal Relevance* component on the basis of the *Perceptions of Behavioral Attributes* dimension of the TPB. This cognitive component concerns a child's perception of the value of expressing epistemic questions and ideas in class to improve one's own learning. Children will be more likely to engage in curious thinking in school when they perceive the positive outcomes of doing so for their own learning performance (Claxton 2007; Claxton and Carr 2004; Lucas et al. 2013; Pellegrino and Hilton 2012). Thus, we reasoned that children's perceived personal relevance of being curious might constitute an important consideration in whether children do or do not engage in curious thinking in the classroom.

Personal enjoyment

The *Personal Enjoyment* component refers to the pleasurable feeling of asking questions or coming up with new ideas (Arango-Muñoz 2014; Kashdan and Steger 2007; Piotrowski et al. 2014). We also derived this affective component of attitude based on the *Perceptions of Behavioral Attributes* dimension of the TPB. It should be noted that the mere pleasure of asking questions opposes typical deficit-type

descriptions that define curiosity as a generally unpleasant feeling of ‘not-knowing’ that needs to be reduced by learning (for a literature review of opposing views on this topic, please see Grossnickle 2016). Yet, do we really want children to develop the type of curiosity in school that drives them to reduce their unpleasant feeling of ‘not-knowing’? We do not think so. Moreover, while curiosity and enjoyment are conceptualized in the literature as distinct psychological constructs, they do share important features and are believed to both determine children’s intrinsic motivation to learn (Reeve 1989; Grossnickle 2016). Thus, if we want children to develop an openness to learning, we should show them the joy of question-asking and explanation-seeking, even if their epistemic questions or ideas do not lead to instant answers or solutions. In our view, such an interest-type quality of epistemic curiosity fits better with the twenty-first century education standards that policy-makers aim to achieve (e.g., Lucas et al. 2013; National Research Council 2012; OECD 2015).

Societal relevance

The *Societal Relevance* component was also derived on the basis of the *Perceptions of Behavioral Attributes* dimensions of the TPB. This cognitive component concerns children’s perception of the value of curious thinkers to society. Studies have stressed that typical teacher-directed and scripted forms of education may unintentionally convey to children that our collective understanding of the world is already absolute and complete, without revealing to children the tentative nature of such knowledge and the epistemological value of curious thinkers to society (Abd-El-Khalick 2012; Fouad et al. 2015; Trevors et al. 2017). Such everyday practice may lead children to develop misconceptions about the societal relevance of curiosity (Post and Walma van der Molen 2018). In our view, such misconceptions may prevent them from perceiving the need for curiosity-driven thinkers in society and, thereby, from acting upon their curiosity in school.

Fear of negative judgment

The *Fear of Negative Judgment* component concerns children’s fears of their peers’ or teachers’ negative judgments about being curious in class. We derived the *Fear of Negative Judgment* component on the basis of the *Perceptions of the Social Norm* dimension of the TPB. Research suggests that, in many countries, primary teachers generally feel uncomfortable when children ask diverse questions about topics that teachers themselves do not know the answers to (Ramey-Gassert et al. 1996; Schoon and Boone 1998; van Aalderen-Smeets and Walma van der Molen 2015; Van Booven 2015). Consequently, many teachers tend to shy away from stimulating children’s curious thinking or to

judge children’s curious questions and ideas as disruptive to their instruction (Claxton and Carr 2004; Claxton 2007). It seems plausible to assume that, over time, this may lead children to develop negative perceptions of what their peers and teachers may think about their curious behavior, which will discourage them from enacting such behavior in the classroom (Marx and Harris 2006; McCombs et al. 2008; Post and Walma van der Molen 2018). If children’s implicit negative perceptions are not explicitly attended to by teachers, these may well persist throughout (and beyond) primary school even when teachers provide children opportunity to inquire and curiously explore study subject matter (Post and Walma van der Molen 2018).

Negative opinion

Apart from fearing the judgments by others, we expect that many children may also hold negative judgments about other people’s curious question-asking and explanation-seeking behavior. Thus, the *Negative Opinion* component refers to a child’s negative opinion about other people’s curiosity-related behavior. Similar to the *Fear of Negative Judgment* component, we derived the *Negative Opinion* component on the basis of the *Perceptions of the Social Norm* dimension of the TPB, and likewise expect that children’s negative opinions about curious thinkers prevent them from expressing their own epistemic questions and ideas in class.

Self-efficacy

Lastly, the *Self-Efficacy* component refers to children’s perceived capability to express epistemic questions or ideas in class when sufficient opportunity is provided. Self-efficacy forms an essential component of the *Perception of Behavioral Control* dimension of the TPB (Ajzen 1991; Eagly and Chaiken 1993; Olson and Zanna 1993) and is widely regarded to be one of the most important determinants of behavior (Bandura 1997; Palmer 2006). In line with attitude and self-efficacy theories, we reason that children who perceive themselves to be capable question-askers and explanation-seekers are more likely to carry out such behaviors in comparison to children who feel insecure about their capabilities in this respect.

Hypotheses

We believe that the above-described image and attitude components may represent children’s primary considerations to perform (or not perform) epistemic curiosity-related behaviors in school. We differentiate between children’s social and epistemic *images* of curiosity in the ‘image scale’ of our survey. In addition, we propose six components of children’s

attitudes towards epistemic curiosity in our ‘attitude scale’. Four attitude components represent perceptions that may positively contribute to children’s curiosity behavior (*Personal Relevance*, *Personal Enjoyment*, *Societal Relevance*, and *Self-Efficacy*). The other two components represent perceptions that may negatively influence children’s curiosity behavior (*Fear of Negative Judgment* and *Negative Opinion*). We hypothesized that our proposed ‘positive components’ would correlate positively. Similarly, we expected that our ‘negative components’ would show a positive correlation. In addition, we expected that the ‘positive’ and ‘negative’ attitude components would be either unrelated or negatively correlated.

We also hypothesized, on the basis of attitude theory, that children’s image scores would predict their attitude scores (Ajzen 2001; Eagly and Chaiken 1993). In particular, we expected that scores on the *Epistemic Image* component would positively predict scores on the *Personal Relevance*, *Personal Enjoyment*, *Societal Relevance*, and *Self-Efficacy* components, because these attitude components positively relate to the epistemic value of curiosity. We expected that scores on the *Epistemic Image* component would negatively predict scores on the *Negative Opinion* component, because children who associate curiosity with its epistemic use would likely not perceive others’ epistemic curiosity behavior to be inappropriate. We expected the *Epistemic Image* component to be unrelated or negatively related to the *Fear of Negative Judgment* component, because children’s own epistemic images of curiosity may exist independently from their perceptions of other people’s negative opinions about epistemic curiosity behavior. In addition, we expected the *Social Image* component to be unrelated to our attitude components, because children’s association of curiosity with social behavior has little relation to the use of curiosity for academic learning and may thus function independently of a positive attitude towards epistemic curiosity.

In line with the TPB, we also considered the predictive power of children’s attitude scores on their motivation to be curious (Conner and Armitage 1998; Fishbein and Ajzen 1974). Research on curiosity suggests that interest-type curiosity may be positively related to mastery orientation motivation in particular (Grossnickle 2016; Litman 2008). However, because our proposed components of children’s attitudes towards epistemic curiosity stem from different underlying attitudinal dimensions (i.e., *Perception of Behavioral Attributes*, *Perception of the Social Norm*, and *Perception of Behavioral Control*), we hypothesized that our attitude components might predict different types of motivations. Therefore, we examined the predictive power of each individual attitude component on two largely distinct but relevant motivational components of children’s motivation to be curious: children’s *Mastery Orientation Motivation* and their *Performance Avoidance Motivation*. *Mastery*

Orientation Motivation is the desire to achieve competence or an understanding, for the joy and personal use of mastering new tasks (Dweck and Leggett 1988). *Performance Avoidance Motivation* is the desire to avoid performing, for not revealing one’s possible incompetence to others (Darnon et al. 2007; Elliot 1999).

On the basis of the above theories on attitude, curiosity and motivation, we expected that the attitude components that relate to positive perceptions about the personal value of curiosity would positively influence mastery orientation motivation. Thus, we expected scores on the *Personal Relevance*, *Personal Enjoyment*, and *Self-Efficacy* components to positively predict children’s scores on *Mastery Orientation Motivation*. Because our *Societal Relevance* component refers to children’s perception about the relevance of ‘curious people’ to society in general, and not to the relevance of children’s own curiosity, we expected that this attitude component would be either unrelated or positively related to *Mastery Orientation Motivation*. In addition, we expected our *Fear of Negative Judgment* and *Negative Opinion* components to be either unrelated or negatively related to *Mastery Orientation Motivation*, because children with such negative perceptions about epistemic curiosity will likely feel less inclined to show such behavior themselves.

We expected that scores on our *Fear of Negative Judgment* and *Negative Opinion* components would predict children’s performance avoidance motivation, because we hypothesized that children who perceive that (their) epistemic curiosity is inappropriate would avoid expressing their own epistemic curiosity. We expected scores on our *Personal Relevance*, *Personal Enjoyment*, and *Self-Efficacy* components to be negatively related to *Performance Avoidance Motivation*, because children with positive perceptions about the use of epistemic curiosity would probably actively seek more opportunities to engage in curiosity behavior, rather than avoid them. Because the *Societal Relevance* component does not refer to children’s own curiosity, we expected this component to be either unrelated or negatively related to *Performance Avoidance Motivation*.

Finally, some studies suggest that teacher-directed and standardized approaches to learning are limiting children’s natural tendency to be inquisitive learners (Claxton and Carr 2004; Engel 2006; Engel and Randall 2009). As we described in the introduction of this paper, such approaches may easily teach children that there is just one correct answer to teachers’ questions and that being inquisitive is disruptive to classroom instruction (e.g., Post and Walma van der Molen 2018). This effect might be especially prominent among children who transition towards the upper grades of primary school due to teachers’ increasing efforts to help children pass national high-stake tests by means of ‘teaching to these tests’ (Jones et al. 2003). Such classroom practice is generally regarded to leave children with little opportunity

to curiously explore new learning content on their own and may lead them to develop negative notions about the educational value of curiosity-driven learning (Engel 2006, 2015; Post and Walma van der Molen 2018). Although little empirical evidence exists to support this proposition, a prevailing hypothesis is that children's attitudes towards epistemic curiosity worsen as they progress through primary school (Engel 2015). We sought to test this hypothesis in our current study.

Development of the CIAC questionnaire

We followed the framework for construct validity described by Trochim and Donnelly (2006) to develop our questionnaire. Below, we describe in detail what methods we employed to attain and examine construct validity of the CIAC questionnaire.

Establishing translation validity

Content validity

The CIAC questionnaire consists of two separate scales that are based on the image and attitude components described in our theoretical framework: the *Images of Curiosity* scale and the *Attitudes towards Epistemic Curiosity* scale. For each sub-component, we constructed a minimum of four items to allow the removal of possible problematic items later on in the validation process. The original version of the CIAC questionnaire consisted of 41 items.

The *Images of Curiosity* scale includes two hypothesized subscales. The *Social Image of curiosity* subscale aims to measure the extent to which children associate curiosity with questions about social matters, such as wanting to know what you will receive as a birthday present or finding out about other people's personal secrets. The *Epistemic Image of curiosity* subscale was designed to measure to what degree children associate curiosity with cognitive or epistemic questions, such as wanting to know how a computer works or how mathematics was invented. Every question in the *Images of Curiosity* scale has a similar format, stating: 'Suppose you wanted to know how someone obtained knowledge about certain gossip/how the human body works/and so on, indicate how much this has to do with curiosity.' Children's answers were measured on a four-point Likert scale, with response scale options: (1) a very small amount, (2) a small amount, (3) a large amount, (4) a very large amount.

The *Attitudes towards Epistemic Curiosity* scale included six subscales. The first subscale, *Personal Relevance*, was designed to measure to what extent children perceive the relevance or value of posing epistemic questions at school for their own educational development (e.g., whether they find it important for their own learning to pose epistemic

questions). The second subscale, *Personal Enjoyment*, was designed to measure to what degree children enjoy posing epistemic questions about lesson content. The *Societal Relevance* subscale was constructed to measure to what extent children attach societal relevance to epistemic curiosity (e.g., whether curious thinkers foster economic or societal welfare). The fourth subscale, *Negative Opinion*, was constructed to measure whether children hold negative feelings or opinions about other people's epistemic questions (e.g., whether they find curious people to be 'meddlesome' or 'acting smart'). The *Fear of Negative Judgment* subscale was designed to measure the extent to which children fear the negative judgments of others in class (teachers and peers) when expressing epistemic questions or ideas (e.g., the fear that classmates may find you 'nerdy'). The sixth subscale, *Self-Efficacy*, was designed to measure the extent to which children perceive themselves capable of posing epistemic questions about lesson content. Each item in the *Attitudes towards Epistemic Curiosity* scale was formulated as a statement. Children's answers were assessed on a four-point Likert scale, with response scale options: (1) strongly disagree, (2) disagree, (3) agree, (4) strongly agree.

Additional considerations concerning item development

In addition to the specific considerations for item design per subscale, we ensured the clear wording of items and a random sequencing of items in the final format of the questionnaire (Schwarz 2008). In addition, we linked each item statement to primary school children's everyday context. For example, we used children's own reported experiences with 'being curious' about birthday presents or about things happening in class, which were based on a previous study (Post and Walma van der Molen 2018). We also ensured that each statement contained a singular, unambiguous, and appropriate attitude object (Ajzen and Fishbein 2005). For instance, asking children whether they enjoy posing *questions* involves a different 'object' of attitude than asking about their enjoyment of posing *questions to explore new lesson content*.

Likert scale

Although the wording and meaning of the response options differed between the Image and Attitude scales, a four-point Likert response scale was used for both parts of the questionnaire. Although there are different methods for measuring attitudes, such as semantic differential scales, direct interviews, or implicit testing, we chose a Likert scale because this method has several advantages. Likert-scale instruments are suitable when including a large number of items organized in multiple subscales, they can be administered to a large number of respondents, the items can be answered

quickly, and Likert items can be combined to form composite scales that enable parametric testing (Boone and Boone 2012; Schwarz 2008).

We used four-point Likert scales, rather than using an uneven number of response options, because primary school children may misinterpret the midpoint in an uneven scale as a *neutral*, *uncertain*, or *do not know* response, which conceptually differs from a midpoint on a sliding scale (Krosnick and Fabrigar 1997; Kulas and Stachowski 2009). We kept the number of response options to a minimum, since primary school children often hold fairly dichotomous opinions about all sorts of subject-matter and thus may experience difficulty in responding to scales with too fine-grained response options (Mellor and Moore 2014). Four-point Likert response scales have been used effectively in other studies that investigated children's images and attitudes (for examples of such questionnaires, please see Frantom et al. 2002; Block 1995; Post and Walma van der Molen 2014).

Face validity

After we formulated the items for each hypothesized subscale, we asked several experts in the fields of questionnaire design and attitude research to indicate whether the items were clearly formulated and representative of their corresponding subscales. In addition, several teachers and children from two primary schools that did not take part in the validation study were asked to assess the comprehensibility, clarity, and appropriateness of the items for our target group. Based on the comments we received from the experts, teachers, and children, we identified minor issues with some items and made necessary improvements. Overall, this qualitative pilot study aided us in improving the face validity of the CIAC.

Validation study for the CIAC questionnaire

Respondents

We administered the CIAC questionnaire among a large sample of children ($N = 737$) in the 4th, 5th and 6th grades of nine public primary schools in The Netherlands. The schools were located in different school districts, but were selected on the basis of similar SES background characteristics. The effective sample size included 369 boys (50.1%) and 367 girls (49.9%). Respondents' ages ranged from 8 to 13 years ($M = 10.61$; $SD = .99$).

Six of the nine schools participated in the study as part of a larger school development project that concerned a 6-month, school-wide teacher-training program on the topic of inquiry-based teaching. The questionnaire was administered around the time the program had started. The

remaining three schools did not participate in the teacher-training program, but took part in the study on a voluntary basis.

Procedure and materials

A paper-and-pencil version of the questionnaire was administered by a research assistant to all the children in their own classroom during normal school hours. After an introduction for the whole class by the assistant, the children were given the time needed, about 10 min, to complete the questionnaire. If a child did not understand a particular item, the researcher provided feedback individually. The objectives, time requirements, and nature of the data collection procedure were explained to the school management, teachers, and parents a few weeks before the start of the data collection. Informed consent was obtained from the parents in accordance with the ethical guidelines of our university.

Data analysis

The criterion validity of the two scales of the CIAC was investigated in several consecutive steps. First, we checked for missing data. We also investigated the discriminant power of each item by evaluating the range of responses and the standard deviation of respondents' scores on each item. Next, using Mplus version 7.4 (Muthén and Muthén 1998–2015), we performed an exploratory factor analysis (EFA) on the raw (ordinal) data for both the *Image* and *Attitude* scales, using the option 'Categorical' and weighted least squares estimation with means and variance adjusted (WLSMV) and Geomin oblique rotation. Since the CIAC questionnaire had been newly developed on the basis of a new theoretical framework, we first explored and identified the latent factors underlying children's images of and attitudes towards curiosity by using EFA (Prudon 2015). This approach also helped us to remove any problematic items from the questionnaire before conducting a confirmatory factor analysis (CFA). Unlike EFA, CFA allows for testing model fit, that is, how well the observed data fit a pre-defined hypothesized factor structure (for an extensive review of best practices using both EFA and CFA for instrument validation, please see Prudon 2015; Schmitt 2011; Worthington and Whittaker 2006). CFA was also performed using Mplus and WLSMV estimation.

While Mplus does provide model fit estimations, it does not provide direct information about the convergent and discriminant validity of the subscales under investigation (Carter 2016). Therefore, we followed the computational formulas provided by Fornell and Larcker (1981) and Raykov (1997) to calculate the necessary additional measures of average variance extracted (AVE), composite reliability (CR), average shared variance (ASV), and maximum

shared variance (MSV). Convergent validity of a subscale is considered satisfactory when AVE is equal to or greater than .50 (i.e., the amount of shared variance among items that belong to a subscale). In addition, the CR value of a subscale should be equal to or greater than .70 and greater than the AVE value of the subscale (Fornell and Larcker 1981). Discriminant validity of a subscale is met when the AVE of the subscale is greater than its MSV and greater than its ASV.

The EFA and CFA analyses were conducted with two different subsamples of the data. A random sampling procedure was used to extract subsample 1 ($n=368$) and subsample 2 ($n=369$) from the total respondent group. Equivalence of the subsamples was investigated for gender and grade using Chi square tests. Results indicated that boys and girls ($\chi^2=.60$, $df=1$, $p=.46$) and children from different grades ($\chi^2=2.20$, $df=2$, $p=.33$) were equally distributed across both subsamples.

As will be described in “Results” section, we also assessed the ability of the CIAC questionnaire to make fair comparisons between the image and attitude scores of children from different grade levels, since we assume that the CIAC questionnaire measures the same constructs the same way for children across different grade levels. To test this assumption, we performed a confirmatory factor analytic test of measurement invariance (Vandenberg and Lance 2000).

Finally, we assessed the predictive validity of the CIAC by examining the extent to which children’s scores on the *Epistemic Image* subscale predicted their scores on each separate attitude component and the extent to which children’s attitudes scores predicted their scores on *Mastery Orientation Motivation* and *Performance Avoidance Motivation*. To this end, we used the validated Achievement Goal Questionnaire by Elliot and McGregor (2001) to assess children’s goal achievement motivations (see also Pekrun et al. 2011). Predictive validity was examined by use of structural equation modelling (SEM) through fitting our hypothesized models of relationships among the image, attitude and motivation variables to our acquired data (Muthén and Muthén 1998–2015). In addition, using multivariate analysis of variance (MANOVA) in SPSS version 24, we examined children’s overall attitude scores and tested whether the children in Grade 6 showed decreased attitude scores in comparison to the children in lower grades.

Results

Preliminary data checks

We carried out preliminary data checks to detect the presence of missing data. The percentage of missing data was 2.77% in subsample 1 and 2.93% in subsample 2. Little’s (1988) MCAR test results indicated that the

missing data were missing completely at random in subsample 1 ($\chi^2=3394.66$; $df=3277$; $p=.07$) and subsample 2 ($\chi^2=2906.911$; $df=2901$; $p=.47$). Both EFA and CFA were performed with the raw data from subsample 1 and 2 that included these missing data, using the default procedure for handling missing data in MPlus. We also examined the discriminant power of each item by computing the standard deviation and range of responses. For each item, the standard deviation should hover around 1 and each response option should be used at least once (Coulson 1992; Schwarz 2008). The standard deviations of all items in our questionnaire ranged between .77 and .97 and all response options were used at least once, indicating sufficient discriminant power. The data were thus considered adequate for subsequent factor analyses.

Exploratory factor analysis

Iterative exploratory factor analyses with WLSMV were conducted on subsample 1 to investigate the latent factor structure in the data. Since we expected the subscales to correlate, Geomin oblique rotation was used to determine the best rotated solution (Reise et al. 2000). No maximum number of factors was preset. Items were omitted from consecutive EFAs if they showed a factor loading lower than .40 or cross loadings less than a .15 difference from an item’s greatest factor loading (Floyd and Widaman 1995; Hair et al. 2006; Worthington and Whittaker 2006).

Images of curiosity

We first performed EFA on the items of the *Images of Curiosity* scale. Two items that belonged to the *Social Image* subscale were removed from the questionnaire because of poor factor loadings ($<.40$). Subsequent EFA analysis revealed the presence of two factors with Eigenvalues above 1. As shown in Table 1, all items designed to address each of the two image components loaded onto one particular factor, resulting in a two-factor structure that corresponds to the two hypothesized subscales of the questionnaire design. Factor one (Eigenvalue = 2.94) contains the four items that refer to children’s *Social Image* of curiosity. Factor two (Eigenvalue = 1.66) includes the five items that refer to children’s *Epistemic Image* of curiosity. The factor loadings ranged between .48 and .76. Please see Table 1 for the obtained factor structure and factor loadings of the *Images of Curiosity* scale.

Attitudes towards epistemic curiosity

We conducted separate exploratory factor analyses for the items on the *Attitudes towards Epistemic Curiosity* scale. After several EFA iterations, three items were removed

Table 1 Factor structure solution for the Images of Curiosity scale as obtained by exploratory and CFA

Indicate how much curiosity has to do with...	EFA ^a		CFA ^b	
	Social	Epistemic	Social	Epistemic
Wanting to know how someone found out about a secret	.76		.81	
Wanting to know how someone found out about a certain rumor	.60		.76	
(D) Wanting to figure out what others are thinking or feeling	.50		–	
(D) Eavesdropping on private conversations to figure out what they’re talking about	.48		–	
Wanting to know how the human body works		.73		.66
Wanting to know how a car works		.69		.76
Wanting to know how computers were invented		.57		.63
Wanting to know how birds are able to fly		.56		.58
Wanting to do how math was invented		.55		.62

Factor loadings are only displayed for items with loadings > .40 on their expected factors and with cross loadings > .15 difference from the item’s greatest factor loading.

(D): indicates that the item was eventually deleted from the questionnaire due to poor factor loading (< .40) based on results obtained by CFA

^aEFA conducted with weighted least squares estimation with means and variance adjusted and Geomin rotation with subsample 1. Values represent rotated factor loadings

^bCFA conducted with weighted least squares estimation with means and variance adjusted with subsample 2. Values represent factor loadings

from the questionnaire because of poor factor loadings or cross loadings, leaving a total of 18 items. The final factor solution consisted of five factors with Eigenvalues above 1. The factor loadings ranged between .50 and .80. However, the observed factor structure differed in two ways from our hypothesized factor structure. Below, we outline these differences and describe what adaptations we made to our subscales on the basis of the EFA.

First, while we hypothesized that the *Personal Relevance* and *Personal Enjoyment* subscales would exist as independent subscales, the EFA revealed that both constructs loaded on a joint factor. In line with the TPB, this result may not be too surprising, since the TPB conceptualizes both cognitive and affective perceptions as part of an overarching concept of ‘attitudinal perceptions of behavioral attributes’ (Ajzen 1991). Unlike the *Societal Relevance* factor—which refers to the value that children may attribute to people’s epistemic curiosity for society *in general*—the *Personal Relevance* and *Personal Enjoyment* factors both refer to children’s perceptions of the value of epistemic curiosity for their *own* development. Therefore, in accordance with the obtained EFA factor structure, we decided to unify the hypothesized *Personal Relevance* and *Personal Enjoyment* subscales into a revised, joint subscale, which we labeled *Personal Inclination*. The *Personal Inclination* subscale contains seven items that cover possible cognitive and affective aspects of the learning-related value that a child may attach to his or her own epistemic curiosity.

Second, the EFA revealed that our hypothesized *Fear of Negative Judgment* subscale seemed to be more accurately described in terms of children’s *Fear of Classmates’*

Negative Judgment (Eigenvalue = 1.05), because the three items of the original subscale that survived the EFA relate particularly to children’s fear of their *classmates’* judgments, rather than possible fears of their *teacher’s* judgments. The items about teachers’ judgments did not load onto a distinct factor of their own.

On the basis of the factor structure obtained by the EFA, we derived the following five attitude subscales: *Personal Inclination* (revised) (Eigenvalue = 4.98); *Societal Relevance* (Eigenvalue = 1.42); *Negative Opinion* (Eigenvalue = 1.27); *Fear of Classmates’ Negative Judgment* (Eigenvalue = 1.05); and *Self-Efficacy* (Eigenvalue = 2.62). Please see Table 2 for the factor structure and factor loadings of the *Attitudes towards Epistemic Curiosity* scale.

Confirmatory factor analysis

We performed CFA with WLSMV to assess how well the data of subsample 2 fitted the obtained factorial structures for the image and attitude subscales that we derived by EFA. To determine model fit, we examined multiple goodness-of-fit indices. Because the χ^2 statistic is highly sensitive to large respondent groups, we supplemented the conventional model fit indices with the absolute model fit estimate weighted root mean square residual (WRMR) (Prudon 2015; Worthington and Whittaker 2006; Yu 2002). The following model fit indices were used to examine whether the factor structures of our scales fitted the data: (1) the WRMR, (2) the Comparative Fit Index (CFI), (3) the Tucker–Lewis Index (TLI), and (4) the RMSEA (Floyd and Widaman 1995). WRMR should be below 1.0 to indicate good fit, CFI and TLI values should

Table 2 Factor structure solution for the Attitudes towards Epistemic Curiosity scale as obtained by exploratory and CFA

Item	EFA ^a					CFA ^b				
	PI	SR	NO	FCNJ	SE	PI	SR	NO	FCNJ	SE
I really like to wonder about all the things I learn at school	.72					.67				
It is very important to me to come up with interesting questions at school, because then I learn more about the things around me	.70					.65				
It is very important to me to wonder about lots of things in class, because then I learn more about all sorts of different things	.59					.70				
I really like to ask questions about all sorts of things in class	.55					.62				
I think people who want to know a lot are very important to the economy of the Netherlands	.77						.67			
I think people who ask good questions have a big impact on society	.66						.65			
I think people who often come up with interesting questions are very important to society	.65						.77			
I feel classmates are being stubborn when they always want to know all about everything in class			.67					.60		
I find classmates to be annoying when they ask a lot of smart questions in class			.65					.64		
I feel people who ask a lot of questions come across as dumb			.60					.69		
I'm afraid that my classmates will think I'm a nerd if I ask a lot of smart questions in class			.79						.71	
I'm afraid that my classmates will think it's stupid if I want to know more about something we're learning in class			.71						.77	
I find it scary to show that I'd like to know more about a topic in class			.58						.69	
I am really good at coming up with smart questions in class				.80						.67
I am really good at coming up with new questions about all sorts of topics in lessons at school				.69						.76
I think I am really good at figuring out new things at school				.53						.76
I am really good at coming up with smart questions about all sorts of subjects at school				.50						.65

Factor loadings are only displayed for items with loadings > .40 on their expected factors and with cross loadings > .15 difference from the item's greatest factor loading.

PI Personal Inclination subscale, SR Societal Relevance subscale, NO Negative Opinion subscale, FCNJ Fear of Classmates' Negative Judgment subscale, SE Self-Efficacy subscale

^aEFA conducted with weighted least squares estimation with means and variance adjusted and Geomin rotation with subsample 1. Values represent rotated factor loadings

^bCFA conducted with weighted least squares estimation with means and variance adjusted with subsample 2. Values represent factor loadings

exceed .95, and RMSEA values should be below .08 (Hu and Bentler 1999; Mueller and Hancock 2008; Yu 2002). It should be noted that the above model fit threshold values are simply guidelines and should not be interpreted as strict rules (Prudon 2015). In addition, we calculated the AVE, CR, MSV and ASV values for each subscale as obtained by CFA to assess their convergent and discriminant power (Carter 2016).

Images of curiosity

We performed CFA on the basis of the factor structure solution that had been previously obtained by EFA. The CFA results revealed some additional poor factor loadings for items that belonged to the *Social Image* subscale. These items were then removed from the analysis, leaving a total of seven items for the *Images of Curiosity* scale, two items for the *Social Image* subscale and five items for the *Epistemic Image* subscale. Although more than two items should ideally represent a subscale, the use of two items per subscale is considered satisfactory when the items are fairly strongly correlated by measure of Spearman–Brown correlation (Eisinga et al. 2013; Worthington and Whittaker 2006). The two items of the *Social Image* subscale are reasonably correlated

($r = .56$), so we decided to keep this factor with these two items. Subsequent CFA of the revised *Images of Curiosity* scale confirmed good model fit (WRMR = .60; CFI = .99; TLI = .98; RMSEA = .06). Please see Table 1 for the factor structure and factor loadings of the *Images of Curiosity* scale as obtained by CFA. Inspection of the factor correlation matrix (see Table 3) showed that the *Social Image* and *Epistemic Image* of curiosity subscales were marginally correlated ($r = .38$). As depicted in Table 3, AVE values for the *Social Image* subscale (AVE = .74) and the *Epistemic Image* subscale (AVE = .55) indicated sufficient convergent validity of the subscales. In addition, CR values for both subscales exceeded the threshold value of .70 and exceeded the AVE values for each respective subscale. The discriminant power of the subscales was also found to be sufficient, as indicated by AVE values that exceeded MSV and ASV.

Attitudes towards epistemic curiosity

We performed CFA on the basis of the factor structure solution that had been previously obtained by EFA. The CFA results confirmed good model fit (WRMR = .84; CFI = .98; TLI = .97; RMSEA = .04). Please see Table 2 for the factor structure and factor loadings of the *Attitudes towards*

Table 3 Factor correlations, CR, AVE, MSV, ASV, scale means and standard deviations for the Image and Attitude subscales as obtained by CFA

	Images of Curiosity		Attitudes towards Epistemic Curiosity				
	Social Image	Epistemic Image	Personal Inclination	Societal Relevance	Negative Opinion	Fear of Classmates' Negative Judgment	Self-Efficacy
Images of Curiosity							
Social Image	–						
Epistemic Image	.38**	–					
Attitudes towards Epistemic Curiosity							
Personal Inclination			–				
Societal Relevance			.70**	–			
Negative Opinion			–.38**	–.28**	–		
Fear of Classmates' Negative Judgment			–.07	.11	.62**	–	
Self-Efficacy			.77**	.60**	–.17*	–.03	–
Composite reliability (CR)	.76	.79	.76	.74	.68	.77	.80
Average variance extracted (AVE)	.74	.55	.56	.61	.57	.66	.63
Maximum shared variance (MSV)	.14	.14	.59	.49	.38	.38	.59
Average shared variance (ASV)	.14	.14	.31	.24	.16	.10	.25
Subscale mean	2.98	2.67	2.65	2.71	1.94	1.82	2.46
Subscale standard deviation	.77	.62	.62	.62	.67	.68	.64

Subscale means are weighted averages and scores could range between 1 and 4

*Factor correlation is statistically significant at $p < .05$

**Factor correlation is statistically significant at $p < .01$

Epistemic Curiosity scale as obtained by CFA. However, our subsequent examination of the convergent and discriminant power of the attitude subscales did reveal two minor issues. As for the convergent power of individual subscales, Table 3 shows that while the AVE and CR values of most of the attitude subscales exceeded the threshold values of .50 and .70 respectively (thus indicating that the subscales possess sufficient convergent power), the CR value of the *Negative Opinion* subscale (.68) was slightly lower than .70.

As for the discriminant power of the individual attitude subscales, Table 3 shows that the AVE did exceed the ASV, but not the MSV, for the *Personal Inclination* and *Self-Efficacy* subscales. This result indicates that both subscales share considerable variance, which may suggest that the items of the *Personal Inclination* and *Self-Efficacy* subscales were interpreted by children as being conceptually similar. However, because MSV only marginally exceeded AVE and the discriminant power of the subscales was corroborated by adequate convergent power, model fit and conceptual design, we think that the amount of shared variance of the *Personal Inclination* and *Self-Efficacy* subscales is acceptable.

The factor correlation matrix (see Table 3) revealed that the subscales correlated as we expected. The positive attitude components, *Personal Inclination*, *Societal Relevance*, and *Self-Efficacy*, all showed statistically significant positive correlations. Similarly, the negative attitude components, *Negative Opinion* and *Fear of Classmates' Negative Judgment*, also showed statistically significant positive inter-correlations. In addition, as hypothesized, the factor correlation matrix shows that positive and negative attitude subscales were either unrelated (insignificant factor correlations) or negatively correlated. These findings further add to the construct validity of the CIAC questionnaire.

Measurement invariance analysis

Next, we conducted multiple-group CFA to test measurement invariance of the CIAC questionnaire across the 4th ($n = 214$), 5th ($n = 242$) and 6th ($n = 281$) grade groups, using the total respondent sample ($N = 737$). We examined

configural invariance (i.e., equality of factor structures) and metric invariance (i.e., equality of factor loadings) of the observed data for both the *Images of Curiosity* and *Attitudes towards Epistemic Curiosity* scales. To assess configural invariance, we allowed item parameters (i.e., factor loadings, item intercepts, and item uniqueness), factor variances, and latent means to vary freely across groups. To assess metric invariance, we fixed the factor loadings of the factor structure across groups.

We examined changes in CFI (ΔCFI) and χ^2 ($\Delta\chi^2$) as our primary tests of measurement invariance, where a ΔCFI less than or equal to .01 indicates invariance and non-significant p -values ($< .05$) for the $\Delta\chi^2$ between the measurement models indicate invariance (Cheung and Rensvold 2002). As presented in Tables 4 and 5, for each of our measurement invariance tests per grade level group comparison, the ΔCFI did not exceed .01 and the p -values for $\Delta\chi^2$ were non-significant. We therefore conclude that both the *Images of Curiosity* and *Attitudes towards Epistemic Curiosity* scales may be used to compare scores across the 4th, 5th, and 6th grade groups.

Predictive validity

Predictive validity of images on attitudes

We assessed the ability of the *Epistemic Image* subscale to predict children's attitude scores, as described in our Hypothesis section. Predictive validity was examined by testing the hypothesized structural model using SEM analysis, which indicated good model fit (WRMR = 1.07; CFI = .97; TLI = .96; RMSEA = .04). Table 6 summarizes the observed regression coefficients among the image and attitude components. As hypothesized, scores on the *Epistemic Image* component significantly predicted scores on the attitude components *Personal Inclination* ($R^2 = .19$, $p < .01$), *Societal Relevance* ($R^2 = .12$, $p < .01$), and *Self-Efficacy* ($R^2 = .13$, $p < .01$). In addition, the *Epistemic Image* component negatively predicted the *Negative Opinion* component ($R^2 = .02$, $p = n.s.$). The *Epistemic Image* component

Table 4 Measurement invariance results of the Images of Curiosity scale for the total respondent sample, with grade level as the group comparison

Group comparison	χ^2	df	TLI	CFI	RMSEA	ΔCFI	$\Delta\chi^2$	Δdf	p
Grade 4 versus Grade 5									
Configural invariance	34.835	26	.99	.99	.04	–	–	–	–
Metric invariance	43.947	31	.98	.99	.04	.00	8.702	5	.12
Grade 5 versus Grade 6									
Configural invariance	45.023	26	.99	.99	.05	–	–	–	–
Metric invariance	47.597	31	.99	.99	.05	.00	4.773	5	.44
Grade 4 versus Grade 6									
Configural invariance	48.735	26	.98	.99	.06	–	–	–	–
Metric invariance	53.414	31	.98	.99	.05	.00	6.786	5	.24

Table 5 Measurement invariance results of the Attitudes towards Epistemic Curiosity scale for the total respondent sample, with grade level as the group comparison

Group comparison	χ^2	<i>df</i>	TLI	CFI	RMSEA	Δ CFI	$\Delta\chi^2$	Δ <i>df</i>	<i>p</i>
Grade 4 versus Grade 5									
Configural invariance	375.959	218	.95	.96	.06	–	–	–	–
Metric invariance	390.616	230	.95	.96	.06	.00	18.735	12	.10
Grade 5 versus Grade 6									
Configural invariance	394.665	218	.96	.97	.06	–	–	–	–
Metric invariance	409.750	230	.96	.96	.06	.01	18.154	12	.11
Grade 4 versus Grade 6									
Configural invariance	429.956	218	.94	.95	.06	–	–	–	–
Metric invariance	442.284	230	.94	.95	.06	.00	18.877	12	.09

Table 6 Summary of regression analyses for the Image components predicting the Attitude components (*N*=737)

	Personal Inclination			Societal Relevance			Negative Opinion			Fear of Classmates' Negative Judgment			Self-Efficacy		
	<i>B</i>	SE <i>B</i>	β	<i>B</i>	SE <i>B</i>	β	<i>B</i>	SE <i>B</i>	β	<i>B</i>	SE <i>B</i>	β	<i>B</i>	SE <i>B</i>	β
Social Image	–.01	.05	–.01	.09	.05	.10	.04	.05	.05	.02	.05	.02	–.01	.05	–.01
Epistemic Image	.46	.05	.43**	.32	.06	.32**	–.15	.05	–.16**	.01	.05	.01	.39	.05	.37**

*Beta is statistically significant at *p* < .05

**Beta is statistically significant at *p* < .01

was unrelated to the *Fear of Classmates' Negative Judgment* component ($R^2 = .00, p = \text{n.s.}$). The *Social Image* component did not predict children's scores on any of the attitude components ($R^2 < .04, p > .06$), which confirms the hypothesis that children's attitudes towards epistemic curiosity are independent of children's social images of curiosity. These findings thus reveal that the *Social Image* and *Epistemic Image* components predicted results that they were hypothesized to predict, thereby supporting the predictive validity of the *Images of Curiosity* scale.

Predictive validity of attitudes on motivations

We examined the ability of the *Attitudes towards Epistemic Curiosity* subscales to predict scores on *Mastery Orientation Motivation* and *Performance Avoidance Motivation*. In preparation of this analysis, we first performed CFA to test the supposed two-factor structure underlying the AGQ items from Elliot and McGregor (2001). The CFA results confirmed that the items that belonged to the *Mastery Orientation Motivation* scale and the items that belonged to the *Performance Avoidance Motivation* scale fitted the data well (WRMR = .99; CFI = .99; TLI = .98; RMSEA = .06). Tests of convergent power of each motivation scale proved to be sufficient, as indicated by AVE and CR values that were greater than .50 and .70 respectively. In addition, we assessed the discriminant power of the motivation scales compared to the attitude subscales to exclude the possibility of multicollinearity. The results showed that the amount of shared

variance among the separate motivation scales and attitude subscales was acceptable, as indicated by AVE values of each motivation scale that exceeded MSV and ASV. In sum, these preparatory analyses showed that the motivation scales can be regarded as sufficiently independent from each other and from the attitude subscales and that, thereby, these latent variables are suitable to be part of a measurement model to test our hypothesized structural models (Prudon 2015).

To examine the extent to which scores on the *Attitudes towards Epistemic Curiosity* components predicted children's *Mastery Orientation Motivation* and *Performance Avoidance Motivation*, we tested a unified structural model that related each attitude component to each motivation factor. Table 7 summarizes the observed regression coefficients among the attitude components and motivation factors. The observed statistical relationships among the attitude components and the *Mastery Orientation Motivation* factor ($R^2 = .50, p < .01$) and the relationships among the attitude components and the *Performance Avoidance Motivation* factor ($R^2 = .33, p < .01$), largely confirmed our hypotheses (WRMR = 1.21; CFI = .96; TLI = .96; RMSEA = .04). Scores on the *Personal Inclination* component significantly predicted scores on *Mastery Orientation Motivation*. In addition, the components *Societal Relevance*, *Negative Opinion*, and *Fear of Classmates' Negative Judgment* related to *Mastery Orientation Motivation* as expected as well. However, in contrast to our predictions, the *Self-Efficacy* component showed no statistically significant relationship with *Mastery Orientation Motivation*.

Table 7 Summary of regression analyses for the Attitude components predicting the Motivation components ($N = 737$)

	Mastery Orientation Motivation			Performance Avoidance Motivation		
	<i>B</i>	SE <i>B</i>	β	<i>B</i>	SE <i>B</i>	β
Personal Inclination	.81	.12	.73**	-.27	.12	-.24*
Societal Relevance	-.09	.08	-.07	-.18	.08	-.15*
Negative Opinion	-.18	.09	-.13*	.38	.09	.28**
Fear of Classmates' Negative Judgment	-.10	.07	-.09	.22	.07	.19**
Self-Efficacy	-.09	.09	-.07	.09	.09	.07

*Beta is statistically significant at $p < .05$ **Beta is statistically significant at $p < .01$

As hypothesized as well, the components *Personal Inclination* and *Societal Relevance* showed to be negatively related to *Performance Avoidance Motivation*. In addition, the scores on the *Negative Opinion* and *Fear of Classmates' Negative Judgment* components significantly predicted children's *Performance Avoidance Motivation*. However, contrary to our predictions, the *Self-Efficacy* component showed no statistically significant relationship with *Performance Avoidance Motivation*. In sum, these findings reveal that, for the most part, children's attitudes towards curiosity predict their mastery orientation and performance avoidance motivations and thereby support the predictive validity of the *Attitudes towards Epistemic Curiosity* scale.

Children's attitude scores per grade level

Lastly, we investigated the degree to which children's attitude scores declined as a function of children's grade level. To this end, we first computed children's weighted sum scores for each attitude subscale per grade level (see Table 8). Notably, these data reveal that, on average, children's scores for the *Personal Inclination*, *Societal Relevance*, and *Self-Efficacy* components lied around the scale's midpoint (2.5), which indicates that children's attitudes towards these matters in the present sample,

irrespective of their grade level, were generally only moderate at best. Contrary, children's scores on *Negative Opinion* and *Fear of Classmates' Negative Judgment* lied below the scale's midpoint, which indicates that, on average, children did not so much perceive their classmates' epistemic curiosity behavior in negative terms, nor did they perceive that their classmates negatively judged their epistemic curiosity behavior.

We further examined differences in children's attitude scores per grade level by performing MANOVA with grade level as the between-subject factor and the five attitude components as multivariate dependent variables. The omnibus test of between-subject effects using Wilk's statistic revealed a significant main effect of grade, $\lambda = .93$, $F(10, 1436) = 5.385$, $p = .00$, $\eta^2 = .04$. As shown in Table 8, post-hoc univariate analyses for each separate attitude component confirmed statistically significant but small differences only between the attitude scores of children from Grade 6 with the children from either Grade 4 or Grade 5 for the attitude components *Personal Inclination*, *Societal Relevance*, and *Self-Efficacy*. In these cases, the attitude scores of children from Grade 6 were somewhat lower than the scores of children from Grade 4 and Grade 5. This result suggests that, on the basis of the present data sample, children's perceptions about the learning value and use of (their) epistemic

Table 8 Comparison of children's weighted sum scores for each separate attitude component per grade level

Attitude subscales	Grade 4 ($n = 214$)		Grade 5 ($n = 242$)		Grade 6 ($n = 281$)		Grade 4 versus Grade 5		Grade 5 versus Grade 6		Grade 4 versus Grade 6	
	<i>M</i>	SD	<i>M</i>	SD	<i>M</i>	SD	η^2	<i>p</i>	η^2	<i>p</i>	η^2	<i>p</i>
Personal Inclination	2.81	.64	2.69	.59	2.51	.62	–	n.s.	.03	.00	.05	.00
Societal Relevance	2.83	.64	2.75	.62	2.52	.65	–	n.s.	.03	.00	.05	.00
Negative Opinion	1.97	.70	1.84	.66	1.93	.67	–	n.s.	–	n.s.	–	n.s.
Fear of Classmates' Negative Judgment	1.83	.76	1.83	.67	1.79	.64	–	n.s.	–	n.s.	–	n.s.
Self-Efficacy	2.57	.68	2.45	.62	2.27	.61	–	n.s.	.02	.01	.05	.00

Weighted sum scores could range between 1 (strongly disagree) to 4 (strongly agree)

curiosity may indeed decline as they progress through primary school.

Discussion

This study focused on the development and validation of a questionnaire to measure primary school children's images of curiosity and their attitudes towards epistemic curiosity. To the best of our knowledge, the development of the CIAC is the first attempt to measure these aspects in children. Based on curiosity and attitude research, we expect that a fruitful approach to stimulating children's epistemic wonderment, questions, and ideas in the classroom is to foster their positive, epistemic images of curiosity and their positive beliefs and feelings about being curious learners in school. To measure these images and attitudes validly and reliably over time, we developed and tested the CIAC.

Because the CIAC questionnaire and its underlying theoretical components had not been empirically tested before, we employed both qualitative and extensive quantitative methods to verify the construct validity of the CIAC questionnaire. We also assessed the ability of the CIAC questionnaire to measure the same image and attitude components the same way for children across the 4th, 5th and 6th grades (Hirschfeld and Brachel 2014; Milfont and Fischer 2010; Vandenberg and Lance 2000).

Main findings of the study

The results of our study provide good evidence to support the construct validity of the CIAC. Although our results showed some minor deviations from the constructs that we originally hypothesized, overall, our data fitted the main underlying dimensions of the TPB well. Furthermore, results showed that the CIAC demonstrated full configural and metric measurement invariance for children across the middle and upper grades of primary school. In addition to the CIAC's internal validity, the observed relationships between the image, attitude, and motivation variables proved to be largely consistent with theory (e.g., Eagly and Chaiken 1993; Osborne et al. 2003) and thus provided support for the predictive validity of the CIAC as well. Also, in accordance with theories proposed by others in the field (e.g., Engel 2006; Jirout and Klahr 2012), the attitude scores revealed that the children in our sample generally did not hold very positive perceptions about the value and use of being curious about new subject matter in school and that the children in Grade 6 generally felt less positive about these matters than the children in the lower grades. Below, we discuss some issues that might improve certain aspects of the CIAC.

Images of curiosity

While EFA revealed that the items of the *Social Image* subscale all loaded on the specified factor, CFA indicated that two of the four items loaded poorly ($< .40$), leaving a total of only two items for the *Social Image* subscale. This finding stresses the importance of employing CFA as well as EFA to thoroughly assess the factor structures obtained by EFA. Research indicates that it is not ideal to represent a subscale by only two items (Worthington and Whittaker 2006). Thus, the number of items of the *Social Image* subscale should be increased in an improved version of the CIAC, with items that better represent children's social images of curiosity. It is plausible to assume that this addition of items will lead to improved convergent and discriminant power of the *Social Image* subscale. Based on previous research by Litman and Pezzo (2007) on the measurement of social curiosity, examples of such items may be: 'Suppose you wanted to figure out what others are thinking or feeling, indicate how much this has to do with curiosity'; 'Suppose you try to figure out what someone is hiding from you, indicate how much this has to do with curiosity'; 'Suppose you eavesdrop on a private conversation to figure out what is being talked about, indicate how much this has to do with curiosity'.

Attitudes towards epistemic curiosity

Our factor-analytic examination of the *Attitudes towards Epistemic Curiosity* scale indicated that most of our developed items loaded on their expected latent factors. However, EFA did produce some unexpected results. First, the EFA indicated that the items that were originally developed for the *Personal Relevance* and *Personal Enjoyment* subscales loaded on one, joint factor. As we already described in "Results" section, this finding is in line with the proposition in the TPB that the attitudinal dimension 'Perceptions of Behavioral Attributes' comprises both cognitive and affective perceptions (Ajzen 1991). Therefore, we decided to continue our analyses with a revised scale, which we labeled *Personal Inclination*. CFA confirmed that the *Personal Inclination* subscale fitted the data well.

Second, the results of the EFA indicated that *Fear of Negative Judgment* predominantly consisted of children's *Fear of Classmates' Negative Judgment*. In retrospect, it is not surprising that children may attribute different social norms for 'being curious' in class to their peers or to their teachers. This finding again underlines the value of employing both EFA and CFA. CFA indicated that the *Fear of Classmates' Negative Judgment* fitted the data well and that the subscale possessed sufficient convergent and discriminant power.

However, this result does not necessarily imply that children *only* fear their classmates' judgments. To investigate whether an improved version of the CIAC should include a

separate scale that measures children's fears of their teachers' judgments, we conducted follow-up interviews with groups of children from the 4th, 5th, and 6th grades from one primary school that participated in the present study. Each group of children consisted of two boys and two girls who were randomly selected from each grade. We first asked the children to what extent they ever experienced having epistemic questions or ideas about lesson topics in class. Second, we asked them if they ever felt afraid of posing their curious questions in class. All children reported having regular curious questions about lesson topics in class. About half of the children in the sample indicated being afraid to express such curious questions. Of these children, all of them indicated that they predominantly feared their peers as likely to make fun of their curiosity. In contrast, the children in this sample did not fear the possibility of their teacher judging their curious questions or ideas in a negative way. Given these additional findings, we propose that children's fear of classmates' negative judgments is indeed an important underlying attitudinal dimension that may hinder children's curiosity behavior, while their fear of teachers' negative judgments seems to be a less important underlying attitudinal factor.

Lastly, the observed relationships among the attitude and motivation variables largely confirmed our hypotheses, with *Self-Efficacy* as the only exception. Children's self-efficacy scores appeared to be unrelated to their mastery orientation motivation and performance avoidance motivation. This result is surprising, because self-efficacy is found to be a central predictor of motivation in a wide range of studies on attitude (Ajzen 1991; Eagly and Chaiken 1993; Olson and Zanna 1993) and achievement (Bandura 1997; Pekrun 2006). A tentative explanation could be that the children in our sample had little experience with curiosity-focused or inquiry learning. An explorative study by Post and Walma van der Molen (2018) suggests that children may indeed only barely engage in curiosity-driven learning in primary school. This may have made it difficult for the children to rate their efficacy and might explain why children's self-efficacy scores did not relate to their motivation scores.

Directions for future research

An improved version of the CIAC should be re-validated by administering the questionnaire again among a large group of primary school children. In addition, it should be noted that the current version of the CIAC was developed for children in the 4th, 5th, and 6th grade. We expect that the CIAC is not suited for children younger than 8 or 9 years old, because the survey items might be too demanding for younger children to comprehend. However, we do expect that the CIAC is useful for measuring older children's images and attitudes (up to 14 or perhaps even 15 years old),

which would allow researchers and educators to use the CIAC to track children's developing images and attitudes over longer periods of time. Nonetheless, this expectation needs to be tested as well, by means of factor-analytic tests of measurement invariance that examine the equality of factor structures and factor loadings of children across these different age groups.

The current version of the CIAC consists of Dutch items, which were developed and refined on the basis of the feedback that we received from Dutch primary school children and teachers. Although we have carefully translated the items of the CIAC from Dutch to English, cross-cultural validation studies should determine the construct validity of the CIAC for children in other countries. To this end, we invite fellow researchers to collaborate with us and validate a translated and improved version of the CIAC in their country.

It is also worth noting that the attitude components *Societal Relevance* and *Fear of Classmates' Negative Judgment* both appeared to be unrelated to *Mastery Orientation Motivation*, as we hypothesized. In our view, the absence of these relationships may be explained by the fact that both components refer to children's perceptions about the effects that *other people's* behavior may have, rather than to the effects of children's own behavior. Nevertheless, we do believe that both attitude components constitute important elements of children's attitudes towards curiosity. To assess the predictive power of these attitude components on motivation, however, we suspect that other types of goal motivational measures might be more appropriate, such as performance approach motivation (Reeve 2015). For example, children may feel driven to ask questions or seek new knowledge to meet normative performance standards or to outperform others. With the addition of this measure, one could also examine the ability of the CIAC to differentiate between children who seek to attain knowledge for their own benefit (i.e., mastery orientation motivation) and those who seek to perform well (i.e., performance approach motivation). Due to practical limitations of the current study, however, we did not include such measures. Future studies may investigate these possible relations further.

Future research could also examine which components of the CIAC are most important for fostering epistemic curiosity in children. For example, it could be that children's curiosity is best fostered by first attending to their perceived value of epistemic curiosity for their own learning (*Personal Inclination*) before attending to their perceived social norms or behavioral control (for research about such considerations, please see Vogel and Wänke 2016). In addition, further research into possible determinants of children's epistemic curiosity might also include children's beliefs about the malleability of their own learning abilities through active learning engagement (e.g., Blackwell et al. 2007).

Importantly, future research should investigate the ability of the CIAC to predict children's epistemic behavior in school. Although we provided evidence for the predictive validity of the CIAC on the basis of motivational measures (in accordance with the TPB), the absence of behavioral data in the present study can be seen as a limitation. However, as others have argued as well (e.g., Engel 2006), children may not exclusively 'express' their epistemic curiosity through *overt* behavior that is observable in the classroom, such as by verbalized question-asking or explanation-seeking behavior, but as much so through *covert* behavior, such as by engaging in curiosity-driven thought while working on study assignments. Therefore, a multi-method approach to assessing children's (developing) epistemic curiosity behavior is needed that includes specialized classroom observation and in-depth interviews with children and teachers to measure the range of possible behaviors and thought-processes associated with children's epistemic curiosity in formal education settings. To our knowledge, however, no such validated measurement instrument yet exist. For this reason, we were unable to reliably assess children's epistemic curiosity behavior as part of the study. Future directions of curiosity research should thus include the development of such behavioral measures.

Conclusion

With this study, we hope to contribute to research on children's epistemic curiosity in formal education settings by broadening the scope of research beyond children's curiosity *behavior* to children's *images* of and *attitudes* towards being curious learners. In our view, research on children's attitudes towards curiosity will lead us to a more complete understanding of why children behave at school the way they do. Our findings provide empirical evidence that supports propositions in previous research (e.g., Claxton and Carr 2004; Engel 2015; Engel and Randall 2009) that children's epistemic curiosity in school may only be moderate at best and seems to decline throughout primary school.

In line with this research, we believe that teachers should explicitly cultivate a positive classroom climate in which children are inspired to adopt epistemic images of curiosity and are made aware of the value of asking curious questions for their own learning and for knowledge acquisition or innovation in general. For example, teachers may lead group discussions among children to expose their pre-existing narrow, naïve and negative perceptions about the epistemic value of curiosity (e.g., Abd-El-Khalick 2012; Deng et al. 2011). In addition, teachers should pay attention to children's potential negative opinions or their fears of their classmates' judgments and strive to cultivate a sense of pleasure and pride in asking epistemic questions or posing new ideas (Post and Walma van der Molen 2018). Simple reward systems could

further convey to children that their epistemic curiosity is part of the assessment of their learning in school. The CIAC questionnaire may provide researchers and educators with a useful measurement tool to evaluate the effectiveness of such pedagogical interventions.

Funding TechYourFuture, center of expertise in technology education, funded the research reported in the manuscript. TechYourFuture was not involved with the study design; with the collection, analysis and interpretation of the data; with the writing of the manuscript; nor with the decision to submit the article for publication.

Compliance with ethical standards

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed passive consent was obtained from the parents of all the children who participated in the study.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Abd-El-Khalick, F. (2012). Examining the sources for our understandings about science: Enduring connotations and critical issues in research on nature of science in science education. *International Journal of Science Education, 34*, 353–374.
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*, 179–211.
- Ajzen, I. (2001). Nature and operation of attitudes. *Annual Review of Psychology, 52*(1), 27–58.
- Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood-Cliffs: Prentice-Hall.
- Ajzen, I., & Fishbein, M. (2005). The influence of attitudes on behavior. In D. Albarracín, B. T. Johnson & M. P. Zanna (Eds.), *The handbook of attitudes* (pp. 173–221). Mahwah: Erlbaum.
- Arango-Muñoz, S. (2014). The nature of epistemic feelings. *Philosophical Psychology, 27*, 193–211.
- Armitage, C. J., & Conner, M. (2001). Efficacy of the theory of planned behaviour: A meta-analytic review. *British Journal of Social Psychology, 40*, 471–499.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Berlyne, D. E. (1954). A theory of human curiosity. *British Journal of Psychology, 45*, 180–191.
- Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development, 78*, 246–263.
- Blalock, C. L., Lichtenstein, M. J., Owen, S., Pruski, L., Marshall, C., & Toepperwein, M. (2008). In pursuit of validity: A comprehensive

- review of science attitude instruments 1935–2005. *International Journal of Science Education*, 30, 961–977.
- Block, M. E. (1995). Development and validation of the children's attitudes towards integrated physical education-revised (CAIPE-R) inventory. *Adapted Physical Activity Quarterly*, 12, 60–77.
- Boone, H. N. Jr., & Boone, D. A. (2012). Analyzing Likert data. *Journal of Extension*, 50(2), 1–5.
- Carter, S. R. (2016). Using confirmatory factor analysis to manage discriminant validity issues in social pharmacy research. *International Journal of Clinical Pharmacy*, 38, 731–737.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9, 233–255.
- Claxton, G. (2007). Expanding young people's capacity to learn. *British Journal of Educational Studies*, 55, 115–134.
- Claxton, G., & Carr, M. (2004). A framework for teaching learning: The dynamics of disposition. *Early Years*, 24(1), 87–97.
- Conner, M., & Armitage, C. J. (1998). Extending the theory of planned behavior: A review and avenues for further research. *Journal of Applied Social Psychology*, 28, 1429–1464.
- Coulson, R. (1992). Development of an instrument for measuring attitudes of early childhood educators towards science. *Research in Science Education*, 22, 101–105.
- Darnon, C., Harackiewicz, J. M., Butera, F., Mugny, G., & Quiazade, A. (2007). Performance-approach and performance-avoidance goals: When uncertainty makes a difference. *Personality and Social Psychology Bulletin*, 33, 813–827.
- Deng, F., Chen, D. T., Tsai, C. C., & Chai, C. S. (2011). Students' views of the nature of science: A critical review of research. *Science Education*, 95, 961–999.
- Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95, 256–273.
- Eagly, A., & Chaiken, S. (1993). *The psychology of attitudes*. Belmont: Wadsworth Group/Thomson Learning.
- Eisinga, R., Te Grotenhuis, M., & Pelzer, B. (2013). The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? *International Journal of Public Health*, 58, 637–642.
- Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist*, 34, 169–189.
- Elliot, A. J., & McGregor, H. A. (2001). A 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology*, 80, 501–519.
- Engel, S. (2006). *Open Pandora's Box: Curiosity and imagination in the classroom. Sarah Lawrence occasional paper series* (pp. 1–10). Bronxville: Child Development Institute.
- Engel, S. (2011). Children's need to know: Curiosity in schools. *Harvard Educational Review*, 81, 625–645.
- Engel, S. (2015). *The hungry mind: The origins of curiosity in childhood*. Cambridge: Harvard University Press.
- Engel, S., & Randall, K. (2009). How teachers respond to children's inquiry. *American Educational Research Journal*, 46, 183–202.
- Fishbein, M., & Ajzen, I. (1974). Attitudes towards objects as predictors of single and multiple behavioral criteria. *Psychological Review*, 81, 59–74.
- Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, 7, 286–299.
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18, 39–50.
- Fortus, D. (2014). Attending to affect. *Journal of Research in Science Teaching*, 51, 821–835.
- Fouad, K. E., Masters, H., & Akerson, V. L. (2015). Using history of science to teach nature of science to elementary students. *Science & Education*, 24, 1103–1140.
- Frantom, C. G., Green, K. E., & Hoffman, E. R. (2002). Measure development: The children's attitudes towards technology scale (CATS). *Journal of Educational Computing Research*, 26, 249–263.
- Galen, B. R., & Underwood, M. K. (1997). A developmental investigation of social aggression among children. *Developmental Psychology*, 33, 589–600.
- Godin, G., & Kok, G. (1996). The theory of planned behavior: A review of its applications to health-related behaviors. *American Journal of Health Promotion*, 11, 87–98.
- Grossnickle, E. M. (2016). Disentangling curiosity: Dimensionality, definitions, and distinctions from interest in educational contexts. *Educational Psychology Review*, 28(1), 23–60.
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis* (6th ed.). Upper Saddle River: Prentice Hall.
- Hausenblas, H. A., Carron, A. V., & Mack, D. E. (1997). Application of the theories of reasoned action and planned behavior to exercise behavior: A meta-analysis. *Journal of Sport and Exercise Psychology*, 19(1), 36–51.
- Hirschfeld, G., & von Brachel, R. (2014). Multiple-group confirmatory factor analysis in R—A tutorial in measurement invariance with continuous and ordinal indicators. *Practical Assessment, Research & Evaluation*, 19(7), 1–12.
- Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55.
- Jepma, M., Verdonshot, R. G., Van Steenbergen, H., Rombouts, S. A., & Nieuwenhuis, S. (2012). Neural mechanisms underlying the induction and relief of perceptual curiosity. *Memory and Motivational/Emotional Processes*, 6, 5.
- Jirout, J., & Klahr, D. (2012). Children's scientific curiosity: In search of an operational definition of an elusive concept. *Developmental Review*, 32, 125–160.
- Jirout, J. J. (2011). *Curiosity and the development of question generation skills*. Paper presented at the 2011 AAAI 2011 Symposium. Retrieved March 23, 2017, from <https://www.aaai.org/ocs/index.php/FSS/FSS11/paper/viewFile/4194/4490>.
- Jones, M. G., Jones, B. D., & Hargrove, T. (2003). *The unintended consequences of high-stakes testing*. Lanham: Rowman & Littlefield.
- Kashdan, T. B. (2004). Curiosity. In C. Peterson & M. E. P. Seligman (Eds.), *Character strengths and virtues: A handbook and classification* (pp. 125–141). New York: Oxford University Press.
- Kashdan, T. B., & Steger, M. F. (2007). Curiosity and pathways to well-being and meaning in life: Traits, states, and everyday behaviors. *Motivation and Emotion*, 31, 159–173.
- Krosnick, J. A., & Fabrigar, L. R. (1997). Designing rating scales for effective measurement in surveys. In L. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz & D. Trewin (Eds.), *Survey measurement and process quality* (pp. 141–164). New York: Wiley.
- Kulas, J. T., & Stachowski, A. A. (2009). Middle category endorsement in odd-numbered Likert response scales: Associated item characteristics, cognitive demands, and preferred meanings. *Journal of Research in Personality*, 43, 489–493.
- Litman, J., Hutchins, T., & Russon, R. (2005). Epistemic curiosity, feeling-of-knowing, and exploratory behaviour. *Cognition & Emotion*, 19, 559–582.
- Litman, J. A. (2008). Interest and deprivation factors of epistemic curiosity. *Personality and Individual Differences*, 44, 1585–1595.
- Litman, J. A., Crowson, H. M., & Kolinski, K. (2010). Validity of the interest-and deprivation-type epistemic curiosity distinction in non-students. *Personality and Individual Differences*, 49, 531–536.

- Litman, J. A., & Pezzo, M. V. (2007). Dimensionality of interpersonal curiosity. *Personality and Individual Differences*, *43*, 1448–1459.
- Litman, J. A., Robinson, O. C., & Demetre, J. D. (2016). Intrapersonal curiosity: Inquisitiveness about the inner self. *Self and Identity*, *16*, 231–250.
- Litman, J. A., & Spielberger, C. D. (2003). Measuring epistemic curiosity and its diverse and specific components. *Journal of Personality Assessment*, *80*, 75–86.
- Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*, 1198–1202.
- Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, *116*(1), 75–98.
- Lucas, B., Claxton, G., & Spencer, E. (2013). *Progression in student creativity in school: First steps towards new forms of formative assessments* (OECD Education Working Paper No. 86). Paris: OECD.
- Marx, R. W., & Harris, C. J. (2006). No Child Left Behind and science education: Opportunities, challenges, and risks. *The Elementary School Journal*, *106*, 467–478.
- McCombs, B. L., Daniels, D. H., & Perry, K. E. (2008). Children's and teachers' perceptions of learner-centered practices, and student motivation: Implications for early schooling. *The Elementary School Journal*, *109*(1), 16–35.
- Mellor, D., & Moore, K. A. (2014). The use of Likert scales with children. *Journal of Pediatric Psychology*, *39*, 369–379.
- Metz, K. E. (2008). Narrowing the gulf between the practices of science and the elementary school science classroom. *The Elementary School Journal*, *109*, 138–161.
- Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research*, *3*, 111–130.
- Mueller, R. O., & Hancock, G. R. (2008). Best practices in structural equation modeling. In J. W. Osborne (Ed.), *Best practices in quantitative methods* (pp. 488–508). Thousand Oaks: Sage.
- Muis, K. R., Pekrun, R., Sinatra, G. M., Azevedo, R., Trevors, G., Meier, E., & Heddy, B. C. (2015a). The curious case of climate change: Testing a theoretical model of epistemic beliefs, epistemic emotions, and complex learning. *Learning and Instruction*, *39*, 168–183.
- Muis, K. R., Psaradellis, C., Lajoie, S. P., Di Leo, I., & Chevrier, M. (2015b). The role of epistemic emotions in mathematics problem solving. *Contemporary Educational Psychology*, *42*, 172–185.
- Muthén, L. K., & Muthén, B. O. (1998–2015). *Mplus User's Guide* (7th ed.). Los Angeles: Muthén & Muthén.
- National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas. Committee on a conceptual framework for new K-12 science education standards. Board on science education division of behavioral and social sciences and education*. Washington, DC: The National Academies Press.
- Olson, J. M., & Zanna, M. P. (1993). Attitudes and attitude change. *Annual Review of Psychology*, *44*, 117–154.
- Opdal, P. M. (2001). Curiosity, wonder and education seen as perspective development. *Studies in Philosophy and Education*, *20*, 331–344.
- Organization For Economic Co-Operation And Development [OECD]. (2015). *OECD skills outlook 2015: Youth, skills and employability*. Paris: OECD.
- Osborne, J., & Dillon, J. (2008). *Science education in Europe: Critical reflections* (Vol. 13). London: The Nuffield Foundation.
- Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education*, *25*, 1049–1079.
- Palmer, D. H. (2006). Sources of self-efficacy in a science methods course for primary teacher education students. *Research in Science Education*, *36*, 337–353.
- Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, *18*, 315–341.
- Pekrun, R., Goetz, T., Frenzel, A., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology*, *36*, 36–48.
- Pellegrino, J., & Hilton, M. (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. Washington, DC: National Research Academy.
- Piaget, J. (1952). *The origins of intelligence in children*. New York: International Universities Press.
- Piotrowski, J. T., Litman, J. A., & Valkenburg, P. (2014). Measuring epistemic curiosity in young children. *Infant and Child Development*, *23*, 542–553.
- Post, T., & Walma van der Molen, J. H. (2014). Effects of company visits on Dutch primary school children's attitudes towards technical professions. *International Journal of Technology and Design Education*, *24*, 349–373.
- Post, T., & Walma van der Molen, J. H. (2018). Do children experience curiosity at school? Exploring children's own conceptions of curiosity inside and outside the school context. *Learning, Culture and Social Interaction*, *18*, 60–71.
- Prudon, P. (2015). Confirmatory factor analysis as a tool in research using questionnaires: A critique. *Comprehensive Psychology*, *4*, 10.
- Ramey-Gassert, L., Shroyer, M. G., & Staver, J. R. (1996). A qualitative study of factors influencing science teaching self-efficacy of elementary level teachers. *Science Education*, *80*, 283–315.
- Raykov, T. (1997). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement*, *21*, 173–184.
- Reeve, J. (1989). The interest-enjoyment distinction in intrinsic motivation. *Motivation and Emotion*, *13*, 83–103.
- Reeve, J. (2015). *Understanding motivation and emotion* (6th ed.). Hoboken: Wiley.
- Reid, N. (2006). Thoughts on attitude measurement. *Research in Science & Technological Education*, *24*(1), 3–27.
- Reise, S. P., Waller, N. G., & Comrey, A. L. (2000). Factor analysis and scale revision. *Psychological Assessment*, *12*, 287–297.
- Rosnow, R. L. (2001). Rumor and gossip in interpersonal interaction and beyond: A social exchange perspective. In R. M. Kowalski (Ed.), *Behaving badly: Aversive behaviors in interpersonal relationships*. Washington, DC: American Psychological Association.
- Schmitt, T. A. (2011). Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment*, *29*, 304–321.
- Schoon, K. J., & Boone, W. J. (1998). Self-efficacy and alternative conceptions of science of preservice elementary teachers. *Science Education*, *82*, 553–568.
- Schwarz, N. (2008). Attitude measurement. In W. Crano & R. Prislin (Eds.), *Attitudes and attitude change* (pp. 41–60). Philadelphia: Psychology Press.
- Silvia, P. J., & Sanders, C. E. (2010). Why are smart people curious? Fluid intelligence, openness to experience, and interest. *Learning and Individual Differences*, *20*, 242–245.
- Trevors, G. J., Muis, K. R., Pekrun, R., Sinatra, G. M., & Muijselaar, M. M. (2017). Exploring the relations between epistemic beliefs, emotions, and learning from texts. *Contemporary Educational Psychology*, *48*, 116–132.
- Trochim, W. M., & Donnelly, J. P. (2006). *The research methods knowledge base* (3rd ed.). Cincinnati: Atomic Dog.

- van Aalderen-Smeets, S. I., & Walma van der Molen, J. H. (2015). Improving primary teachers' attitudes towards science by attitude-focused professional development. *Journal of Research in Science Teaching*, *52*, 710–734.
- van Booven, C. D. (2015). Revisiting the authoritative–dialogic tension in inquiry-based elementary science teacher questioning. *International Journal of Science Education*, *37*, 1182–1201.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*(1), 4–70.
- Vogel, T., & Wänke, M. (2016). *Attitudes and attitude change* (2nd ed.). London: Routledge.
- von Stumm, S., Hell, B., & Chamorro-Premuzic, T. (2011). The hungry mind: Intellectual curiosity is the third pillar of academic performance. *Perspectives on Psychological Science*, *6*, 574–588.
- Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist*, *34*, 806–838.
- Yu, C.-Y. (2002). *Evaluation of model fit indices for latent variable models with categorical and continuous outcomes* (Doctoral dissertation, University of California). Retrieved May 5, 2017, from Mplus website <http://www.statmodel.com/download/Yudisertation.pdf>.