
Silvia Diazgranados, Robert L. Selman and Michelle Dionne, Harvard Graduate School of Education

Abstract

To understand and assess how early adolescents use their social perspective taking (SPT) skills in their consideration of social problems, we conducted two studies. In study 1, we administered a hypothetical SPT scenario to 359 fourth to eighth graders. Modeled on the linguistic pragmatics of speech acts, we used grounded theory to develop a functional approach that identified three types of SPT acts: (1) the acknowledgment of different actors, (2) the articulation of their thoughts and feelings, and (3) the positioning of the roles, experiences, or circumstances that influence how they resolve problems. Study 2 tested the validity of an expanded instrument, the Social Perspective Taking Acts Measure, with 459 fourth to eighth graders. We confirmed the structure of the construct with a fully saturated confirmatory factor analysis, with factor loadings in the range of .62 and .71, and a factor determinacy of .90. We obtained evidence of criterion-related validity by successfully predicting that girls and older participants would exhibit better performance than boys and younger students, and that SPT would exhibit a negative association with aggressive interpersonal strategies, a positive but moderate association with writing, and non-significant associations with academic language, complex reasoning, and reading skills.

Keywords: social perspective taking; assessment; performance measure; adolescence

Introduction

Often defined as a human developmental ability to ‘put one self in the place of another person and to make inferences concerning the other’s capabilities, attributes, expectations, feelings, and potential reactions’ (Light, 1979), social perspective taking (SPT), nevertheless, has been conceptualized and operationalized in a range of different ways. Although cognitive-representational models based on theories of simulation (Harris, 1994), theories of mind (Gopnik & Wellman, 1994), and executive function (Carlson & Moses, 2001) use mentalistic and individualistic terms that
assume SPT occurs within the minds of individuals, social-relational models (Mead, 1938; Selman, 2003) work under the premise that SPT occurs in the relational space of coordinated human social interactivity, where increasingly complex, differentiated forms of social understanding emerge (Martin, Sokol, & Elfers, 2008). Although cognitive-representational models have dominated the field, we distinguish our work from the mentalistic and individualist assumptions of those models and align ourselves with social-relational approaches that underscore the ways in which humans use their SPT skills to interact with others and coordinate actions in the social world.

The Theoretical and Empirical Context of this Research

As children grow and have more opportunities to interact with others across broader social contexts and experiences, their SPT skills develop in ways that may move their understanding of the social world from undifferentiated and egocentric, to one that is increasingly more complexly differentiated and aware of the existence and variation of multiple points of view (Selman, 2003; Werner, 1957). Both developmental and cultural socialization processes affect the development of SPT, affording different opportunities for young people to consider different perspectives (Eisenberg, Fabes, & Spinrad, 2006). As with many social skills, across the wide swath of childhood and adolescence, females consistently tend to exhibit better SPT skills than males (Maccoby & Jacklin, 1974; Ruble, Martin & Berenbaum, 2006), perhaps because they are often socialized to care about the feelings of others and rely more heavily on relational considerations to make social decisions (Gilligan, 1982).

Research has shown that competent social perspective takers are more able to develop positive, successful relationships with others (Verhofstadt, Buyssee, Ickes, Davis, & Devoldre, 2008), and more likely to exhibit altruistic behaviors (Batson, Early, & Salvarani, 1997; Batson et al., 1997), increased social awareness (Selman, 2003), effective negotiation capabilities (Galinsky, Maddux, Gilin, & White, 2008), and communication skills (Nickerson, 1999) than those who are less skilled. Skillful perspective takers are less likely to stereotype others (Galinsky & Moskowitz, 2000), engage in aggressive behaviors (Richardson, Green, & Lago, 1998) and experience adjustment problems (Gleason, Jensen-Campbell, & Ickes, 2009) than are their counterparts. Competent perspective takers are able to learn more (Bernieri, 1991) and got better grades in school (Halberstadt & Hall, 1980) than less competent perspective takers. However, in the majority of practical, correlational and theoretically framed studies, most measures of SPT have focused on the assessment of this construct within a cognitive-representational framework. Missing in the field, as the following review will demonstrate, is the availability of a measurement approach focused on the functional development of this social skill, in particular as performed in the context of social interactions. The research reported here has two aims: (1) the operationalization of a social-relational definition of SPT performance focused on the production of SPT acts and (2) the creation and validation of a psychometrically robust measure of SPT performance that can be used for the purposes of psychosocial and developmental research.

A Comparative Review of Other Measures

Perspective Taking Self-Efficacy vs. Perspective Taking Performance. Most available validated instruments of SPT are self-report scales in which participants
evaluate their own ability and motivation to take the points of view of others. For example, the perspective-taking scale of the interpersonal reactivity index (Davis, 1983; Davis & Franzoi, 1991) contains seven items that use a five-point Likert scale (Does not describe me well = 0, Describes me very well = 4) to capture participants’ reported tendency to spontaneously adopt the psychological point of view of others in their everyday lives. For example: ‘I sometimes try to understand my friends better by imagining how things look from their perspective’. Although self-report measures are helpful for research focused on motivation and self-efficacy, they are not as useful for research that aims to capture individuals’ actual SPT performance. In fact, people’s ability to assess their own SPT competences is often inaccurate, and reports of SPT self-efficacy have been shown to have minimal relationship with actual SPT ability (Ames & Kammrath, 2004; Hall, Andrzejewski, & Yopchick, 2009; Realo et al., 2003).

Other instruments in which participants rate the abilities of others to take on different points of view may also suffer from ‘accuracy bias’. In the dyadic perspective-taking scale (Long, 1990), participants respond to 33 items in which they first rate their own abilities to take on the perspective of their partners in intimate relationships, and then provide their perceptions about their partners’ ability to do the same. For example: ‘When my partner is upset with me, he/she tries to put him/herself in my shoes for a while’. Similar to self-efficacy questions, these items only capture beliefs about the motives and skills relevant to someone else’s performance, which also suffer from bias. We concluded that these instruments do not provide an adequate measure of SPT performance and thus decided to focus on assessments of how participants respond to SPT challenges rather than on self-reports.

**Perspective Taking Accuracy vs. Perspective Taking Acts.** In contrast to self-assessments, instruments that introduce SPT challenges aim to measure SPT performance. However, most available measures of SPT performance use a cognitive-representational approach that emphasizes SPT accuracy. For example, Gehlbach (2004) developed a video task, in which participants who observed a dyadic conversation were asked to report which of a list of thoughts and feelings they believed each individual had experienced during their interaction. Consensual validation scores of SPT accuracy were calculated by matching participants’ answers against the self-reports of individuals from the videotaped conversations.

By presupposing that the best answer to a SPT challenge is the one that matches the mental reality of the target, Gehlbach’s test captures SPT performance, with a focus on accuracy. In this cognitive-representational framework, competent perspective takers are thought to be those who can accurately describe the mental content of others. However, assigning values based on content accuracy can be problematic because people can infer the feelings and thoughts of actors in ways that do not correspond to what an actor feels or thinks, but which are nevertheless plausible. Furthermore, someone with highly developed SPT skills may have a more comprehensive understanding of the point of view of the interlocutor than the interlocutor her/himself. For example, a mother may recognize that a child’s tantrum about not getting a toy derives primarily from fatigue or hunger.

For these reasons, we decided to move beyond measures based on cognitive-representational frameworks that focus on SPT accuracy, toward measures using a social-relational approach that could help us capture what participants actually do...
with their SPT skills when responding to social situations that involve the consideration of the perspectives of multiple actors.

**Perspective Taking Comprehension vs. Perspective Taking Production.** One available measure of SPT performance based on a social-relational framework is the RelQ test (Schultz, Selman & LaRusso, 2003). This measure, validated for ages 5–16, uses dilemmas and developmentally informed multiple-choice options to capture different levels of SPT integration that indicate whether someone shows evidence of understanding: only his own perspective (level 0), a second person’s perspective as distinct from his own (level 1), that a second person can also have a view of one’s own perspective (level 2), that a third person can have a view of our (first and second person combined) perspectives (level 3), and finally that his/her own perspective can be contextualized among many others (level 4).

However, the RelQ, as well as other tests that ask children to choose the best response from among a given set of multiple options that correspond to a continuum of increasing social awareness, may effectively capture SPT comprehension, but are not indicative of SPT production. In fact, the ability to express a differing perspective works in a similar fashion as does the relation of receptive and expressive language. Just as children’s receptive language skills are often more developed than their expressive language skills, children’s ability to recognize sophisticated SPT acts produced by someone else is more developed than their ability to produce similar SPT acts themselves.

As related to measurement, this means that children’s ability to select from among a set of multiple options the social choice that contains the most sophisticated SPT response will be better than their ability to produce an equally adequate or sophisticated response. All other factors equal, recognizing the best perspective-oriented response from a menu is almost always easier than supplying it because SPT comprehension requires less cognitive effort than SPT production. For this reason, multiple-option tests in which children do not retain autonomy on how to approach a SPT challenge capture the development of their SPT comprehension but not their SPT production, and reflect their optimal, but not necessarily their functional performance.

We concluded that for purposes of research oriented toward assessing what individuals do functionally with their SPT skills, we needed to operationalize a construct and create a measure of SPT performance that would capture the responses young people produce when responding to the demands of SPT tasks, as opposed to the responses they choose from a set of multiple options provided to them as measures of comprehension do.

**The Present Studies**

With these issues in mind, we conducted two studies. In Study 1, we created an instrument to capture participants’ SPT with a focus on (1) SPT performance rather than SPT self-efficacy, (2) SPT acts rather than SPT accuracy, and (3) SPT production rather than SPT comprehension. Our goal was to gain an understanding of what adolescents do with their SPT skills in the responses they produce to hypothetical tasks that involve the consideration of multiple perspectives. In Study 2, we built on this emergent framework, expanded the instrument constructed in Study 1 into a psychometrically robust measure—the Social Perspective Taking Acts Measure.
(SPTAM), and evaluated its validity and reliability. Our goals were to identify whether the hypothesized latent structure of the SPT construct is consistent with the observed measures of the SPTAM, and to find evidence of criterion-related validity.

**Study 1 Method**

**Participants**

The data used in Study 1 were collected in the Fall of 2011. The sample comprised 359 fourth to eighth grade students from diverse socioeconomic backgrounds in five public schools along the east coast of the USA. The sample was balanced by gender and ethnically diverse, including African-American or black, Caucasian and Latino students. The majority of students were fourth (42 percent) and sixth (51 percent) graders, and a minority seventh (4 percent) and eighth (3 percent) graders. We chose to work with this population because SPT develops rapidly during early adolescence. The emergence of formal operational thinking (Piaget, 1955) allows early adolescents to consider the thoughts and feelings of others more than ever before. Additionally, the psychological struggle they experience to form a unique identity that will gain acceptance by friends, family, and society at large (Erikson, 1968) makes early adolescents highly motivated to use their social competencies to understand the perspectives of multiple others.

**Instrument**

The questionnaire introduced children to a scenario that reflected a situation of school bullying in which a young student who is observing the harassment of a victim does not know what to do and is asking different classmates for advice. Survey respondents were prompted to think about the recommendations this observer might receive from three other actors, each portrayed in one of the following three positions: (P1) someone who is often teased for similar reasons as the victim; (P2) someone who often socializes with the teasers; and (P3) participant (student) respondents themselves. Then, questionnaire respondents were prompted to answer three questions: (Q1) what would (the prompted actor) recommend to the observer? (Q2) why would (the prompted actor) make that recommendation? and (Q3) what might go wrong with this recommendation?

**Procedure**

During the Fall 2011, a SPT challenge was administered as part of an educational intervention and, therefore, we did not obtain parental consents. Trained research assistants read the instructions to classes of students and walked the groups through the questions in the measure, providing them with 30 min to respond in writing. Students were allowed to ask clarifying questions. Students who completed earlier sections more quickly than the allotted time were allowed to proceed without waiting for later prompts to be read aloud. Students’ names were replaced by ID codes to protect confidentiality.
Study 1 Results

Analytic Strategy

We used grounded theory (Charmaz, 2006) to analyze students’ written responses to the SPT scenario, with a focus on the way responses took into account the different actors embedded in the scenario. This method facilitated the process of theory generation by allowing different categories to emerge from the data. We used a non-linear process of open, selective, and axial coding to analyze how students use their SPT skills—what they do linguistically in their written responses. To conduct a nuanced analysis, we conducted line-by-line open coding, identifying the ways in which students used their SPT skills to respond to a scenario from a given perspective (i.e., position 1, 2, or 3). In the following selective phase, we identified focused codes by looking at the most salient SPT categories that had emerged from the data. We wrote memos to explore not only the explicit, but also the implicit, unstated SPT acts that the study participants engaged in when responding to SPT challenges. Finally, we developed an axial framework to understand how early adolescents engage in SPT in response to social situations that require the consideration of multiple points of view.

As a result, we developed categories of analysis and a conceptual framework to make meaning of our data. The research team held group discussions about the emerging codes, examined typical and atypical cases, and held inter-rater reliability meetings. We agreed on a codebook that was deemed stable when we achieved an indicator of inter-rater reliability of .90, which reflects the proportion of units on which raters agreed out of the total number of units coded.

Findings

Our analysis of participants’ responses to the SPT challenges led us to reliably identify how early adolescents performed a variety of SPT acts, which vary in their (1) function and (2) level of integration. Following Austin’s (1955) linguistic work on speech acts, we defined SPT acts as the basic elements that constitute the action(s) of taking the social perspective of other(s). We postulated that:

1. The function of a SPT act refers to the social-relational effect of a SPT statement; what the statement does in the context of a social interaction as communicated through language. In our data, we observed three types of SPT acts whose functions were to: (1) acknowledge the existence of actors in a given situation; (2) articulate how actors think, feel or are inclined to behave; and (3) identify their position in the scenario, according to the roles, circumstances, and their experiences that may influence how they think and feel about an issue.

2. The level of integration of a SPT act refers to the degree to which a response acknowledges, articulates, and identifies the positions of an increasingly larger number of actors in a given situation. Responses vary in their degree of integration because for every scenario, there are multiple potential perspectives (of actors) that could be acknowledged, articulated, and positioned. Although some participants provide responses that acknowledge the presence of few actors, others provide responses that acknowledge many actors. Some responses articulate or position the perspectives of many of the actors acknowledged whereas others articulate or position the perspective of only a few of those acknowledged.
Based on these findings, we developed a codebook (Diazgranados, Selman, & Dionne, unpublished) to rate participants’ SPT performance according to both the functions and degree of integration of the SPT acts included in their responses (Table 1).

**Study 2 Method**

**Participants**

The data used in Study 2 were collected during the summer of 2012. They included responses from 459 students in grades four to eight from one urban school located in a large city along the east coast of the United States. The study formed part of an educational intervention and, therefore, we did not obtain parental consents. The sample was balanced by gender, with 80 percent of students coming from low income families and receiving free or reduced-price lunch. Participants were from diverse ethnic backgrounds, with 66 percent African-American or black, 22 percent Caucasian and 7 percent Latino. The sample was comprised of 25 percent fourth graders, 21 percent fifth graders, 18 percent eighth graders, 16 percent seventh graders, and 16 percent eighth graders.

**Procedure**

We administered the SPTAM, identified for students as ‘The Advice on Making Social Choices Measure’, in whole classrooms. Students received a gel pen for their participation. Research assistants read the instructions and walked participants through the scenarios and questions, providing them with 4 min to answer each prompt, for a total of 30 min of administration. If participants completed a section before others, they were allowed to move on at their own pace. Names were replaced by ID codes to protect confidentiality.

On the same day, participants also responded to other group-administered assessments of their performance on measures of reading, writing, academic language, and complex reasoning skills, as well as a measure of aggressive interpersonal strategies (AINS).

**Instruments**

*The Social Perspective Taking Acts Measure.* The SPTAM is an instrument that aims to capture early adolescents’ functional ability to: (1) acknowledge the existence of different actors, (2) articulate how others think and feel about a given situation, and (3) position how different actors see the situation given their roles, experiences, and circumstances in their own social world. The measure is designed to capture (1) SPT performance, as opposed to self-efficacy, (2) SPT acts, as opposed to SPT accuracy, and (3) SPT production, as opposed to SPT comprehension.

The SPTAM builds on the instrument we used in Study 1, introducing participants to two illustrated narrative scenarios that reflect social problems that often occur in schools (See Appendix). In each scenario, an actor who is observing a social problem (i.e., a witness to teasing) does not know what to do and is asking different people for advice. Survey respondents are prompted to think about the recommendations this observer might receive from the following three advisors: (1) someone who is often...
<table>
<thead>
<tr>
<th>Function</th>
<th>Definition and coding system</th>
<th>Example</th>
<th>Level of integration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acknowledgement</td>
<td>The act of identifying the various actors involved in a given social situation. It can be determined by counting, only once per unit of analysis, the names and pronouns that refer to any particular actor that is included in the unit of analysis.</td>
<td>‘Get to know Jariah’. &lt;br&gt;‘She would want to help him’. &lt;br&gt;‘Ali would tell Casey to talk to the teacher’ &lt;br&gt;‘Ali would recommend Casey to tell the teasers to stop teasing Jariah’ &lt;br&gt;‘I think that Ali would recommend to Casey to help Jariah by hanging out with her so that her classmates can see and then they would start hanging out with Jariah’</td>
<td>1 point 2 points 3 points 4 points 5 points</td>
</tr>
<tr>
<td>Articulation</td>
<td>The act of describing the thoughts, feelings, or orientations to action of distinct actors in a given social scenario. It can be determined by counting, only once per unit of analysis, the actors whose feelings, opinions, beliefs, preferences, and orientations to action are described in the scenarios.</td>
<td>‘She wouldn’t care about his problems’. &lt;br&gt;‘Lee is going to be really embarrassed but classmates will think it’s funny’. &lt;br&gt;‘Sam recommends to annoy her because the teasers think it’s fun’. &lt;br&gt;‘I think she [Ali] would make that recommendation because she knows Jariah is sad nobody likes getting teased’.</td>
<td>1 point 2 points 3 points 4 points</td>
</tr>
<tr>
<td>Positioning</td>
<td>The act of identifying the roles, circumstances, or</td>
<td>‘Ali would do this because she has been teased and</td>
<td>1 point</td>
</tr>
</tbody>
</table>
teased for similar reasons as the victim; (2) someone who often socializes with the teasers; and (3) participant respondents themselves. Then, questionnaire respondents were prompted to answer three questions: (1) what would (the prompted actor) recommend to the observer? (2) why would (the prompted actor) make that recommendation? (3) what might go wrong with this recommendation?

This structure (two scenarios*three advisors), provides participants with the opportunity to produce open-ended responses to six distinct sets of prompts. Answers to all three questions provided by each advisor constitute one unit of analysis, which receives one score for each of the three subscales: acknowledgement, articulation, and positioning. Using the coding system, we developed in Study 1, we scored each unit of analysis (scenario by advisor) according to the function and level of integration of the SPT acts present in their responses. We obtained subscale scores by calculating participants’ mean performance across the six units of analysis. A detailed guideline with examples about how to code and how to obtain subscale scores can be found in the coding manual (Diazgranados et al., unpublished).

**Aggressive Interpersonal Strategies Measure.** This measure assesses participants’ disposition to recommend the use of aggressive strategies in response to scenarios that contain socially ambiguous intentions that focus on potential rejection or provocations (Dalhberg, Toal, & Behrens, 1998). For each story, participants choose the response they would use with one option indicating an aggressive strategy (scored as 1) and three options indicating non-aggressive strategies. Responses to the

<table>
<thead>
<tr>
<th>Function</th>
<th>Definition and coding system</th>
<th>Example</th>
<th>Level of integration</th>
</tr>
</thead>
<tbody>
<tr>
<td>attributes that qualify the position distinct actors hold in a social scenario. It can be determined by counting, only once per unit of analysis, the actors whose roles, attributes, experiences, or circumstances are identified in the scenario.</td>
<td>knows how it feels’. ‘Ali would recommend this because <strong>she was teased before</strong> and it’s not fair to tease Jariah just because <strong>her culture is different</strong>’. ‘Carson was teased before so he will say for Rene to tell the teacher since <strong>a teacher is supposed to help</strong> and so is the <strong>classroom president</strong>’.</td>
<td>2 points</td>
<td>3 points</td>
</tr>
</tbody>
</table>
respective vignettes are added to create a score for tendency to recommend the use of AINS. The measure showed good internal reliability (Cronbach’s alpha = .80).

Written Language Scale of the Oral and Written Language Scale (OWLS-II). This measure evaluates written expression, for ages 5–21 (Carrow-Woolfolk, 1995). Each set of 15–18 age-graded items involves writing tasks similar to those students encounter in their classroom. Some tasks are open-ended and require students to complete a story or interpret a statement. Others are structured, requiring students to copy words or write a dictated sentence. The examiner presents oral, written, and pictorial prompts, and examinees write their answers in a Response Booklet. Descriptive analysis tables and worksheets help scorers generate a qualitative summary of student writing. The test has excellent inter-rater reliability ($\alpha = .95$) and a half-split reliability in the range of .84–.91.

GATES-MacGinitie Reading Test. This measure evaluates students’ levels of reading achievement using grade-level texts (MacGinitie & MacGinitie, 1988). The test examines language and reading abilities, including literary concepts, oral language concepts, letter recognition, and letter-sound relationships. Readers between first and 12th grade receive scores reflecting their vocabulary and comprehension. The test has excellent internal consistency ($\alpha = .96$) and test–re-test reliability in the range of .85–.90.

Core Academic Language Skills. The core academic language skill (CALS) is a group-administered assessment that captures the development of academic language skills for children in fourth through eighth grades (Uccelli et al., 2014). Academic language comprises several dimensions, including grammatical processes such as subordination and nominalization, sophisticated vocabulary, discourse structure, and differentiating formal from informal registers. Reliability for the CALS was robust, with a Cronbach’s alpha of .92 and split reliability at .82.

Lectical Assessment System. Students’ complex reasoning was assessed using the lectical assessment system (LAS), a measurement approach based on the dynamic skill theory of cognitive development (Fischer & Bidell, 2006; Dawson, 2002). This framework describes 13 levels of cognition ordered along a dimension of hierarchical complexity. Inter-rater reliability for the LAS across a number of studies consistently is in the range of .85–.95, depending on age range of the sample (Dawson & Gabriellen, 2003).

Study 2 Results

Analytic Strategies

To assess whether our hypothesized latent construct of SPT performance had a structure that was consistent with the observed measures of the SPTAM, we conducted a confirmatory factor analysis (CFA), using the statistical program MPlus version 7. CFA is a statistical method that uses analysis of covariance to confirm hypothesized relationships among a set of variables. We first established that the indicators had a normal distribution and that the subscales exhibited good reliability.
Then, we tested the hypothesis that acknowledgement, articulation, and positioning are indicators of a single-latent construct: SPT (See Figure 1).

Given that a one-factor model with three indicators is just-identified, that is, it has the same number of free parameters as observations and, therefore, a unique solution that perfectly fits the data, we did not use fit indices to evaluate goodness-of-fit. Instead, we relied on a careful examination of the parameter loadings, R-square of each item, standardized residuals, and factor determinacy of the model. Factor loadings reflect the degree to which each item is linked to the hypothesized factor. If an item is hypothesized to load on a particular factor, then its loading should be large, positive, and statistically significant, with values larger than .50 being desirable. The standardized residuals reflect the ratio of a covariance residual over its standard error, which are interpreted as a z test of whether the population covariance is zero. Desirable standardized residuals approximate zero and do not exceed .1 points. The R-square measures the proportion of variance in Y that is explained by the model and will be between 0 and 1 with larger values indicating a higher proportion of variance in the endogenous variable explained by the model. Factor determinacy is an estimation of the internal reliability of the latent variable. Values range from 0 to 1, with larger values indicating better measurement of the factor by the observed indicators (Kline, 2011).

In an effort to assess whether the SPT latent variable is a valid measure of SPT, we tried to find evidence of criterion-related validity. Criterion-related validity refers to the degree to which the score of an instrument predicts an outcome based on information from external instruments (Kline, 2011). To assess the validity of our instrument, we predicted the following relationships between the SPTM and other available instruments and information:

1. Children in higher grades will exhibit better performance in the SPTAM than children in lower grades, because the ability to take on the perspectives is a developmental skill that progresses with age (Selman, 2003).

![Figure 1. CFAs for the Internal Structure of the Perspective Taking Measure.](image-url)
Girls will exhibit better SPT performance than boys in the SPTAM, because females have been consistently found to exhibit better social skills than their male counterparts (Ruble et al., 2006) and because girls are often socialized to consider relationships in their social and moral reasoning (Gilligan, 1982).

The SPTAM will have a negative and moderate association with a measure of (AINS) because competent perspective takers are less likely to engage in aggressive behaviors than their less competent counterparts (Richardson et al., 1998), and have been documented to exhibit good social problem solving (Gehlbach, 2004), and communication skills (Nickerson, 1999).

The SPTAM will have a positive and moderate association with a measure of writing (OWLS) because our coding system uses a pragmatic approach that focuses on production of SPT acts. Although SPT performance is inherently linked to language production, it cannot be reduced to writing skills, and therefore, we should not find a high correlation between these measures.

The SPTAM will exhibit a positive, low to moderate association with measures of complex reasoning (LAS), academic language (CALS), and reading (GATES). Given that the SPTAM is read out loud and uses illustrations, we expect that the cognitive skills involved in reading will not interfere with those involved in SPT, and therefore, that our measure of SPT will not predict reading skills in a significant way.

We used structural equation modeling (SEM) to estimate the correlations between our measurement model for the SPT factor and a set of observed demographic variables (See Figure 2). We evaluated our model on the basis of their parameter estimates and a combination of goodness-of-fit indicators, which include the chi-square statistic, Comparative Fit Index (CFI), root mean square error of approximation (RMSEA), and Standardized Root Mean Square Residual (SMRS). Additionally, we used a maximum likelihood estimator to compare the fit of the models against each

\[ \text{Figure 2. SEM Model of the Relationship Between a Latent Measure of Perspective Taking and Measures of Aggressive Interpersonal Strategies, Writing, Reading, Academic Language, Complex Reasoning, Gender, and Grade.} \]
other. The chi-square statistic assesses absolute fit of the model to the data, but it is highly sensitive to sample size and assumes that the model is correct. A chi-square statistic larger than .05 is indicator of good fit. The CFI is an incremental fit index that compares the theoretical model with a null model, and uses the non-central chi-square distribution. A CFI greater than .95 indicates a good fit of the model. The RMSEA corresponds to the root mean square error of approximation; an RMSEA of .05 is considered to demonstrate optimal fit of the model. The SMRS is a measure of the mean absolute value of the correlation residuals; an SMRS smaller than .05 indicates good fit (Kline, 2011).

**Findings**

Table 2 shows the means, standard deviations, inter-scale correlations, and reliability of the scales. All the subscales of the measure exhibited positive and statistically significant correlations with each other. All scales had good levels of internal consistency as reflected by their Cronbach’s alpha.

Table 2. Means, Standard Deviations, Minimum, Maximum, Range, Inter-scale Correlations, and Cronbach’s Alpha for the Acknowledgement, Articulation, and Positioning Subscales

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
<th>Ackn</th>
<th>Artic</th>
<th>Posit</th>
<th>Cronbach’s alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Acknowledgment</td>
<td>3.33</td>
<td>.84</td>
<td>.4</td>
<td>5.2</td>
<td>1</td>
<td></td>
<td></td>
<td>.80</td>
</tr>
<tr>
<td>2. Articulation</td>
<td>2.03</td>
<td>.84</td>
<td>0</td>
<td>4.4</td>
<td>.46</td>
<td>1</td>
<td></td>
<td>.83</td>
</tr>
<tr>
<td>3. Positioning</td>
<td>1.09</td>
<td>.43</td>
<td>0</td>
<td>2.5</td>
<td>.43</td>
<td>.40</td>
<td>1</td>
<td>.70</td>
</tr>
</tbody>
</table>

**Figure 3.** Distribution of Perspectival Acknowledgement.
The distributions of SPT acknowledgement (see Figure 3), articulation (see Figure 4), and positioning (see Figure 5), subscales approximated normality. We hypothesized that taken together, these scales represented a latent construct: SPT. To test this hypothesis, we conducted a CFA, which provided support for a three-dimensional model in which SPT is a factor comprising acknowledgement, articulation, and positioning. Standardized coefficients between the observed variables and the latent factors for SPT are depicted in Table 3. All parameter estimates were positive, statistically significant, and exhibit loadings in the range of .62–.71. The R-square suggests that the model

Figure 4. Distribution of Perspectival Articulation.

Figure 5. Distribution of Perspectival Positioning.
explained 50 percent of the variation of acknowledgment, 43 percent of articulation, and 38 percent for positioning. The model exhibited a factor determinacy of .90, indicating a highly reliable measure. Based on the evidence, we concluded that SPTAM had three observable subscales that exhibited acceptable levels of reliability, as well as a latent score of SPT with an excellent level of internal consistency.

With regard to criterion-related validity, simple correlations between our latent factor of SPT and each of the other measures used in our analysis indicated that our construct exhibited moderate, positive and statistically significant associations with reading \((rs = .37; p < .001)\), writing \((rs = .42, p < .001)\), academic language \((rs = .18, p < .001)\), and complex reasoning \((rs = .03, p < .001)\). As predicted, SPT exhibited a negative and statistically significant association with a measure of AINS \((rs = .05, p < .001)\). As expected, the parameter estimates showed that the SPTAM captured significant differences in the SPT performance of students, with scores favoring females \((p < .001)\), as well as grade-related differences favoring students in higher grades \((p < .01)\) (See Table 4).

To understand the unique contribution of each measure used in the SEM (see Figure 2), we explored the relationship of our latent SPT construct with each of these different variables, after controlling for the rest (e.g., relationship of SPT with reading, after controlling for writing, academic language, complex reasoning, age, and gender). Consistent with our predictions, we found that that after using controls, only the relationships that SPT had with writing, gender, grade, and AINS remained

| Table 3. Correlations Between Indicators, Unstandardized Parameter Estimates, Standard Errors, p Values, Standardized Parameter Estimates, and R-Squares for a One Factor CFA with Three Indicators of Perspective Taking |
|----------------------------------|----------------|----------------|----------------|--------------------|
| Parameter                       | Unstandardized | S.E.           | p value        | Standardized       | R-square          |
|----------------------------------|----------------|----------------|----------------|--------------------|
| Acknowledgment                   | 2.21           | .25            | .000           | .62                | .50              |
| Articulation                     | 2.03           | .23            | .000           | .71                | .43              |
| Positioning                      | 1              | 0              | .000           | .65                | .38              |

<p>| Table 4. Unstandardized Parameter Estimates, Standard Errors, Standardized Parameter Estimates, and p Values for Correlations Between our Measurement Model of PT and Nine Measures of Criterion Validity |
|----------------------------------|----------------|----------------|----------------|--------------------|</p>
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Unstandardized</th>
<th>SE</th>
<th>p value</th>
<th>Standardized</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggressive int. strategies</td>
<td>-.33</td>
<td>.08</td>
<td>&lt;.001</td>
<td>-.23</td>
</tr>
<tr>
<td>OWLS (reading)</td>
<td>11.54</td>
<td>1.53</td>
<td>&lt;.001</td>
<td>.51</td>
</tr>
<tr>
<td>GATES (writing)</td>
<td>14.19</td>
<td>2.53</td>
<td>&lt;.001</td>
<td>.36</td>
</tr>
<tr>
<td>Academic language</td>
<td>1.75</td>
<td>.37</td>
<td>&lt;.001</td>
<td>.30</td>
</tr>
<tr>
<td>Complex reasoning</td>
<td>.37</td>
<td>.15</td>
<td>&lt;.010</td>
<td>.16</td>
</tr>
<tr>
<td>Female</td>
<td>1.37</td>
<td>.011</td>
<td>&lt;.001</td>
<td>.43</td>
</tr>
<tr>
<td>Grade</td>
<td>.10</td>
<td>.313</td>
<td>&lt;.010</td>
<td>.12</td>
</tr>
</tbody>
</table>

*Note: Chi-square (42.19, df = 14) = .0001, RMSEA = .06 (.04, .09), CFI = .96, SRMR = .03.*
and that the relationships of SPT with academic language, complex reasoning, and reading became weak and non-significant. The model suggested that on average in the population, for every additional point in the OWLS writing test, students scored 11.54 more points on the SPTAM ($p < .001$). Similarly, on average in the population, for every additional unit in the AINS measure, students obtained .33 fewer points on the SPTAM ($p < .10$). Girls scored 1.37 points higher on the SPTAM than boys. Finally, for every additional grade level, students scored .37 points more on the SPTAM ($p < .001$) (See Table 5).

General Discussion

We conducted two studies to understand and assess the ways in which early adolescents use their SPT skills to respond to social challenges that involve multiple actors. We first used qualitative methods to gain an understanding of functional SPT, and then developed and tested the validity and reliability of a performance measure of SPT.

An Emerging Framework: The Functional Dimensions of SPT

In Study 1, we used grounded theory to analyze responses from 359 participants in grades four to eight who responded to a SPT challenge that involved the consideration of multiple perspectives. Our aim was to operationalize SPT in a way that would capture SPT performance as opposed to SPT self-efficacy, SPT acts as opposed to SPT accuracy, and SPT production as opposed to SPT comprehension.

Based on our analysis, we posit a social-relational approach to SPT based in the linguistic pragmatics of speech acts (Austin, 1955), which identifies SPT as composed of acts that serve different functions. We found that when challenged to resolve social situations, participants produced responses that (1) acknowledged the existence of different actors, (2) articulated the thoughts, feelings and orientations to action of those actors, and (3) positioned these actors according to their characteristics, roles, or circumstances in the scenario. Responses varied in their levels of integration, as participants demonstrated different abilities to acknowledge, articulate and position the perspectives of multiple actors in the scenario.

### Table 5. Unstandardized Parameter Estimates, Standard Errors, Standardized Parameter Estimates, and $p$ Values for a Path Analysis of a Measurement Model of PT and Nine Measures of Criterion Validity

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Unstandardized</th>
<th>SE</th>
<th>$p$ value</th>
<th>Standardized</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggressive int. strategies</td>
<td>-.111</td>
<td>.061</td>
<td>.07</td>
<td>-.140</td>
</tr>
<tr>
<td>OWLS (writing)</td>
<td>.015</td>
<td>.005</td>
<td>.002</td>
<td>.298</td>
</tr>
<tr>
<td>GATES (reading)</td>
<td>.002</td>
<td>.003</td>
<td>.46</td>
<td>.030</td>
</tr>
<tr>
<td>Academic language</td>
<td>.013</td>
<td>.019</td>
<td>.50</td>
<td>.092</td>
</tr>
<tr>
<td>Complex reasoning</td>
<td>.007</td>
<td>.029</td>
<td>.80</td>
<td>.024</td>
</tr>
<tr>
<td>Female</td>
<td>.226</td>
<td>.045</td>
<td>&lt;.001</td>
<td>.378</td>
</tr>
<tr>
<td>Grade</td>
<td>.067</td>
<td>.020</td>
<td>&lt;.001</td>
<td>.244</td>
</tr>
</tbody>
</table>

Note: Chi-square (18.08, df = 14) = .20, RMSEA = .03 (.00, .08), CFI = .98, SRMR = .02.
Our operationalization of SPT moves the field forward using a pragmatic-based approach that begins to elucidate the functional dimensions of SPT. Our emergent construct aligns with social-relational frameworks (Martin et al., 2008; Mead, 1934; Selman, 2003) that differ from the cognitive-representational approaches that have dominated the field for years. Whereas the latter focus on the mental processes that enable people to infer accurately the thoughts and feelings contained within the minds of others, our work identifies the functions and levels of integration of the SPT acts individuals perform in their attempt to consider resolutions to social problems that involve multiple actors. From a functional view, competent perspective takers are not defined as those who accurately capture the thoughts and feelings of others, but those who effectively, and habitually, acknowledge, articulate, and position multiple perspectives when considering social problems.

Future research can explore the functional uses of other SPT acts in the contexts of developmental and cultural studies. In this regard, we do not claim that acknowledgement, articulation, and positioning are the only SPT acts that people can produce, but only the SPT acts we observed using the instrument we developed with early adolescents in the northeast region of the United States. Particularly longitudinal, but also cross-sectional studies across a wider chronological age span would be essential to gaining deeper insights of this emerging construct. It is possible that by altering our instrument or by expanding our population, chronologically, culturally, and contextually, we would find SPT acts that serve other functions. For example, older or more socially sophisticated individuals will likely produce acts whose function are to interpret how the thoughts, feelings, and orientations to action of actors are shaped by the positions they hold in a given scenario, and to critique the different points of view that have been articulated by identifying the positions that had not been considered by a given actor. From a cultural perspective, acts of acknowledgement, articulation, and positioning may be produced under different circumstances. Some types of actors may be more likely to be acknowledged in some cultures than in others, and some perspectives and positions may be more readily articulated and positioned in some contexts than in others.

A Novel Instrument to Measure SPT Performance

During Study 2, we built on the operational definition of SPT acts and created the SPTAM, a robust instrument that we examined in terms of its psychometric properties. In fact, despite the existence of a range of different measures that capture SPT (Davis, 1983; Davis & Franzoi, 1991; Gehlbach, 2004; Schultz et al., 2003), none assessed young people’s functional ability to produce responses that integrate the perspectives of multiple actors in a given scenario.

After administering the SPTAM to 459 fourth to eighth grade students, we conducted correlations among subscales, analysis of internal consistency and a CFA, and concluded that our hypothesized construct of SPT adequately fits the observed data. Specifically, we observed that perspective acknowledgement, articulation, and positioning are reliable scales that can be seen as related but distinct constructs, which come together to serve as indicators of a latent variable: SPT. The three subscales exhibited positive, moderate, and statistically significant correlations with each other. The CFA showed high loadings of the parameter estimates of these variables, and a good fit of the model according to several indicators, including the
Chi-square, SRMR, RMSEA, and CFI. A high factor determinacy confirmed that our latent construct is a reliable performance measure of SPT.

We also conducted SEM analyses to find evidence of criterion-related validity. Consistent with the literature, we confirmed that girls are more competent perspective takers than boys (Ruble et al., 2006), that older students are more competent perspective takers than younger students (Selman, 2003), and that competent perspective takers are significantly less likely to recommend the use of aggression as a way to solve conflicts (Gehlbach, 2004; Nickerson, 1999; Richardson et al., 1998). We also confirmed the existence of a positive and moderate association between the SPTAM and participants’ writing skills, which we had predicted based on our emphasis on SPT production. In this regard, we see the use of expressive language skills as inherently related to, but different from, SPT performance. Although becoming a good or practiced writer might improve one’s SPT skills, just as it may improve one’s thinking and language abilities, from a theoretical and logical point of view, SPT is prior to writing. Writing presupposes SPT even as it may subsequently strengthen SPT functionality, and in this sense, the variations found within and across SPT production skills are expected to be related to writing skills, but cannot be fully accounted for by or reduced to them. Future research needs to explore the relationship between SPT performance production using data from students’ oral responses to the SPTAM.

Finally, we also confirmed that SPT is a construct that overlaps, but is distinct from skills assessed by the LAS, CALS, and OWLS. In fact, analyses that looked at the unique contribution each measure makes to explaining SPT show that the associations between SPT and complex reasoning, academic language, and reading comprehension skills that were initially positive and statistically significant, became weaker and non-significant when controls were included, and that the expected associations of SPT with grade, gender, AINS, and writing skills are the associations that remained. We concluded that the SPTAM demonstrates the properties of a valid measure of early adolescents’ developing functional ability to actively: acknowledge the existence of different stakeholders in scenarios of social conflict, to articulate their feelings, thoughts, and dispositions to action, and to identify the attributes, roles, or circumstances that quality their positions in a given situation.

Future studies need to determine other types of reliability for the SPTAM, including (1) test–re-test reliability or the degree to which the SPT scores are stable over different testing occasions; (2) measurement invariance across equivalent versions of the SPTAM that feature other scenarios with a distinct focus on social groups, social institutions, cultures, and historical figures beyond those assessed here; (3) measurement invariance across oral and written forms of administration; (4) measurement invariance across different methods with data generated through tests vs. in vivo social interactions in laboratory or naturalistic settings; (5) measurement invariance across groups from different cultures; and (6) measurement invariance among other age groups.

**Implications for Practice and Evaluation Research**

With regard to our emerging framework, we believe a functional approach to SPT can inform educators, program designers, and evaluators interested in scaffolding the development of SPT among young people. Education programs can support early adolescents’ social and academic development by providing them with
opportunities to practice different SPT acts. They can also encourage students to
develop their own points of view in ways that integrate the consideration of the
existence of multiple actors, their thoughts and feelings, and the positions that affect
them. Our framework can also be particularly relevant for educators interested in
enhancing classroom-based activities such as discussions and debates (Elizabeth,
Ross, Snow, & Selman, 2012), reading of printed texts, and writing of persuasive
essays (Dray, Selman, & Schultz, 2009). Specifically, educators can encourage stu-
dents to improve their SPT performance by helping them identify which functional
acts they are using and with which degree of integration, as well as which other
acts they could learn to use, as developmentally appropriate.

With regard to the measure, the SPTAM fills a gap by providing researchers
with a tool to assess early adolescents’ ability to produce SPT acts that serve differ-
ent social-relational functions. This innovative instrument can be particularly useful
in the context of impact evaluations of programs whose theory of change includes
SPT performance as a mechanism of change or outcome. The SPTAM can also be
used to conduct psychosocial research, such as longitudinal studies that identify
developmental trajectories in SPT performance, as well as SPT acts as related to
SPT self-efficacy, accuracy, and comprehension.

References
Austin, J. A. (1955). In J.O. Urmson & Marina Sbisà (Eds.), *How to do things with words:
The William James lectures delivered at Harvard University in 1955*, 1962. Cambridge,
MA: Harvard University Press.
feels versus imagining how you would feel. *Personality and Social Psychology Bulletin,
23*, 751–758.
empathy-induced helping due to self-other merging? *Journal of Personality and Social
Psychology, 73*, 495–509.
Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and child-
Carrow-Woolfolk, E. (1995). *Oral and written language survey (OWLS)*. Circle River, Min-
beliefs, and behaviors among youths: A compendium of assessment tools*. Atlanta, GA:
Centers for Disease Control and Prevention, National Center for Injury Prevention and
Control.
of Applied Measurement, 3*, 146–189.
Diazgranados, S., Dionne, M. O., & Selman, R. L. (2011). *The social perspective taking
acts measure (SPTAM)*. Unpublished manuscript, Graduate School of Education, Harvard
University, Cambridge, MA.

© 2015 John Wiley & Sons Ltd
Social Development 00, 00, 2015


**Acknowledgments**

The research reported here was supported by the Institute of Education Sciences, US Department of Education through Grant R305F100026, which was awarded to the Strategic Education Research Partnership (SERP) as part of the Reading for Understanding Research Initiative. The opinions expressed are those of the authors and do not represent the views of the institute or the US Department of Education. We express our gratitude to our colleague, Catherine Snow for her advice on and support of this initiative; to the members of the IES Word Generation Project at Harvard and SERP for their constant collegiality over the past five years, to the students and teachers who shared their valuable time and insights with us, and to numerous doctoral students and faculty for their helpful comments as we conducted this work. Finally, we are thankful for the valuable comments from three anonymous reviewers and the editor of this journal.
Appendix

The Social Perspective Taking Acts Measure
Diazgranados, S; Selman, R. L., & Dionne, M.
Harvard Graduate School of Education

In school, students can learn about reading, math, history, and science. They can also learn how to get along and talk with each other. This questionnaire is about situations that could happen to anyone, even you. We are interested in what you think about these situations, and what you think people your age might do under these circumstances. Many students have told us that the stories in this survey are just like things that have happened to them. We hope you will find our questions interesting.

• There are no right or wrong answers to these questions.
• Each student will have different opinions, thoughts and feelings about different issues or situations. We are interested in what YOU think.
• Please be as clear and thoughtful as possible.
• No one in your school will know how you answered.
• If you do not understand a question, please try to answer it anyway and then write us a note below the question letting us know what you didn’t understand.

Before we begin, here are some more vocabulary words that you may not know, and the definitions of these words. You can use these words to help you in this survey.

Recommend: This word means “to suggest doing something.” If you recommend asking a friend for advice, that means you think it is a good idea to talk to a friend and listen to what they have to say.

Situation: This survey is about different things that happen to students and that could happen to you or in your school. For example, you may find yourself in a situation where you are asked to give advice to a friend about how to behave or respond.
1. Jariah’s Weird Sense of Style

You have lots of friends in your class this year and you are enjoying school. In January, your teacher introduces a new student to the class, Jariah, whose family arrived from a different far away city very recently. Jariah’s hairstyle is really strange. Also, no one in the class likes the music Jariah listens to. Some students are teasing Jariah because they think Jariah is weird. Casey is a student who has many friends and gets along with most classmates. Casey has been observing the situation and does not know what to do. Casey is asking different people for advice.
Ali is a student who was recently teased by other students, but found a way to solve the problem.

a. What do you think Ali would recommend to Casey?

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

b. Why do you think Ali would make that recommendation?

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

c. What might go wrong with Ali’s recommendation?

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________
Sam is a student who often hangs out with the teasers.

a. What do you think Sam would recommend to Casey?

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

b. Why do you think Sam would make that recommendation?

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

c. What might go wrong with Sam’s recommendation?

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________
Now that you thought about what other people would recommend to Casey, think about what you would say to Casey. Imagine Casey comes to you and asks: What should I do?

a. What would **you** recommend to Casey?

b. Why would **you** make that recommendation to Casey?

c. What might go wrong with your recommendation?
2. The Journal

Lee is a student who often gets excluded from games at recess, because others think that Lee doesn’t fit in the group. One day, Terry found a journal that Lee had written with lots of private information. Terry and other students were reading it out loud, laughing, and talking about how they were going to tell everybody about Lee’s secrets. Rene is a student who was recently elected by other students as the classroom president, and works with others students and with teachers to make the school a better place. Rene has been observing the situation and doesn’t know what to do. Rene is looking for advice.
a. **Carson** is a student who is often teased by others.

b. What do you think Carson would recommend to Rene?

c. Why do you think Carson would make that recommendation?

d. What might go wrong with Carson’s recommendation?
Andy is one of the students who laughed about Lee’s private journal.

a. What do you think Andy would recommend to Rene?

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

b. Why do you think Andy would make that recommendation?

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

c. What might go wrong with Andy’s recommendation?

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

________________________________________________________________________

________________________________________________________________________
Now imagine Rene comes to you and asks: What should I do?

a. What would you recommend to Rene?

b. Why would you make that recommend that?

c. What might go wrong with your recommendation?