Development of a Prosocial-Antisocial Tease Comprehension Measure

Thesis

Presented in Partial Fulfillment of the Requirements for the Degree Master of Science in the Graduate School of The Ohio State University

By

Lauren Nicole McFadden Pino, B.S.

Graduate Program in Psychology

The Ohio State University

2021

Thesis Committee:

Dr. John C. Gibbs, Advisor

Dr. Stephen Petrill

Dr. Sarah Schoppe-Sullivan

Abstract

Teasing is a common but complex part of communication, especially when needing to distinguish between prosocial and antisocial intents/types. Its complexity may also explain the protracted development of prosocial-antisocial tease comprehension (PATC) into late childhood and continued difficulty with PATC into adolescence and adulthood. To understand this comprehension process, its development, and its difficulties, a measure is needed to study the nuances of PATC, but previous PATC measures are undermined by the lack of a theoretical basis and more intensive testing of and/or support for validity and reliability. This pair of studies begins the process of developing a new PATC measure based in social information processing theory (e.g., Dodge & Crick, 1990), which postulates that contextual and situational cues are used to navigate and comprehend complex, ambiguous social interactions. The measure features prosocial, antisocial, and ambiguous teases that vary in the number of cues and the specific cue categories (i.e., facial expression, gesture/body language, and relationship information) included. Study 1 tested the preliminary measure with a small sample, focusing on evaluation of tease types and participants' explanations for those evaluations to move beyond face validity and test the expectations for the measure. Study 2 improved the measure and identified items for a more finalized measure by having a much larger sample evaluate the tease types of possible items. The expectations for the measure from

social information processing theory and past research were all supported to some extent and replicated across the studies. Namely, participants used cues present in a tease scenario to guide their reasoning about PATC. The more cues an item had, the better the PATC and accuracy were. There were differences in item and general measure performance by tease type, and, finally, different cue categories had different effects on PATC and accuracy. Altogether, these findings expand what is known about PATC, while starting the process of developing a theoretically-based, valid, and reliable PATC measure that can be used to further explore these findings and the other aspects of PATC. More research will be needed to continue developing this measure and clarifying these findings, but these studies are a vital starting point for learning more about teasing and its comprehension.

Dedication

I would like to dedicate this to my parents, Sean, other family, and friends for their continued support and understanding. Too many times to count, my research and work limited my time with them and my ability to support them as they deserved, but, still, they stayed by my side, always encouraging me. For that, I am forever grateful.

Acknowledgments

I must recognize the many individuals who have supported, assisted, and guided me. The first should rightfully be Dr. Gibbs, whose guidance, understanding, conversation, and compassion aided me. The next acknowledgements are much deserved by Drs. Petrill, Schoppe-Sullivan, Opfer, and Johnson, who all provided vital perspectives, information, suggestions, and direction as I approached this type of research for the first time. Drs. De Boeck, Rhodes, and Rockwood must also be thanked for their invaluable advice on the measure creation and statistical analyses. Drs. Gibbs, Petrill, and Opfer additionally need to be acknowledged for their teaching on topics pertinent to prosocial-antisocial tease comprehension and measurement creation/analysis. Drs. Wegener, Fazio, and Hothersall should likewise be acknowledged for their own teaching and discussions, which led me to new theories and ideas and led me to make new connections from other areas of psychology. I must also acknowledge Dr. Beers, my fellow PSYCH1100 GTAs, and my students for continually making me think in new and creative ways about communicating and understanding psychology. Finally, for putting in the work on this research with me, I must thank my undergraduate research assistants, Vrinda Sawhney, Hanna Detwiler, and Emma Bernardi, and the individuals who consented to participate and provided the data described here.

Vita

SP 2013……Blue Valley West High School

AU 2015……..Teaching Apprentice, Department of Psychological Sciences, Kansas State University

SP 2015 – SU 2016.....Research Assistant (2 Labs), Department of Psychological Sciences, Kansas State University

SP 2016……B.S. Psychology, Kansas State University

AU 2016 to present….....Research Associate, The Ohio State University

AU 2017 to present…….....Lab Manager, Social Development Lab, The Ohio State University

AU 2017 to present…….....Graduate Teaching Associate, Department of Psychology, The Ohio State University

Publications

Jones, T. L., Wadian, T. W., Barnett, M. A., Hellmer, M. K., & Pino, L. N. (2018). Young children's ability to discriminate between antisocial and prosocial teases. *Journal of Genetic Psychology, 179*(1), 1-8.

Pino, L. N., Wadian, T. W., Jones, T. L., Sonnentag, T. L., Langley, C. A., & Barnett, M. A. (2018). Social technology and video games: A comparison of responses from adolescents and their parents. *Journal of Psychological Inquiry, 22*(1), 59-66.

Fields of Study

Major Field: Psychology

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

Teasing is a common, but complex phenomenon. An examination of 62 conversations between friends found that "jocularity" or teasing with friendly, affiliative intent made up 4% of conversational turns and more than 2% involved sarcasm, some of which would be considered teasing with mean, hurtful intent (Gibbs, 2000). Teasing is also frequently found within the home and with family cross-culturally (e.g., Eisenberg, 1986; Miller, 1986; Schieffelin, 1986). So, it is unsurprising that children begin to experience teasing from young ages (Warm, 1997), but their comprehension of teasing is not completely understood. This is especially true of the ability to distinguish prosocial (friendly) teases from antisocial (mean, hurtful) teases. Cross-sectional studies with children of different ages have found that young children (ages 5 to 8) have more difficulty with prosocial-antisocial tease comprehension (PATC) than older children (ages 10 to 12; Barnett, Burns, Sanborn, Bartel, & Wilds, 2004; Barnett, Barlett, Livengood, Murphy, & Brewton, 2010; Jones, Wadian, Barnett, Hellmer, & Pino, 2018), suggesting a developmental trend.

However, even when those of their age group demonstrated relatively consistent and differing comprehension for prosocial and antisocial teases, some children still displayed difficulty with PATC (e.g., Barnett et al., 2010). It could be suggested that they simply develop this ability later than usual, but these difficulties do not seem to always

disappear over time and development. Adolescents have also shown this continued difficulty in PATC studies (Barnett, Nichols, Sonnentag, & Wadian, 2013; Sonnentag, Barnett, Wadian, & Nichols, 2016), and a clinical treatment manual for adult residents even discusses the need for caution when teasing because the possible misunderstanding of prosocial teases as antisocial teases (Potter, Gibbs, Robbins, & Langdon, 2015). These age groups are clearly well beyond the age range when the ability to differentiate between tease types appears to be fully developed and/or ready for use (i.e., ages 10 to 12; e.g., Barnett et al., 2004), but individuals in these age groups still have difficulty using this ability for some unknown reasons and factors. These continued difficulties in older children, adolescents, and adults have not been directly explored.

To begin researching this topic, a formalized measure to gauge participants' PATC must be developed. Currently, most teasing studies use scenarios and questions that have not been validated nor tested formally for reliability, other than internal consistency, which has been reported as low as $\alpha = 0.53$ in one case (Jones et al., 2018). At most, teases that seem to work as intended in research are simply adapted and/or reused in subsequent studies. This approach leaves itself open to criticism, so, to address some of the resulting concerns, a measure of teasing scenarios with questions adapted from these past PATC studies (e.g., Jones et al., 2018) is proposed for validity and reliability testing. Once supported through psychometric investigations, this measure will make studies of possible factors in the continued difficulty with PATC more comparable and generalizable, while also enabling studies of PATC development, the underlying cognitive processes of PATC, and possible covariates.

To put into perspective the uses of the proposed measure, past PATC research suggests that having worse PATC may be associated with other sociocognitive issues, such as hostile attribution bias (e.g., Barnett et al., 2010), in which a person tends to see others' actions toward them as more negative/hostile than they likely are in reality. Thus, having a poorer PATC ability may serve as a "symptom" of sociocognitive issues that can impact interpersonal relationships and communication. For example, those with autism spectrum disorder have difficulty with nonliteral communication (e.g., Happé, 1994), and, correspondingly, their PATC ability might be impaired. Once the proposed measure is completed and found to be valid and reliable, we can directly explore these possible factors and others that may be contributing negatively or positively to PATC or may be related to differences in PATC. This research will be useful to school administrators, parents, and mental health professionals/clinicians, who all must deal with teasing, which is common in their settings and/or known to cause problems (e.g., Potter et al., 2015). The research produced with this measure will elucidate teasing and PATC and aid in understanding, so that there are fewer instances of misinterpretation and more successful communication.

Defining Teasing

Before discussing PATC measure development and the present studies, teasing must be precisely defined, and aspects of its comprehension must be described to understand the reasoning behind the studies and dissuade misconceptions. Teasing can be best defined as an "intentional provocation" that addresses "something relevant to the

target" (Keltner, Capps, Kring, Young, & Heerey, 2001). Some may consider this to be a highly unspecific definition, but that is necessary to capture the diverse forms teasing can take. For example, teasing can include things like irony (Pexman, Glenwright, Krol, & James, 2005), but it is not inherently necessary. Similarly, the definition lacks information on the valence of teasing intent as teasing can have a friendly, affiliative intent or a mean, hurtful intent (Warm, 1997; Keltner et al., 2001; Barnett et al., 2004). This is important as "teasing" is often used to solely describe teasing with hurtful intent, as can be seen in weight-based teasing research (e.g., Schvey et al., 2019), which uses "teasing" as an umbrella term that includes antisocial behaviors like bullying, victimization, and stigmatization, and other research that treats "teasing" as interchangeable with "bullying" or as a verbal form of bullying (e.g., Cornell, Gregory, Huang, & Fan, 2013; Huang & Cornell, 2019; Vessey & O'Neill, 2011). In the last decade, there has been a push to clarify for researchers and educators that teasing is not just antisocial (e.g., Mills & Carwile, 2009), and some recent research has at least started to include caveats and notes to mention that reality (e.g., Cornell et al., 2013).

It is also important to make that distinction as teasing with friendly intent may be a beneficial communicative tool. Keltner and colleagues (2001) describe how friendly teases may be used to resolve interpersonal conflicts and to teach social rules and norms. For instance, a tease may be used to diffuse a tense situation, or a person may tease a friend about their shirt being put on backwards or the clothing tag sticking out to light-heartedly call out what could be perceived as a social misstep so that it can potentially be fixed. As such, treating teasing as inherently negative and something to prevent in

general without clarifying this distinction may lead to the removal of a useful tool from people's communication toolbox (cf. Keltner, 2008).

Properly defining teasing additionally calls direct attention to the difference between intent and impact, with the major definitions of teasing and different types of teasing focusing on intentionality (Keltner et al., 2001; Barnett et al., 2004). By centering on the teaser's intent, these definitions accurately describe that intent and impact, as in how the person being teased perceives the tease, may not align consistently. Due to the provoking and thus potentially ambiguous nature of a tease, teasing a friend about their clothing tag sticking out may have all good intentions but could still be taken as a personal attack instead. This disconnect between intent and impact, especially with friendly teases, can be seen across the lifespan and notably in young children (Jones et al., 2018; Potter et al., 2015; Sonnentag et al., 2016) and speaks to the complexity of PATC, which needs proper exploration and understanding.

Following these considerations, the tease type definitions in the present studies will be adapted from Barnett and colleagues (2004) to investigate PATC: Prosocial teases are teases with friendly or affiliative intent that are meant to support relationships with others, whereas antisocial teases are teases with mean or hurtful intent that are meant to harm others.

Prosocial-Antisocial Tease Comprehension (PATC) and Its Measurement

The ability to distinguish between prosocial and antisocial teases is clearly important as teasing is a common feature of language (Gibbs, 2000), but PATC is developed over a prolonged timeline compared to other aspects of development. However, to be more specific, prosocial tease comprehension experiences protracted development, whereas antisocial tease comprehension appears to develop and mature earlier. In a study with 5- to 8-year-old children (Jones et al., 2018), even 5- and 6-year-old children were successful at evaluating the antisocial teases as unfavorable and at identifying the intent as antisocial, and 7- and 8-year-old children were similarly successful at antisocial tease comprehension and still experienced some improvement with their identification of the antisocial intent. On the other hand, 5- to 8-year-old children struggled with prosocial tease comprehension. They rated prosocial teases as unfavorable, although still significantly more favorable than antisocial teases, and the children were unsuccessful at identifying the intent as prosocial. A study by Barnett and colleagues (2004) suggests that this specific difficulty with prosocial tease comprehension may persist until ages 10 to 12, when 10- to 12-year-old children demonstrate success at evaluating the prosocial teases as favorable and at identifying the intent as prosocial. This is also when children report experiencing and using prosocial teases more than antisocial teases at home and school (Barnett et al., 2004).

In general, the development of PATC and the continued difficulties experienced with PATC (e.g., Barnett et al., 2010) are largely based in this difference between prosocial tease comprehension and antisocial tease comprehension, wherein mistaking

prosocial teases as antisocial teases is the main issue (Potter et al., 2015) and successful prosocial tease comprehension is the major developmental milestone. This may lead one to wonder whether investigating PATC as a whole is meaningful, when a researcher could focus solely on prosocial tease comprehension. However, if antisocial teases are ignored, a lot of information and understanding about the similarities and differences between prosocial and antisocial teases would not be examined. Being able to compare both types of teases and to examine them individually enables the in-depth investigation needed to study the full and complex nature of PATC.

PATC Covariates

Prior to explaining theories and research that will guide the measure creation process that will allow that nature of PATC to be properly explored, some variables that may or may not be associated with PATC ability need to be addressed in the present studies. As previously mentioned, there are likely associations between PATC and other abilities, traits, and variables, such as hostile attribution bias (e.g., Barnett et al., 2010) and autism spectrum disorder (e.g., Happé, 1994), but those associations have not yet been directly examined. However, there are variables that have already been investigated in relation to PATC. The first is age, and as described, there is a general developmental trend. However, after PATC improves with age until about 10 to 12 years old (Barnett et al., 2004), adolescents demonstrate no age-related differences in PATC ability (e.g., Barnett et al., 2013; Sonnentag et al., 2016). Yet, recent research found comprehension differences among adults of different ages (Rothermich, Giorio, Falkins, Leonard, &

Roberts, 2021), specifically finding that increased age is associated with worse comprehension. As the study used different definitions of teasing than what is used for the present studies, the possibility of age differences in adults should be explored with that limitation in mind. Another variable that has been explored is gender with past PATC measures finding no gender differences (e.g., Barnett et al., 2004; Jones et al., 2018). It is important to revisit age and gender with an improved PATC measure to better capture their potential effects on PATC or their lack of effect. Another variable to explore with an improved PATC measure is time or task duration. There is no information on the potential effects of time on PATC, but it is possible that those who spend more time considering the tease scenarios and potentially the cues that are present may have better PATC than others who spend less time on the measure. That is an important possibility to consider, and it is considered in the present studies alongside age and gender.

Theories of PATC and Measuring PATC

Looking back at previous research on PATC reveals major issues in measure development. Namely, past PATC measures suffer from minimal work and testing in terms of validity and reliability. The major measures are not validated beyond face validity, which may call all the findings of the measure into question. The reliability of these measures is mixed at best with a Cronbach's alpha at 0.53 (Jones et al., 2018), which may speak to the complexity of PATC or to issues with the items selected. To study PATC, a new measure must be developed and must address these issues. The present studies seek to do so by a) adapting the tease scenarios from past PATC measures

(e.g., Barnett et al., 2004; Jones et al., 2018) to enable review and comparison, b) developing the measure with a meaningful theoretical basis to promote construct validity, c) moving beyond face validity, and d) selecting a finalized PATC measure that aligns with expectations given the nature of PATC and the theoretical basis. With the tease scenarios from past PATC measures available, the next step is to identify a theoretical basis to act as the foundation for the measure, and that theoretical basis needs to be a possible explanatory model of what has been found thus far with PATC.

The irony comprehension literature is an excellent source for identifying a theoretical basis in this measure development process, due to how prosocial and antisocial teases often include irony and how similar the findings are to those of the teasing literature, especially the developmental and age-related trends. The age at which children begin developing a comprehension of verbal irony is around ages 5 and 6 (e.g., Dews et al., 1996; Harris & Pexman, 2003; Pexman, Zdrazilova, McConnachie, Deater-Deckard, & Petrill, 2009), and children appreciate the speaker's intent closer to ages 7 and 8 (e.g., Harris & Pexman, 2003; Pexman et al., 2005). This corresponds with how Jones and colleagues (2018) found that, although PATC ability was not fully evident in their 5- to 8-year-old participants and the overall evaluations were all still negative, the children did evaluate prosocial teases significantly more favorably than antisocial teases. Their antisocial tease comprehension was good but even better at age 7 and 8. Like prosocial tease comprehension that continues developing into late childhood (Barnett et al., 2004), aspects of irony comprehension and appreciation, especially for more complex irony, continue developing into late childhood and adolescence (e.g., Capelli, Nakagawa,

& Madden, 1990; Glenwright, Tapley, Rano, & Pexman, 2017). Thus, PATC and irony comprehension may follow similar paths in development. Additionally, PATC and irony comprehension also converge on this continued difficulty phenomenon. In one study, adults were found to only recognize ironic statements as sarcastic about half the time (Demorest, Meyers, Phelps, Gardner, & Winner, 1984). This recognition is much better than that of the children and adolescents, but adults' difficulty with recognizing sarcasm from outright deception is clear and distinct, much like continued problems with PATC.

Additionally, irony researchers have similarly sought to develop formalized measures of irony comprehension. A method to capture irony comprehension has been created, but, unfortunately, this approach is ill-suited to the desired PATC measure. Specifically, irony researchers use an eye gaze method for irony comprehension (e.g., Climie & Pexman, 2008; Pexman et al., 2011). This involves having objects present as an ironic statement is said, wherein in one object is associated with the literal meaning and the other the ironic interpretation. Eye gaze tracking allows researchers to see when the ironic interpretation is considered and chosen. At this stage of PATC measure development, this method is beyond what is needed and possible, especially with only written tease scenarios versus the more interactive approach of the eye gaze method. Adapting a PATC measure in the future to be compatible with this method may help explore the full nuance of PATC and its processing. Rather than adapting this eye gaze method right now, prominent theoretical approaches to irony comprehension will be able to help guide the creation of this new PATC measure for the present studies.

Verbal irony and, by extension, teasing rely on a number of cues to lead to the appropriate comprehension of the situation. Studies have found strong support for specific cues for verbal irony, including context (Ivanko & Pexman, 2003), personal information (i.e., occupation; Pexman & Olineck, 2002), and relationship information (e.g., Pexman & Zvaigzne, 2004; Whalen, Doyle, & Pexman, 2020). This corresponds with the social information processing theory that suggests that cues guide social interactions in complex and/or ambiguous situations (e.g., Dodge & Crick, 1990). Equivalently in the irony literature, researchers use the parallel-constraint-satisfaction framework (Pexman, 2008; Pexman, in press). To comprehend irony, this theory contends that multiple cues are processed at once and quickly, and, when the combination of cues (with situational social inferences and past experiences in mind) surpasses a threshold level of evidence/support, the ironic interpretation is considered (Pexman, 2008). Recent research extending from the parallel-constraint-satisfaction framework has clarified that no specific cue is inherently necessary for irony comprehension to occur and that, instead, different cues produce additive effects on irony comprehension (Rivière & Champagne-Lavau, 2020; Pexman, in press). Based on these approaches and related research, the theoretical basis for this PATC measure development process is the social information processing theory because some aspects of the parallel-constraint-satisfaction framework cannot be confirmed to occur in PATC at this point and thus the simpler cue-based model will be most useful at this stage. Following from the social information processing theory, a measure with tease scenarios that vary on tease type, cue categories, and numbers of cues is proposed. The cues to be included are facial expressions,

gesture/body language, and relationship information (see Figure 1 for example tease item with cues marked), which are in line with cues included and discussed in irony research (e.g., Pexman, in press; Whalen et al., 2020). However, as with all cues and scenarios, there can be very specific and personal interpretations of what they should indicate, but this is planned to be at least partially addressed in the present studies.

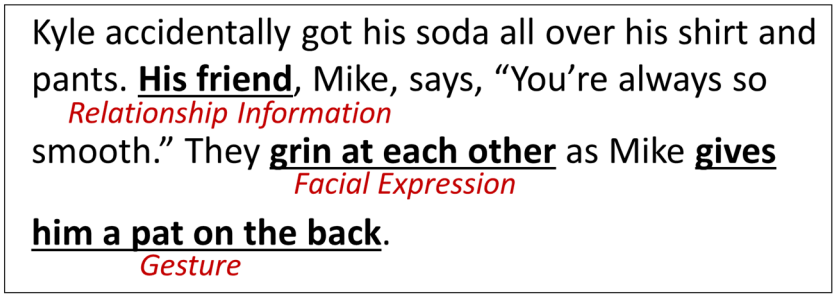

Figure 1. Example of Tease Scenario (Prosocial Tease with 3 Cues)

Measure Development

When developing a measure, there is a traditional approach, classical test theory, and a more recent and increasingly used approach, item response theory (also known as latent trait theory; Crocker & Algina, 2008; Embretson & Poggio, 2012). Classical test theory assumes that the observed measure score as a whole represents a person's true score and some amount of measurement error (Embretson & Poggio, 2012). Specific items may be investigated for difficulty and discrimination, but the evaluation of a measure in classical test theory often relies on proportion of correct items and item-item

and item-total (measure) correlations (Crocker & Algina, 2008; Embretson & Poggio, 2012). Thus, for the trait or ability a measure is designed to assess, classical test theory item analyses do not provide information about how participants at different ability levels perform on a specific item nor how and whether performance on that item can inform the researcher on a participant's true ability level (Crocker & Algina, 2008; Embretson & Poggio, 2012). In comparison, item response theory is focused on how specific item performance can be associated with and explained by an underlying latent trait that the measure captures (Crocker & Algina, 2008; Embretson & Poggio, 2012). As such, items function as individual predictors of a person's ability level on the trait, and the difference between any two scores on the measure and on particular items has specific invariant meaning, unlike classical test theory in which such differences are often contextual and different in meaning (Embretson & Poggio, 2012). This difference produces other issues for classical test theory, such as scaling artifacts and inaccurate estimates and conclusions about individual differences on the trait of interest (Embretson & Poggio, 2012), which led Embretson and Poggio (2012) to argue that item response theory has "more theoretically justifiable scaling of test scores than [classical test theory]" (p. 106).

Item response theory's invariance in meaning for score and item performance differences also enables more justifiable generalization and use of the measure with people and groups not included during the measurement development process and across equated but different versions of the measure (Crocker & Algina, 2008). Additionally, item response theory can describe individual differences in the trait of interest and individual differences in trait growth more accurately (Embretson & Poggio, 2012),

which is useful as the hope for this PATC measure is to eventually adapt the measure for children and adolescents to potentially understand developmental changes in PATC over time. Clearly, item response theory would be considered the preferred method for measurement development.

A final point on the measurement development process is that, in both classical test theory and item response theory, there may be a need to test different measure versions to allow for more items to be tested than can be given to a single sample of people. Equating across these measure versions is possible in both classical test theory and item response theory but necessitates the inclusion of an anchor test, a set of items that is administered to all participants, in addition to their specific measure version (Crocker & Algina, 2008). This anchor test is usually shorter than any of the measure versions to be equated, often about 20% of the number of items on either of the measure versions as a typical practice (Angoff, 1971, as cited in Crocker & Algina, 2008, p. 466). A general necessity of these anchor tests in classical test theory is that the anchor score distributions for the two measure versions need to be similar to meet the assumptions for equating (Angoff, 1971, as cited in Crocker & Algina, 2008, p. 465). In item response theory, that quality is not required, but the anchor tests in item response theory must instead ensure that the items on the anchor test measure the same latent trait assessed in both measure versions (Crocker & Algina, 2008). This enables the creation of a common scale on which all item parameter estimates across measure versions can be expressed, which is referred to as item calibration or more specifically person-free item calibration

(Crocker & Algina, 2008). The lack of a proper anchor tests limits many comparisons across measure versions.

The Present Studies

To begin the PATC measure development process, the present studies, Studies 1 and 2, will focus respectively on moving beyond face validity and gauging theoretical feasibility and then on making any necessary improvements upon the Study 1 measure, testing and replicating theoretical feasibility, and selecting items for the "final" PATC measure. To achieve these goals, Study 1 is an initial and smaller test to enable more in-depth data collection to move beyond face validity, while testing a starting point for the PATC measure and identifying any measure issues that would need to be addressed. As for Study 2, it can examine a variety of improved items and measure versions at a larger scale than Study 1 and be used to replicate and qualify any findings from Study 1. Both studies also assess whether the measure and items meet the expectations and hypotheses of the chosen theoretical basis. This informs whether PATC aligns with the theoretical basis and, if so, informs on construct validity.

As noted, social information processing theory (e.g., Dodge & Crick, 1990) has been used as the foundation for developing this measure. Social information processing theory postulates that we rely on cue-based reasoning in complex and ambiguous social interactions, which should include teasing. Extrapolating from this theory, it can be expected that participants use cues in tease scenarios to guide them to their tease type

evaluations and that more cues would improve accuracy in tease type evaluations. These expectations align with these first two hypotheses of interest:

Hypothesis 1: Participants use cues present in a tease scenario to guide their reasoning about prosocial-antisocial tease comprehension.

Hypothesis 2: The more cues an item has, the better the prosocial-antisocial tease comprehension and accuracy will be.

Essentially, if the social information processes are relevant, the use of cue-based reasoning and the presence of more cues in a tease scenario should improve accuracy. In Study 1, the impact of cues on PATC reasoning and performance will be assessed with an open-ended question on why they evaluated a tease scenario as a specific tease type with an eye on when they mention cues from the scenario in their reasoning. Having more cues present in a tease scenario may also be expected to inform on the use of cue-based reasoning, and more cues, potentially through the use of cue-based reasoning, may improve PATC accuracy. Studies 1 and 2 will compare cue-based reasoning (specific to Study 1) and accuracy across items with different numbers of cues present and within the tease types to test these hypotheses.

Additionally, due to the developmental differences between the tease types (e.g., Barnett et al., 2004; Jones et al., 2018) and continued difficulty into adolescence and adulthood (e.g., Barnett et al., 2013; Potter et al., 2015), it could be assumed that the tease types may involve different cognitive strategies and approaches and that different cue categories may contribute differentially to PATC. These possibilities align with the last two hypotheses of interest:

Hypothesis 3: There will be differences in item and general measure performance by tease type.

Hypothesis 4: Different cue categories will have different effects on prosocial-antisocial tease comprehension and accuracy.

In Studies 1 and 2, these hypotheses will be tested by comparing general performance on the tease types and by delving into their item performances separately with additional focus on cue categories present in these items as well as specific cue categories mentioned in evaluation explanations in Study 1. Additionally, the possible covariates of age, gender, and time/task duration and any associations with PATC will be explored in both studies.

Chapter 2. Studies

Study 1

Goals and Objectives of Study 1. Study 1 was planned as a first step in developing a PATC measure. As such, the study addresses some of the issues of past measures. The goals were to move beyond face validity and to construct a theory-based measure, while assessing if the theory aligns PATC. Like past measures, the proposed measure uses tease scenarios and asks participants to answer questions about the scenario with the focus on the identification of the correct tease type. However, past measures relied on face validity, wherein the researchers created the scenarios to demonstrate a specific tease type and used it as such without corroborating that tease type designation with independent judges outside the research team. Accordingly, this first study is designed to promote construct validity of the proposed measure by (a) having judges outside of the research team evaluate the tease scenario items and (b) exploring the item-specific and total performance of the measure to assess its alignment with its theoretical basis.

That theoretical basis, namely social information processing theory (e.g., Dodge & Crick, 1990), leads to specific expectations about participant performance on the measure and leads to our hypotheses of interest, including these first two hypotheses to investigate in Study 1:

Hypothesis 1: Participants use cues present in a tease scenario to guide their reasoning about prosocial-antisocial tease comprehension.

Hypothesis 2: The more cues an item has, the better the prosocial-antisocial tease comprehension and accuracy will be.

To assess the relevance of cues to PATC performance, the participants will be asked to explain their reasoning for their tease type evaluation. Their open-ended responses will be coded as using Cues, Personal Knowledge/Experience, Generalizations about the Nature of Teasing, Self-Generated Information about the Scenario, or Other Reasons. It is expected that most responses will have cue-based reasoning, but this content analysis of evaluation explanations will help determine how often cue-based reasoning is used. Then, cue-based reasoning can be included into statistical models that predict accuracy and clarify whether using cue-based reasoning improves PATC. Cue-based reasoning can also be explored in items with different numbers of cues, as we might expect more cues would lead to more cue-based reasoning. The number of cues and its specific effects on accuracy can likewise be tested using logistic regression and other statistical models to see if more cues translate to greater accuracy.

The next two hypotheses to explore in Study 1 focus on whether PATC varies by tease type and the presence of different cue categories:

Hypothesis 3: There will be differences in item and general measure performance by tease type.

Hypothesis 4: Different cue categories will have different effects on prosocial-antisocial tease comprehension and accuracy.

Both Hypotheses 3 and 4 can be tested by comparison at different levels. Performance on prosocial and antisocial tease items can be compared to each other and explored separately for differences; these differences may come from cue categories working in distinct ways within prosocial and antisocial tease items. It is possible that specific cue categories will provide information that is more useful in one tease type but not the other, or participants will diverge in their references to specific cue categories in their evaluation explanations.

Exploring these expectations and possibilities is the focus of the analyses in Study 1, but other analyses will need to be conducted as well. The first set of these should focus on general performance of the measure and its items. Each item will be analyzed for how many of the participants identified the tease type correctly. For items that are underperforming (i.e., more incorrect than correct responses for prosocial or antisocial teases), the items will be changed to better fit their intended tease type in future studies (e.g., Study 2). For the ambiguous teases, the rates of prosocial and antisocial evaluations should not be significantly different. It is expected that there will be more "Not Sure" responses to these items, but the important criterion is that the responses should not be skewed toward the prosocial tease type nor antisocial tease type more so than any other response. Ambiguous teases will also be changed for future studies, if they do not perform as intended.

The next set of additional analyses should include preliminary explorations of reliability and possible covariates. Past PATC measures have only had weak internal consistency reliability, which suggests that the items are not consistently measuring the

same construct (e.g., as low as $\alpha = 0.53$ in one case; Jones et al., 2018). Internal consistency reliability must be explored, but the construction of items differing by tease type, number of cues, and cue categories may make this difficult. Covariates, such as gender, age, and task duration (time spent on the measure), should also be considered to either replicate lacking effects (e.g., Barnett et al., 2004; Sonnentag et al., 2016; Jones et al., 2018) and/or address mixed findings (Rothermich et al., 2021) and absence of information.

Altogether, Study 1 will enable the first forays into developing a PATC measure that moves beyond face validity and has a theoretical basis as its guiding principle. Support for the four hypotheses of interest will determine whether this measure is viable and informative of PATC.

Methods

Sample. For Study 1, only a small group of participants was needed. A study using a similar scenario-based measure had 50 students for the same purpose for 58 vignettes (Dodge, Price, Bachorowski, & Newman, 1990), so we chose a number of participants for our similar measure that seemed suitable based on Dodge and colleague's number of participants (i.e., 20 students for 16 scenarios).

The sample was 20 students from an introductory psychology course at the Ohio State University. The participants (M_{age} = 19.05 years old; 70% women, 30% men) received course credit for their participation in this study.

Procedure and Measures. This study and its procedures were reviewed and approved by the Ohio State University Social and Behavioral Sciences Institutional Review Board.

At a scheduled time, participants were brought in and taken to a private computer space. A member of the research team was present to set them up and answer general questions about the procedure. As this is a new measure, it is helpful to have a member of the research team present when the participant completes the measure, so that any instances of confusion can be found, clarified, and then fixed before the future studies. The study itself will be completed online through the Qualtrics survey system.

The survey begins with the informed consent form that gives information about the study, their rights, and a place to choose to consent or not to consent to participate in the study. If they did not consent, they would be taken to the end-of-study debriefing and thanked. If they did consent, they would begin the study.

Participants were first asked to give their month and year of birth (to calculate their age at the time of their study participation) and their gender, and then they start the primary measure.

The measure has 16 tease scenarios, and each tease scenario has one of three tease types (seven prosocial teases, seven antisocial teases, and two ambiguous teases) and a certain number of cues present (zero to three) with different cue categories (gesture/body language, facial expression, relationship information). All possible combinations of number of cues and cue categories are included with both the prosocial teases and antisocial teases, whereas the two ambiguous teases both have no cues at all. Each tease

scenario is between two named, fictional boys of unspecified age with one boy teasing the other, which is a format used in previous PATC studies and measures (e.g., Jones et al., 2018). Most of the tease scenarios were adapted from those past PATC measures (e.g., Barnett et al., 2004; Barnett et al., 2010; Jones et al., 2018). The order of the tease scenarios was randomized. The specific tease scenarios are available in the Appendix A.

Each tease scenario has two questions. The first question asks the participant to identify what the tease type of the tease in the scenario was (prosocial ["Friendly"], antisocial ["Mean"], or ambiguous ["Not Sure"]) from their perspective. The second question is open-ended and asks them to explain their reasoning for their response to the first question. These responses were later coded into response categories: Cues (i.e., participant cites the cues presented in the scenario), Personal Knowledge/Experience (i.e., participant bases their reasoning on their own experiences and not what is presented in the scenario), Generalizations about the Nature of Teasing (i.e., participant uses general assumptions about and interpretations of teasing), Self-Generated Information about the Scenario (i.e., participant creates and uses information not actually presented in the scenario), or Other Reasons. This response/evaluation explanation categorization and coding was completed by three raters. The raters' specific instructions can be found in the Appendix B. Any disagreements in categorization were discussed, and a final response category assignment would be made with the agreement of all three raters.

When they finished the PATC measure and thus the study, participants were debriefed on the full purpose of the study and thanked for their participation. They were also given the opportunity to be informed of the study results at a later date.

Results

The analyses for Study 1 focus on general analyses of the measure and content analysis of the open-ended evaluation explanations, as well as analyses of possible covariates. The general analyses and content analysis are intended to examine whether the hypotheses and expectations of the measure are supported and met, while also identifying any issues that must be addressed prior to Study 2 and further measure development.

Reliability. Measuring the reliability of this measure presents specific issues for typical approaches to reliability estimation, especially with estimates of internal consistency reliability. Using Cronbach's alpha to estimate the reliability of and/or the internal consistency of a measure is a common practice (Sijtsma, 2009), and it can be understood as "the correlation of [the] test with itself" or the interconnections among the items (Tavakol & Dennick, 2011, p. 53). Beyond the many criticisms of and issues with Cronbach's alpha (Sijtsma, 2009; Tavakol & Dennick, 2011), Cronbach's alpha is not a useful approach to measure the reliability and the internal consistency of this PATC measure.

To demonstrate, when estimated for the entire PATC measure, the Cronbach's alpha is 0.537, much below typical standards of 0.7 to 0.95 (e.g., Bland & Altman, 1997). Even removing the ambiguous items, which are absent of cues in their construction, actually decreases the Cronbach's alpha to 0.456. Splitting the measure into prosocial items and antisocial items likewise does not fix the problem, respectively producing Cronbach's alphas of 0.535 and 0.194. Part of the issue may be the small sample size and

some items with zero variance (i.e., due to all participants getting the item correct), but the more specific issue is that the differential construction of the items with different numbers and combinations of different cues produces items that may not easily relate to one another.

For example, compare an antisocial 2-cue tease item with cues for relationship information and facial expression to an antisocial 1-cue tease item with a cue for gesture. Although both items are antisocial, they differ in the number of cues and in the type of cues included. It makes sense to expect a weaker correlation, and in this study, that is exactly what you find (r = -0.015, p = 0.951). If instead two more similar items are compared, such as a prosocial 2-cue tease with cues for gesture and relationship information and a prosocial 1-cue tease with a cue for gesture, the relationship is stronger, but not always strong on its own (this example: r = 0.312, p = 0.181).

Another reliability measurement that takes this into account would be preferred, such as composite reliability (e.g., Brunner & Süβ, 2005). However, due to the small sample size and zero variance on some items, such reliability analyses are not possible at this stage. These issues limiting reliability analyses are also exacerbated by item-specific problems that will be discussed throughout the results.

Demographics and Task Duration. The descriptive statistics for age and task duration can be found in Table 1. Information on gender was included in the sample description.

Table 1. Descriptive Statistics for Age and Task Duration in Study 1

Variable	Mean	Standard Deviation	Range
Age (in years)	19.05	1.12	18 - 22
Task Duration in Minutes (excluding errors)	25.63	9.59	9.72 - 46.68

When task duration, gender, and age were used in logistic regression models to predict accuracy for prosocial tease item accuracy, antisocial tease item accuracy, and ambiguous tease item accuracy, all produced nonsignificant results (all $p > 0.05$). This aligns with past research that found no gender difference for PATC in children and adolescents (e.g., Barnett et al., 2004; Jones et al., 2018) and no age difference for PATC in older adolescents (e.g., Barnett et al., 2013; Sonnentag et al., 2016), although recall that recent research suggests that there may be differences among adults by age (Rothermich et al., 2021). There were a handful of significant correlations between task duration, gender, and age and other variables (see Table 2). However, these results need qualification, given two specific issues with this study.

First, due to an error in task delivery, 6 (30%) of the 20 participants' task durations were inaccurate. This led to their exclusion from analyses involving task duration. Second, in this study, 14 (70%) of the 20 participants were women, so, although no significant gender difference was found in the model, it could be due to having fewer men for comparison. Overall, there seem to be little to no effects of task duration, gender,

and age in this study, but more studies, such as Study 2 and beyond, are needed to fully clarify their effects or lack thereof.

Table 2. Significant Correlations Between Task Duration, Gender, and Age and Other Variables in Study 1

Variable	Task Duration	Gender (1 = Man, 2 = Woman)	Age
Accuracy in Antisocial 1-Cue Tease Items	-0.562*	0.122	0.357
Accuracy in Prosocial 2-Cue Tease Items	-0.400	-0.449*	-0.123
Cue-Based Reasoning in the Antisocial 1-Cue (Gesture) Tease Item	-0.050	0.491*	-0.090
Cue-Based Reasoning in the Antisocial 3-Cue Tease Item	-0.122	-0.535*	0.493*
Use of Facial Expression Cue in Evaluation Explanations for Antisocial Tease Items	0.060	-0.045	-0.246*

a. Nonsignificant correlations are not displayed.
*$p < 0.05$

General Analyses

All participants demonstrated above chance performance on prosocial and antisocial tease items, and although the accuracy on ambiguous tease items is not as meaningful to this investigation, 60% of participants exhibited above chance performance. The average accuracy was 80.7% for prosocial items and 80.0% for antisocial items. As accuracy on ambiguous items does not necessarily work the same as

in prosocial and antisocial items, it is no surprise that it was 35% or just slightly above the chance level of 33.3%.

Of the prosocial and antisocial tease items that did not have zero variance, all but two had accuracies that significantly differed from chance (33.3%; all *p*s < 0.01). The two items (prosocial 1-cue tease item with gesture cue and antisocial 1-cue tease item with relationship information cue) that had accuracies that did not differ from chance were items that experienced specific problems with how they were constructed, which will be discussed in the next analyses. As expected, accuracy on the two ambiguous tease items did not differ from chance (both *p*s > 0.05).

As ambiguous tease items should be minimally skewed toward friendly or mean, the skewness of each of the two ambiguous tease items was explored. Shapiro-Wilk tests found that both items are significantly skewed (ambiguous tease item referring to teeth: $W(20) = 0.773$, $p < 0.001$; ambiguous tease item referring to grades: $W(20) = 0.675$, $p < 0.001$). Their values of skewness are also suggesting that the tease type evaluations were skewed (ambiguous tease item referring to teeth: skewness = -0.549, SE = 0.512; ambiguous tease item referring to grades: skewness = 1.283, SE = 0.512). One appears to have a moderate negative skew (fewer mean evaluations than friendly evaluations) and the other a stronger positive skew (more mean evaluations than friendly evaluations).

Additional general information on the measure and accuracy will be presented in relation to specific relevant hypotheses.

Hypothesis 1: Participants use cues present in a tease scenario to guide their reasoning about prosocial-antisocial tease comprehension.

Content Analysis. Examples for different evaluation explanation coding categories can be seen in Table 3, and specific coding instructions for raters can be found in Appendix B.

Table 3. Examples of Different Coding Categories for Content Analysis

Content Analysis Category	Example
Cues (Cue-Based Reasoning)	"**they aren't friends** so the tease was intentionally aimed at hurting the other person."
Personal Knowledge/Experience	"Adam's comment reminds me of something my grandpa would say to me in this situation as well. I don't believe Adam's intent is to hurt Gary."
Generalizations about the Nature of Teasing	"Regardless of Rick's reaction, I think this can still be considered friendly even if Rick didn't laugh. Bags under someone's eyes aren't a permanent physical feature so I don't think it's really mean at all."
Self-Generated Information about the Scenario	"Kyle does this a lot so it seems to be a joke between the two which means that Mike wasn't trying to hurt Kyle's feelings with this joke"
Other Reasons	"Funny interaction."

Interrater Reliability. The 3 raters largely agreed on coding decisions with all three in agreement 59.1% of the time and at least two in agreement 85.3% of the time. Most (68.7%) of all disagreements came from the "Other Reasons" category, and because

the primary focus was the “Cues” category and not the “Other Reasons” category, the interrater reliability statistics will also include variations with or without items marked in the “Other Reasons” category when meaningful.

When more than two raters are used, Cohen’s κ can quantify the agreement between rater dyads, whereas Fleiss’s κ can quantify the agreement between more than two raters (Landis & Koch, 1977). Both approaches to interrater reliability have numerical benchmarks for satisfactory agreement and can provide different but, at times, overlapping information. The rater dyads in this study all had “moderate” agreement with their category assignments (Cohen’s $\kappa_{1,2} = 0.571$; Cohen’s $\kappa_{2,3} = 0.529$; Cohen’s $\kappa_{1,3} = 0.424$). The mean Cohen’s κ across the rater dyads was 0.508, which is equal to the Fleiss’s κ for all three raters when the full data was included. In both cases, this suggests satisfactory or “moderate” agreement among the raters. When data for the “Other Reasons” category is ignored, the Fleiss’s κ is 0.670, which is considered “substantial” agreement. As the “Cues” category was the focus, the Fleiss’s κ for the “Cues” category in the full data is 0.701 (“substantial” agreement), and the Fleiss’s κ for the “Cues” category when the “Other Reasons” category data is ignored is 0.757 (“substantial” agreement). Interrater reliability for the three raters is satisfactory in general and excellent for the category of interest.

Category Use. In the open-ended evaluation explanations, all 320 explanations (1 per item, 16 per participant) were coded into 5 categories as described previously. 57.8% of explanations were categorized as Cues or cue-based reasoning, followed by 16.9% for Other Reasons, 13.4% for Self-Generated Information about the Scenario, 8.4% for

Generalizations of the Nature of Teasing, and 3.4% for Personal Knowledge/Experience. The proportion for cue-based reasoning was significantly greater than the second most common categorization (Other Reasons), $t(319) = 9.607$, $p < 0.001$, and the third most common non-miscellaneous categorization (Self-Generated Information about the Scenario), $t(319) = 11.038$, $p < 0.001$.

This is also true for prosocial and antisocial tease items separately. In prosocial tease items, 64.3% of explanations were categorized as cue-based reasoning. The proportion for cue-based reasoning was significantly greater than the next most common categorizations (equally Other Reasons and Self-Generated Information about the Scenario), $t(140) = 8.761$, $p < 0.001$. In antisocial tease items, 60.7% of explanations were categorized as cue-based reasoning. The proportion for cue-based reasoning was significantly greater than the second most common categorization (Other Reasons), $t(140) = 6.512$, $p < 0.001$, and the third most common non-miscellaneous categorization (Self-Generated Information about the Scenario), $t(319) = 7.927$, $p < 0.001$. Additionally, using cue-based reasoning on prosocial teases was positively correlated with cue-based reasoning on antisocial teases ($r = 0.482$, $p = 0.031$).

This pattern is not seen with ambiguous (0-cue) tease items, wherein 30% of explanations were categorized as Other Reasons, followed by 25% for cue-based reasoning, 22.5% for the Generalizations of the Nature of Teasing, 20% for Self-Generated Information about the Scenario, and 2.5% for Personal Knowledge/Experience. There are no significant differences among the three most common categorizations, Other Reasons, cue-based reasoning, and Generalizations of the

Nature of Teasing, within the ambiguous tease items (all *p*s > 0.05). Using cue-based reasoning on ambiguous teases was positively correlated with cue-based reasoning on antisocial teases ($r = 0.515$, $p = 0.020$).

Cue-Based Reasoning and Accuracy. 96.2% of cases using cue-based reasoning led to the correct answer[1]. Accuracy on prosocial teases was positively correlated with cue-based reasoning on prosocial teases ($r = 0.719$, $p < 0.001$). Accuracy on antisocial teases was positively correlated with cue-based reasoning on antisocial teases ($r = 0.648$, $p = 0.002$). Accuracy on ambiguous teases was positively correlated with cue-based reasoning on ambiguous teases ($r = 0.698$, $p = 0.001$).

Content Analysis Categories and Number of Cues. For all tease items, the evaluation explanation categorizations differed by number of cues in the tease item, $\chi^2(12, N = 320) = 28.353$, $p = 0.005$. This is driven by the significant difference in cue-based reasoning between ambiguous (0-cue) tease items (25% use) and prosocial and antisocial (1- to 3-cues) tease items (63% use), $t(319) = 4.626$, $p < 0.001$. As all categorization differences and cue-based reasoning differences by cue number within prosocial or antisocial tease items were non-significant (all *p*s > 0.05), a future, ideally larger study would need to clarify how cue number interacts with different categories of reasoning and cue-based reasoning specifically.

[1] Only 7 cases of cue-based reasoning (3.8%) led to a wrong answer out of 185 cue-based responses: 6 “Not Sure” responses on prosocial/antisocial teases (All for 1-cue teases) & 1 “Mean” response on an ambiguous tease.

For now, an examination of the number of cues mentioned in the evaluation explanations[2] provides another way of investigating the effect of cue number on cue-based reasoning. For prosocial and antisocial tease items, the number of cues mentioned did not differ by tease type (prosocial or antisocial), $\chi^2(2, N = 175) = 1.931$, $p = 0.381$, but did differ by cue number, $\chi^2(4, N = 175) = 28.172$, $p < 0.001$. Follow-up analyses suggest that this is driven mainly by an increase in cues mentioned from 1-cue tease items to 2-cue tease items, $t(146) = 4.527$, $p < 0.001$.

Additional issues discovered during content analysis. Although the content analysis was meant to help validate the measure and its theoretical foundations as well as give further insight into PATC generally, the participants' written explanations of their tease type evaluations also illuminated some issues with specific tease items. The first issue involved two items being taken as far meaner than intended due to the verbal tease within the scenario, which will be further explained in the discussion section of this specific study. The second issue involved the perspective of specific information and cues, which ran counter to how PATC is postulated to work.

If intent is indeed the determining factor between a prosocial tease and an antisocial tease, then it is important that the cues speak to the intentions of the teaser and not just the perspective of the person being teased. However, a couple of teases included relationship information cues that were from the perspective of the teasee and not the teaser. This error was actually called out by some participants, who expressed difficulty

[2] Cues mentioned within explanations categorized as cue-based reasoning were coded by one person after the three raters categorized all the explanations.

using the cue with that direction of relationship information to identify the intent and tease type. Here is one such example from one participant:

> **Antisocial 1-Cue Tease with Cue for Relationship Information:** Peter arrived late to work and got scolded by his boss. Isaac, a coworker he dislikes, asks, "So, when are you getting your employee of the month award?"
>
> **Tease Type Evaluation:** Not sure
>
> **Evaluation Explanation:** "It's hard to tell whether or not this is a friendly or mean remark. Even though [P]eter doesn't like Isaac, it doesn't mean that Isaac was being mean."

Unfortunately, this issue especially affected the antisocial 2-cue tease items, which had accuracies lower than almost all other 2-cue tease items, antisocial or prosocial, and lower accuracy than 2 out of 3 of the antisocial 1-cue tease items. As there are only three antisocial 2-cue tease items, having two out of the three affected has consequences on the analyses and results discussed next. Table 4 exhibits the items specifically affected, along with the original text and the alternation to address this issue. The performance of altered versions of the items in future studies will be necessary to clarify the role of these items and the nature of antisocial tease comprehension overall.

Table 4. Teases with Relationship Information in the Wrong Direction

Tease Item Identifier	Original Tease Scenario Text	Alteration of the Item for Future Use (e.g., Study 2)
Antisocial 2-Cue Tease with Cues for Facial Expression and Relationship Information	Greg accidentally wore his little brother's shirt to class, and it has Mickey Mouse on the front. Brian, **a classmate Greg hates**, smirks and says, "I had a shirt just like that when I was in preschool." Greg scowls.	Greg accidentally wore his little brother's shirt to class, and it has Mickey Mouse on the front. Brian, **a classmate who hates Greg**, smirks and says, "I had a shirt just like that when I was in preschool." Greg scowls.
Antisocial 2-Cue Tease with Cues for Gesture and Relationship Information	Carter came to class with his shirt on backwards. Will, **who always gets on his nerves**, points at the tag sticking out and says, "Maybe you should get dressed with the lights on next time."	Carter came to class with his shirt on backwards. Will, **who purposefully gets on his nerves**, points at the tag sticking out and says, "Maybe you should get dressed with the lights on next time."
Antisocial 1-Cue Tease with Cue for Relationship Information	Peter arrived late to work and got scolded by his boss. Isaac, **a coworker he dislikes**, asks, "So, when are you getting your employee of the month award?"	Peter arrived late to work and got scolded by his boss. Isaac, **a coworker who dislikes Peter**, asks, "So, when are you getting your employee of the month award?"

Hypothesis 2: The more cues an item has, the better the prosocial-antisocial tease comprehension and accuracy will be.

Whether cues are used in PATC can also be clarified if different performance is found with different amounts of cues and the use of cue-based reasoning, and this was indeed the case (see Figure 2 for accuracy by cue number). When ambiguous tease items are excluded because they have no specific cues, accuracy on prosocial and antisocial

tease items are significantly positively correlated with number of cues in the item (r = 0.209, $p < 0.001$) and the use of cue-based reasoning in the evaluation explanations (r = 0.508, $p < 0.001$). Likewise, in a logistic regression model, cue number significantly predicts accuracy in prosocial and antisocial tease items, β = 1.106, SE = 0.307, Wald $\chi^2(1) = 12.975$, OR = 3.022, $p < 0.001$, 95% CI = [1.656, 5.517], and cue-based reasoning also significantly predicts accuracy in prosocial and antisocial tease items, β = 3.155, SE = 0.450, Wald $\chi^2(1) = 49.251$, OR = 23.447, $p < 0.001$, 95% CI = [9.715, 56.588]. Table 5 shows the logistic regression coefficient, Wald test and p-values, and odds ratio for cue number and cue-based reasoning.

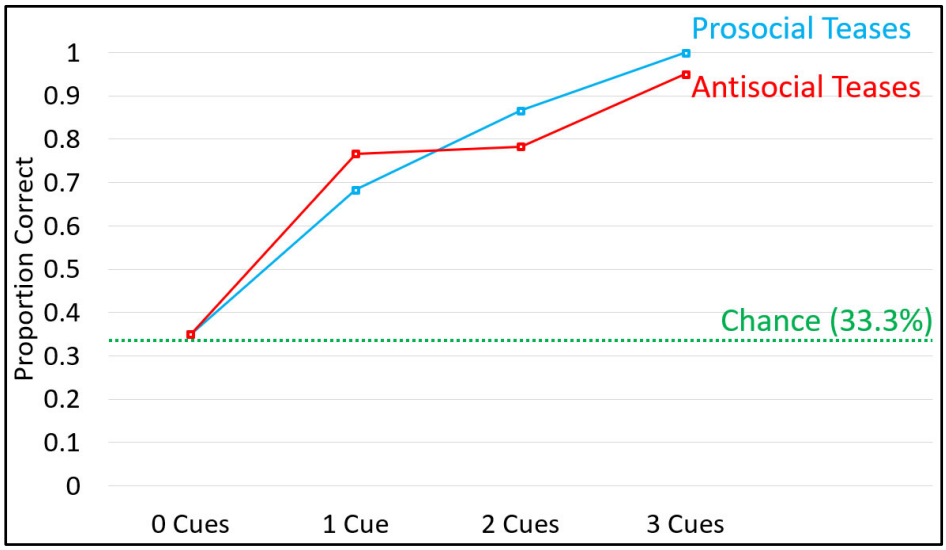

Figure 2. Accuracy for Prosocial and Antisocial Tease Items by Cue Number in Study 1

The shared point on the graph for 0 cues refers to accuracy on ambiguous tease items and is included on the lines for both prosocial and antisocial tease items for comparison.

This model was able to correctly classify 91.1% of those with accurate performance on prosocial and antisocial tease items and 47.3% of those with inaccurate performance on prosocial and antisocial tease items, producing an overall success rate of 82.5%. On its own, cue number explained 7.4% of the variance in accuracy on prosocial and antisocial tease items (Nagelkerke $R^2 = 0.074$), and the model with cue number was significantly more predictive over a model with only the intercept, $\chi^2(1, N = 280) = 13.272$, $p < 0.001$. When the use of cue-based reasoning is added as another predictor to the model, the amount of variance in accuracy on prosocial and antisocial tease items explained increases to 43.3% (Nagelkerke $R^2 = 0.433$), and its inclusion into the model significantly improves the model's predictive ability, compared to the model with just cue number and the intercept, $\chi^2(1, N = 280) = 75.799$, $p < 0.001$. Overall, the presence of more cues and using cues to reason about prosocial and antisocial teases increases accuracy. Cue number and cue-based reasoning explain a large portion of the variance in accuracy on prosocial and antisocial tease items and produce a significantly predictive model of accuracy on prosocial and antisocial tease items and thus PATC.

Table 5. Logistic Regression Predicting Accuracy on Prosocial and Antisocial Tease Items from Cue Number and Use of Cue-Based Reasoning in Study 1

Predictor	β	Wald χ^2	*p*-value	Odds Ratio
Cue Number	1.106	12.975	< 0.001	3.022
Cue-Based Reasoning	3.155	49.251	< 0.001	23.447
Intercept	-1.647			

Prosocial Teases. Accuracy on prosocial tease items is significantly positively correlated with the number of cues in the item ($r = 0.292$, $p < 0.001$) and the use of cue-based reasoning in the evaluation explanations ($r = 0.580$, $p < 0.001$). Likewise, in a logistic regression model, cue number significantly predicts accuracy in prosocial tease items, $\beta = 1.774$, SE = 0.551, Wald $\chi^2(1) = 10.360$, OR = 5.894, $p = 0.001$, 95% CI = [2.001, 17.359], and cue-based reasoning also significantly predicts accuracy in prosocial tease items, $\beta = 4.048$, SE = 0.815, Wald $\chi^2(1) = 24.675$, OR = 57.261, $p < 0.001$, 95% CI = [11.595, 282.777]. Table 6 shows the logistic regression coefficient, Wald test and p-values, and odds ratio for cue number and cue-based reasoning.

This model was able to correctly classify 93.8% of those with accurate performance on prosocial tease items and 63% of those with inaccurate performance on prosocial tease items, producing an overall success rate of 87.9%. On its own, cue number explained 14.9% of the variance in accuracy on prosocial tease items (Nagelkerke $R^2 = 0.149$), and the model with cue number was significantly more predictive over a model with only the intercept, $\chi^2(1, N = 140) = 13.676$, $p < 0.001$. When

the use of cue-based reasoning is added as another predictor to the model, the amount of variance in accuracy on prosocial tease items explained increases to 57.6% (Nagelkerke $R^2 = 0.576$), and its inclusion into the model significantly improves the model's predictive ability, compared to the model with just cue number and the intercept, $\chi^2(1, N = 140) = 48.733$, $p < 0.001$. Overall, the presence of more cues and using cues to reason about prosocial teases increases accuracy. Cue number and cue-based reasoning explain a large portion of the variance in accuracy on prosocial tease items and produce a significantly predictive model of accuracy on prosocial tease items and thus prosocial tease comprehension.

Table 6. Logistic Regression Predicting Accuracy on Prosocial Tease Items from Cue Number and Use of Cue-Based Reasoning in Study 1

Predictor	β	Wald χ^2	*p*-value	Odds Ratio
Cue Number	1.774	10.360	0.001	5.894
Cue-Based Reasoning	4.048	24.675	< 0.001	57.261
Intercept	-2.797			

Consistent with the regression models, prosocial tease accuracy increases with more cues within the overall model. Participants were more accurate on prosocial 2-cue tease items than on prosocial 1-cue tease items, $t(19) = -2.77$, $p = 0.012$. Participants were more accurate on the prosocial 3-cue tease item than on prosocial 2-cue tease items,

$t(19) = -2.99$, $p = 0.008$. Participants were more accurate on the prosocial 3-cue tease item than on prosocial 1-cue tease items, $t(19) = -4.79$, $p < 0.001$.

Next, a persons-as-effect-sizes approach (Grice et al., 2020) is used to further clarify whether accuracy increases with more cues by identifying how many participants followed a general increasing trend in accuracy by number of cues within each tease type. For prosocial teases and in comparison to the t-test results, 40% saw their accuracy increase from prosocial 1-cue tease items to prosocial 2-cue tease items, whereas 55% had equal accuracy on prosocial 1-cue tease items and prosocial 2-cue tease items. Notably, most of that 55% (35% of the participants) were at ceiling with accurate answers to all prosocial tease items. As they would not be able to demonstrate the increasing trend in this case, they will be set aside to better clarify the patterns. When those at ceiling are excluded, 92.31% saw their accuracy increase from prosocial 1-cue tease items to the prosocial 3-cue tease item, whereas 7.69% had equal accuracy on prosocial 1-cue tease items and the prosocial 3-cue tease item.

When those who generally increased their accuracy from the least cues to the most cues are explored, 16.67% increased their accuracy with each additional cue, 33.33% had the same accuracy on 1-cue and 2-cue tease items and then increased with the 3-cue tease item, and 50% increased their accuracy from the 1-cue tease items to the 2-cue tease items and had the same accuracy on 2-cue and 3-cue tease items. Only one participant did not match this general increasing trend and was not at ceiling, but they only missed one prosocial tease item in total. Because the one incorrect item was a 2-cue

tease item, their specific accuracy pattern involved a decrease from 1-cue tease items to 2-cue tease items and then an increase from 2-cue tease items to the 3-cue tease item.

Antisocial Teases. Accuracy on antisocial tease items is significantly positively correlated with the use of cue-based reasoning in the evaluation explanations ($r = 0.439$, $p < 0.001$), but not with the number of cues in the item ($r = 0.128$, $p = 0.133$). Based on that, it is no surprise that cue number did not significantly predict accuracy in antisocial tease items in a logistic regression model, $\beta = 0.692$, SE = 0.372, Wald $\chi^2(1) = 3.468$, OR = 1.997, $p = 0.063$, 95% CI = [0.964, 4.137], whereas cue-based reasoning significantly predicts accuracy in antisocial tease items, $\beta = 2.553$, SE = 0.550, Wald $\chi^2(1) = 21.567$, OR = 12.843, $p < 0.001$, 95% CI = [4.373, 37.720]. Table 7 shows the logistic regression coefficient, Wald test and p-values, and odds ratio for cue number and cue-based reasoning.

This model was able to correctly classify 88.4% of those with accurate performance on antisocial tease items and 32.1% of those with inaccurate performance on antisocial tease items, producing an overall success rate of 77.1%. On its own, cue number explained 2.7% of the variance in accuracy on antisocial tease items (Nagelkerke $R^2 = 0.027$), and the model with cue number was not significantly more predictive over a model with only the intercept, $\chi^2(1, N = 140) = 2.372$, $p = 0.124$. When the use of cue-based reasoning is added as another predictor to the model, the amount of variance in accuracy on antisocial tease items explained increases to 31.5% (Nagelkerke $R^2 = 0.315$), and its inclusion into the model significantly improves the model's predictive ability, compared to the model with just cue number and the intercept, $\chi^2(1, N = 140) = 28.705$, p

< 0.001. Overall, using cues to reason about antisocial teases increases accuracy, but the presence of more cues did not. Cue number and cue-based reasoning explain a large portion of the variance in accuracy on antisocial tease items and produce a significantly predictive model of accuracy on antisocial tease items and thus antisocial tease comprehension, but that is mainly driven by cue-based reasoning, not cue number. As stated previously, this could be due to the issues with the antisocial 2-cue items. However, this could also suggest that the number of cues matters less in antisocial teases, despite cue-based reasoning still being predictive of accuracy on antisocial tease items. This will be revisited in Study 2, which addresses the specific item issues and has a larger sample size, to better understand the role of cues in antisocial tease comprehension.

Table 7. Logistic Regression Predicting Accuracy on Antisocial Tease Items from Cue Number and Use of Cue-Based Reasoning in Study 1

Predictor	β	Wald χ^2	*p*-value	Odds Ratio
Cue Number	0.692	3.468	0.063	1.997
Cue-Based Reasoning	2.553	21.567	< 0.001	12.843
Intercept	-0.856			

Consistent with the regression models, antisocial tease accuracy generally increases with more cues as in the overall model. However, participants were not more accurate on antisocial 2-cue tease items than on antisocial 1-cue tease items, $t(19) = -$

0.20, $p = 0.847$. This may reflect relatively strong performance with the antisocial 1-cue tease items or the issues with the antisocial 2-cue tease items. Despite that, the other t-tests still support a general increasing trend in accuracy with more cues. Participants were more accurate on the antisocial 3-cue tease item than on antisocial 2-cue tease items, $t(19) = -2.70$, $p = 0.014$. Participants were more accurate on the antisocial 3-cue tease item than on antisocial 1-cue tease items, $t(19) = -2.24$, $p = 0.037$.

Using a person-as-effect-sizes approach from the antisocial tease items is complicated by knowing that there were problems with the 2-cue tease items, which likely led to somewhat lower than expected accuracy. This must be considered when evaluating the accuracy patterns. As with the prosocial tease items, 20% of participants were at ceiling for accuracy on antisocial tease items, answering all items correctly, and will be similarly excluded in calculating the next percentages. When those at ceiling are excluded, 68.75% saw their accuracy increase from antisocial 1-cue tease items to the antisocial 3-cue tease item, whereas 25% had equal accuracy on antisocial 2-cue tease items and the antisocial 3-cue tease item.

When those who generally increased their accuracy from the least cues to the most cues are explored, 0% increased their accuracy with each additional cue, 30% had the same accuracy on 1-cue and 2-cue tease items and then increased with the 3-cue tease item, and 70% increased their accuracy from the 1-cue tease items to the 2-cue tease items and had the same accuracy on 2-cue and 3-cue tease items. Unlike with the prosocial tease items, 30% of participants did not match this general increasing trend and were not at ceiling, and most (83.33%) did not follow the increasing pattern because they

had difficulty with at least one 2-cue tease item and namely the 2-cue tease items with issues. For those experiencing difficulty with a 2-cue tease item, their specific accuracy pattern involved a decrease from 1-cue tease items to 2-cue tease items and then an increase from 2-cue tease items to the 3-cue tease item. One participant exhibited a general decreasing accuracy trend, which could reflect issues with the 2-cue tease items and only having one item for 3-cue teases.

Ambiguous Teases. Again, as stated, accuracy on ambiguous items is less meaningful as to be accurate necessitates choosing the middle or "Not Sure" option. Accuracy on ambiguous tease items is significantly positively correlated with the use of cue-based reasoning in the evaluation explanations ($r = 0.67$, $p < 0.001$). Likewise, in a logistic regression model, cue-based reasoning significantly predicts accuracy in ambiguous tease items, $\beta = 3.807$, SE = 1.162, Wald $\chi^2(1) = 10.725$, OR = 45.000, $p = 0.001$, 95% CI = [4.611, 439.164]. Table 8 shows the logistic regression coefficient, Wald test and p-values, and odds ratio for cue-based reasoning.

This model was able to correctly classify 64.3% of those with accurate performance on ambiguous tease items and 96.2% of those with inaccurate performance on ambiguous tease items, producing an overall success rate of 85.0%. The use of cue-based reasoning explained 50.5% of the variance in accuracy on antisocial tease items (Nagelkerke $R^2 = 0.505$), and the model with cue-based reasoning was significantly more predictive over a model with only the intercept, $\chi^2(1, N = 40) = 18.260$, $p < 0.001$. Even when accuracy is less meaningful, using cues to reason about ambiguous teases increases accuracy. Cue-based reasoning explains a large portion of the variance in accuracy on

ambiguous tease items and produces a significantly predictive model of accuracy on ambiguous tease items.

Table 8. Logistic Regression Predicting Accuracy on Ambiguous Tease Items from Use of Cue-Based Reasoning in Study 1

Predictor	β	Wald χ^2	*p*-value	Odds Ratio
Cue-Based Reasoning	3.807	10.725	0.001	45.000
Intercept	-1.609			

Hypothesis 3: There will be differences in item and general measure performance by tease type.

On average, participants were equally accurate on prosocial items (M = 5.65 out of 7 total items, SD = 1.23) and antisocial items (M = 5.60 out of 7 total items, SD = 1.10), $t(19) = 0.149$, $p = 0.883$, 95% CI = [-0.654, 0.754]. Because the sample is small and an outlier might greatly affect those results, the accuracy on prosocial versus antisocial items was looked at on an individual level. 20% of participants were equally accurate on prosocial and antisocial items, getting the same exact number of items correct for each tease type. Including that group, 75% of participants were almost equally accurate on prosocial and antisocial items, only having at most a difference of 1 item between the number of correct items for each tease type. Altogether, the participants largely perform similarly on prosocial and antisocial items in terms of accuracy.

An interesting finding is that, when you compare the proportions of participants doing equally well on both tease types or doing better on one type over the other, 20% do equally well on both types, and the remaining 80% are equally split. 40% do better on prosocial teases than on antisocial teases (10% did better on prosocial items by at least 2 items). 40% do better on antisocial teases than on prosocial teases (15% did better on antisocial items by at least 2 items).

Due to item-specific issues preventing some direct comparisons, differences by tease type will be explored further in tests of the next hypothesis, regarding the differential effects of cues.

Hypothesis 4: Different cue categories will have different effects on prosocial-antisocial tease comprehension and accuracy.

To investigate whether cues may produce different effects on PATC, correlations between the presence of different cue categories and accuracy on prosocial and antisocial items were examined. Only facial expression cues were significantly correlated with accuracy on prosocial and antisocial tease items, $r = 0.226$, $p < 0.001$. Adding the presence of facial expression cues as a predictor to a logistic regression model with cue number predicting accuracy in prosocial and antisocial tease items significantly improves predictive ability, compared to a model with just cue number and the intercept, $\chi^2(1, N = 280) = 5.614$, $p = 0.018$. Even adding it as a predictor to a regression model with cue number and the use of cue-based reasoning predicting accuracy in prosocial and antisocial tease items significantly improves predictive ability, compared to a model with

just cue number, cue-based reasoning, and the intercept, $\chi^2(1, N = 280) = 5.406$, $p = 0.020$. Table 9 shows the logistic regression coefficient, Wald test and p-values, and odds ratio for cue number, cue-based reasoning, and presence of a facial expression cue. Clearly, the presence of facial expression cues contributes something beyond an increase in cue number and eliciting cue-based reasoning.

Table 9. Logistic Regression Predicting Accuracy on Prosocial and Antisocial Tease Items from Cue Number, Use of Cue-Based Reasoning, and Presence of a Facial Expression Cue in Study 1

Predictor	β	Wald χ^2	*p*-value	Odds Ratio
Cue Number	0.809	5.802	0.016	2.246
Cue-Based Reasoning	3.197	49.002	< 0.001	24.461
Facial Expression Cue	0.938	5.271	0.022	2.556
Intercept	-1.645			

Looking at the cues mentioned in cue-based reasoning may also illuminate the effects of difference cues on accuracy. Among explanations that used cue-based reasoning and in terms of correlations, only mentioning relationship cues was significantly associated with accuracy, $r = -0.129$, $p = 0.045$. However, mentioning relationship cues was not found to be a significant predictor of accuracy in prosocial and antisocial tease items in a logistic regression model, $\beta = -1.673$, SE = 1.091, Wald $\chi^2(1) = 2.350$, OR = 0.188, $p = 0.125$, 95% CI = [0.022, 1.594]. To explain the perhaps

unexpected finding that relationship cue mentions were negatively correlated with accuracy in prosocial and antisocial tease items and its non-significance as a predictor, relationship cues were mentioned more than the other two cue categories in evaluation explanations that used cue-based reasoning: 17.8% more than facial expression cues, $t(184) = 2.966$, $p = 0.003$, and 24.3% more than gesture cues, $t(184) = 3.935$, $p < 0.001$. Relationship cues were also mentioned 26% of the time when a relationship cue was not even present. If it is used more and even when it may not be immediately relevant, you might see a decrease in accuracy. However, it is clearly contributing in some way. For example, of the 10 evaluation explanations for ambiguous (0-cue) tease items that used cue-based reasoning, all of them mentioned relationship cues[3].

Prosocial Teases. Only facial expression cues were significantly correlated with accuracy on prosocial tease items, $r = 0.491$, $p < 0.001$. Adding the presence of facial expression cues as a predictor to a logistic regression model with cue number predicting accuracy in prosocial tease items significantly improves predictive ability, compared to a model with just cue number and the intercept, $\chi^2(1, N = 140) = 24.721$, $p < 0.001$. This addition changes cue number into a nonsignificant predictor of accuracy in prosocial tease items, $\beta = 0.568$, SE = 0.503, Wald $\chi^2(1) = 1.274$, OR = 1.764, $p = 0.259$, 95% CI = [0.658, 4.729]. This may be due to the number of cues and the presence of a specific cue being inherently related. After all, the presence of a specific cue automatically means the number of cues is at least one. However, when this same logistic regression model was

[3] One of these evaluation explanations for ambiguous (0-cue) tease items that used cue-based reasoning also mentioned facial expression cues.

used to predict accuracy on prosocial and antisocial items, cue number remained significant when the presence of a facial expression cue was added as a predictor. This may suggest that facial expression cues may have a distinct effect on prosocial tease comprehension specifically. As such, cue number will remain in the logistic regression model for comparisons to the logistic regression model for accuracy in prosocial tease items and to the logistic regression model for accuracy in prosocial and antisocial tease items.

To further clarify the role of facial expression cues in predicting accuracy in prosocial tease items, the presence of facial expression cues was added as a predictor to a logistic regression model with cue number and the use of cue-based reasoning predicting accuracy in prosocial tease items, and this significantly improves predictive ability, compared to a model with just cue number, cue-based reasoning, and the intercept, $\chi^2(1, N = 140) = 17.781$, $p < 0.001$. Table 10 shows the logistic regression coefficient, Wald test and p-values, and odds ratio for cue number, cue-based reasoning, and presence of a facial expression cue. Clearly, the presence of facial expression cues contributes something beyond an increase in cue number and eliciting cue-based reasoning (see Figure 3 for accuracy by cue number and specific cue categories present).

Table 10. Logistic Regression Predicting Accuracy on Prosocial Tease Items from Cue Number, Use of Cue-Based Reasoning, and Presence of a Facial Expression Cue in Study 1

Predictor	β	Wald χ^2	*p*-value	Odds Ratio
Cue Number	0.594	0.710	0.400	1.812
Cue-Based Reasoning	4.089	22.979	< 0.001	59.681
Facial Expression Cue	3.288	12.237	< 0.001	26.780
Intercept	-2.146			

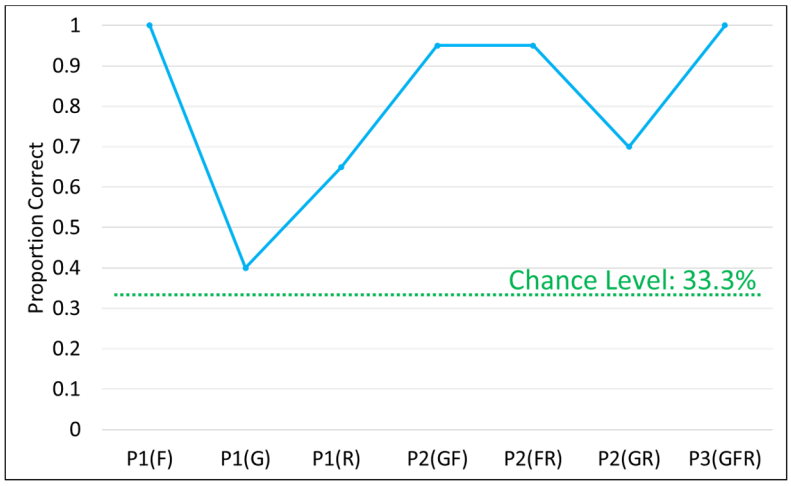

Figure 3. Accuracy of Prosocial Tease Items by Cue Number and Cue Categories Present in Study 1

Accuracy of prosocial (P) tease items is shown by number of cues (1, 2, or 3) and cue category (F = Facial Expression, G = Gesture, and R = Relationship Information). The order within the 1- and 2-cue items is arbitrary.

As facial expression cues were correlated with accuracy on prosocial tease items, prosocial tease items with facial expression cues will be compared in paired t-tests to items without facial expression cues (i.e., when facial expression cues are in addition to

other cues or when items have the same number of cues, but one has a facial expression cue and the other has a different cue). The results of such analyses can be found in Table 11. The apparent accuracy improvements of a prosocial tease item with a facial expression cue are clear, especially in 1-cue teases (the 1-cue tease item with a facial expression cue had better accuracy than the other 1-cue tease items with the other cues, both ps < 0.05, and better accuracy than the 2-cue tease item that did not have a facial expression cue, $t(19) = 2.854$, $p = 0.010$), but this will need to be revisited in a larger sample to understand whether and why facial expression cues might provide interpretive benefits for prosocial tease comprehension.

Table 11. Comparisons of Prosocial Tease Items With and Without the Facial Expression Cue in Study 1

Significant increases **bolded**.

Comparison Type	Prosocial Tease Item with Facial Expression Cue	Prosocial Tease Item without Facial Expression Cue	Mean Difference in Accuracy	95% CI		*t* (df = 19)	*p*-value
				Lower	Upper		
Addition of Facial Expression Cue	3-cue (Facial Expression, Gesture, and Relationship Information Cues)	2-cue (Gesture and Relationship Information Cues)	**0.30**	0.08	0.52	2.85	0.010*
Facial Expression Cue Compared to Different Cues	2-cue (Facial Expression and Relationship Information Cues)	2-cue (Gesture and Relationship Information Cues)	0.25	-0.01	0.51	2.03	0.056
Facial Expression Cue Compared to Different Cues	2-cue (Facial Expression and Gesture Cues)	2-cue (Gesture and Relationship Information Cues)	0.25	-0.01	0.51	2.03	0.056
Addition of Facial Expression Cue	2-cue (Facial Expression and Gesture Cues)	1-cue (Gesture Cue)	**0.55**	0.31	0.79	4.82	<0.001*
Addition of Facial Expression Cue	2-cue (Facial Expression and Relationship Information Cues)	1-cue (Relationship Information Cue)	**0.30**	0.08	0.52	2.85	0.010*
Facial Expression Cue Compared to Different Cues	1-cue (Facial Expression Cue)	2-cue (Gesture and Relationship Information Cues)	**0.30**	0.08	0.52	2.85	0.010*
Facial Expression Cue Compared to Different Cues	1-cue (Facial Expression Cue)	1-cue (Relationship Information Cue)	**0.35**	0.12	0.58	3.20	0.005*

Continued

Table 11 Continued

Comparison Type	Prosocial Tease Item with Facial Expression Cue	Prosocial Tease Item without Facial Expression Cue	Mean Difference in Accuracy	95% CI		*t* (df = 19)	*p*-value
				Lower	Upper		
Facial Expression Cue Compared to Different Cues	1-cue (Facial Expression Cue)	1-cue (Gesture Cue)	**0.60**	0.36	0.84	5.34	<0.001*

*$p < 0.05$

Among explanations that used cue-based reasoning and in terms of correlations, no specific cue category mentions were significantly correlated with accuracy in prosocial tease items (all *p*s>0.05). However, there are significant differences in gesture cue mentions as the number of cues changes in prosocial tease items, $F(2,87) = 3.366$, $p = 0.039$. Follow-up t-tests suggest that the significant difference is specifically due to an increase in gesture cue mentions from 1-cue prosocial tease items to 2-cue prosocial tease items, $t(73) = 2.621$, $p = 0.011$.

Antisocial Teases. Only gesture cues were significantly correlated with accuracy on antisocial tease items, $r = 0.217$, $p = 0.010$. Adding the presence of gesture cues as a predictor to a logistic regression model predicting accuracy in antisocial tease items significantly improves predictive ability, compared to a model with just the intercept, $\chi^2(1, N = 280) = 5.614$, $p = 0.018$. Due to the issues likely arising from the antisocial 2-cue tease items, cue number is not a significant predictor of accuracy in antisocial tease items and will not be included in the next analyses.

Adding the presence of gesture cues as a predictor to a logistic regression model with the use of cue-based reasoning predicting accuracy in antisocial tease items significantly improves predictive ability, compared to a model with just cue-based reasoning and the intercept, $\chi^2(1, N = 280) = 4.365$, $p = 0.037$. Table 12 shows the logistic regression coefficient, Wald test and p-values, and odds ratio for cue-based reasoning and presence of a gesture cue. The presence of gesture cues may contribute something beyond an increase in cue number and eliciting cue-based reasoning (see Figure 4 for accuracy by cue number and specific cue categories present).

Table 12. Logistic Regression Predicting Accuracy on Antisocial Tease Items from the Use of Cue-Based Reasoning and Presence of a Gesture Cue in Study 1

Predictor	β	Wald χ^2	*p*-value	Odds Ratio
Cue-Based Reasoning	2.396	19.443	< 0.001	10.976
Gesture Cue	0.998	4.222	0.040	2.713
Intercept	-0.139			

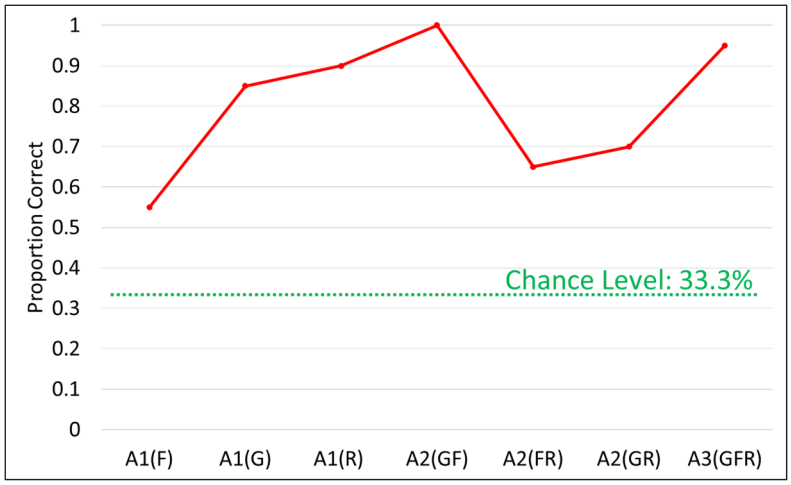

Figure 4. Accuracy of Antisocial Tease Items by Cue Number and Cue Categories Present in Study 1

Accuracy of antisocial (A) tease items is shown by number of cues (1, 2, or 3) and cue category (F = Facial Expression, G = Gesture, and R = Relationship Information). The order within the 1- and 2-cue items is arbitrary.

As gesture cues were correlated with and were a significant predictor of accuracy on antisocial tease items, antisocial tease items with gesture cues will be compared in paired t-tests to items with the same number of cues but a different cue. Unfortunately, the important items for these comparisons are one of the problematic antisocial 2-cue items and the antisocial 1-cue tease item with similar issues. As such, the comparison can only be focused on difference between the antisocial 1-cue tease item with a gesture cue and the antisocial 1-cue tease item with a relationship information cue, which is not significant, $t(19) = -0.438$, $p = 0.6662$. This will need to be revisited in future studies.

Among explanations that used cue-based reasoning and in terms of correlations, no specific cue category mentions were significantly associated with accuracy in antisocial tease items (all ps>0.05). Three interesting correlations suggest that, despite

accuracy in antisocial tease items only being significantly correlated with gesture cues, as the number of cues increases, participants mentioned relationship information cues (r = 0.214, p = 0.049) and facial expression cues (r = 0.310, p = 0.004) significantly more in their evaluation explanations. Of course, as the number of cues increase, they are likely to encounter more of the other cue categories, but gesture cue mentions actually had a nonsignificant and weaker negative correlation with cue number (r = -0.181, p = 0.097). Gesture cues are associated with accuracy, but aren't mentioned in reasoning as the number of cues increases. It is an interesting finding, but with a small sample size and non-significance, it would not be meaningful to interpret at this point. However, it does support future studies double-checking whether this finding was random chance or something systematic.

Ambiguous Teases. Due to ambiguous (0-cue) tease items only producing 10 evaluation explanations that used cue-based reasoning, no additional analyses will be performed for the ambiguous tease items. However, to restate information that should be investigated again in future studies, all of the ambiguous tease evaluation explanations that used cue-based reasoning mentioned relationship cues, and one also mentioned facial expression cues.

Discussion

The hypotheses of interest were all supported to some extent, which provides support for the validity of the measure and its theoretical basis beyond face validity. Participants, indeed, used cues present in a tease scenario to guide their reasoning about

PATC, and they used cue-based reasoning significantly more than other reasoning approaches in prosocial and antisocial tease items, both separately and together. Generally, cue-based reasoning differed by number of cues, as did the number of cues mentioned in their reasoning. These connections are primarily driven by increases in the number of cues in the items.

As additionally expected, increasing the number of cues likewise significantly increased accuracy in prosocial and antisocial tease items combined and prosocial tease items on their own. The more cues an item has the better the prosocial-antisocial tease comprehension and accuracy. When cue-based reasoning was added to models of cue number predicting accuracy, cue-based reasoning was also a significant predictor. This relationship between cue number and accuracy was weaker and, in some analyses, nonsignificant for antisocial tease items, although this is likely due to item-specific problems.

There were not many apparent differences between the tease types on general performance, but exploring the possible differential effects of specific cue categories enabled investigation of item and specific cue category differences by tease type. The performance on prosocial tease items was found to benefit from the presence of a facial expression cue, even beyond the effects of cue number and cue-based reasoning, whereas the performance on antisocial tease items was found to benefit from the presence of a gesture cue, even beyond the effects of cue-based reasoning. Thus, there are differences in measure performance by tease type, and different cue categories appear to have different effects on prosocial-antisocial tease comprehension and accuracy.

Additional analyses demonstrate strong, above-chance PATC performance overall with relatively large proportions of participants at ceiling for prosocial tease items and/or antisocial tease items. Most items also performed as intended, and those that did not had identifiable problems that can be addressed in future studies. Reliability unfortunately could not be explored at this stage, but the possible covariates of gender, age, and task duration could be. However, no major differences in PATC by gender, age, and task duration were found, but should be revisited in a larger sample.

Despite the support for all of the hypotheses, many of these findings and analyses should be reexamined in future studies that address the limitations and problems that had widespread effects on Study 1.

Limitations. The major limitations are mainly the small sample size and the source of the sample. The small sample size limited many of the analyses and results as well as the usefulness of any generalization from the results. However, as just a starting point for developing this PATC measure, the sample size provided enough information to test the measure and identify problems that needed to be addressed before subsequent studies with larger samples are conducted. The source of the sample was a research program pool at a university. Drawing a sample from such a participant pool leads to questions about the representativeness of the sample, as young adults attending college are not representative of all adults in the United States, nor in the world. Conclusions from this small and less representative sample need to be qualified and revisited in larger and more diverse and representative samples in future studies.

Along with these limitations, there are additional problems and possible changes that should be addressed in Study 2 and future studies of this PATC measure.

Looking Ahead to Study 2

Items that are too mean. As teases are playful provocations, the contents of a verbal tease may lean toward criticizing, which makes the verbal tease itself a form of cue on its own. However, to ensure that the effect of specific cues and cue numbers could be explored, the verbal or spoken tease in the tease scenario needs to be as ambiguous as possible. Of course, there will inherently be a bit of meanness as it is meant to be provoking or critical, but for the purpose of this measure, the goal was that the verbal teases would not be overly mean. Unfortunately, based on accuracy and evaluation explanations, two items did not hit that mark.

The first was the prosocial 1-cue (gesture) tease item (see Table 13). As can be seen in Table 13, the verbal tease itself is too long, which may have added too much of a negative connotation. This seems to be the case with how low the accuracy for the item was. Only 40% evaluated it as a "Friendly" tease. This accuracy is the second lowest of all the items and lower than that of one of the ambiguous teases. This tease has been changed to simplify it and make it more ambiguous and less mean for future use.

Another tease, an ambiguous (0-cue) tease (see Table 13) was also taken as too mean. The verbal tease was considered to be a severe personal attack by some participants, and two participants even said it would be considered a mean tease even if it was clear the teaser and teasee were good friends. Other participants suggested that a

close relationship would be enough to consider this tease as friendly, but the unequivocal meanness perceived by some participants warranted altering this item. The verbal tease was changed to be less personal and specifically directed at the teasee and more general and ambiguous.

Table 13. Verbal Portions of Tease Scenarios Interpreted as Meaner than Intended

Tease Item Identifier	Original Tease Scenario Text	Alteration of the Item for Future Use (e.g., Study 2)
Prosocial 1-Cue Tease with Cue for Gesture	Ethan came last in a footrace. Max high-fives Ethan as he finishes the race and says, **"My grandmother could give you a head start and still beat you in a race."**	Ethan came last in a footrace. Max high-fives Ethan as he finishes the race and says, **"My mom could beat you in a race."**
Ambiguous (0-Cue) Tease (1 of 2)	Luke got a bad grade on a test. Adrian sees and says, **"That looks like a new personal best for you."**	Luke got a bad grade on a test. Adrian sees and says, **"It can't get worse than that."**

Direction of relationship information. As discussed previously, three antisocial tease items were affected by the relationship information cue coming from the perspective of the teasee and not the teaser. The major alteration for future studies involves simply changing the direction of the relationship information to come from the teaser specifically, which can be seen in Table 4 with the original and altered versions of the items.

Tests of all possible items, cues, and scenarios. After the issues with specific items, it became clear that a larger test of these items and a more statistics-driven approach to identifying the best items for the final PATC measure were a necessary next step. To do so, all scenarios would need to be transmuted into all possible tease items: prosocial, antisocial, and ambiguous; 0 to 3 cues in all combinations of cue categories. More than 20 participants should complete a specific item, which would also likely address some, but not all, of the problems with a small sample size and items with zero variance. This would give a good pool of tested items to choose from for the final PATC measure, while also providing a combined, large-sample test for hypotheses that could not be addressed completely by Study 1.

More response options. An additional step that may further improve the utility of the measure is to extend the response options from three response options (Friendly, Not Sure, Mean) to seven response options (Very Friendly, Friendly, Slightly Friendly, Not Sure, Slightly Mean, Mean, Very Mean). This would provide more information on the valence of specific items (i.e., just how friendly or mean does a tease seem) and still allow for the options to be used for accuracy (e.g., a tease type evaluation of a prosocial tease item is accurate if any of the three Friendly response options are chosen).

Next in Study 2. Study 1 had several problems and began to illuminate how PATC may work and how a PATC measure needs to be constructed. These problems need to be addressed, and additional changes need to be considered and incorporated for Study 2 to improve upon the PATC measurement in Study 1 and to move forward in the development of a PATC measure.

Study 2

Goals and Objectives of Study 2

If Study 1 was the first step in developing a valid and informative PATC measure, Study 2 is the necessary second step in its development. First, it remedies the problems from and improves upon Study 1, and, second, it allows the reexamination of the hypotheses of interest and the findings of Study 1, all while Study 2 is enabling the identification of items for use in the final PATC measure**.**

Addressing Issues from Study 1. To carry out Study 2, some of the issues from Study 1 must first be addressed. Some of those issues were easily fixed by editing the basic text of an item such as addressing the items that were interpreted as too mean and those wherein the direction of relationship information led to confusion. These fixes can be seen in Table 13 as well as comparing Appendices A and C. Another simple change that was performed was expanding the number of response options from three response options (Friendly, Not Sure, Mean) to seven response options (Very Friendly, Friendly, Slightly Friendly, Not Sure, Slightly Mean, Mean, Very Mean). This will provide those who use the measure more information on tease valence in terms of friendliness/meanness, while still providing information on accuracy by sectioning the response options into friendly options, mean options, and the ambiguous option. The final issue to address is identifying items for the measure using a more statistically driven approach. Specifically, all 16 tease scenarios were used to create all possible combinations of tease type, cue number, and cue category. All resulting 240 items were organized into 16 measure versions with the ambiguous tease items being used in two

measure versions. This approach will allow for the selection of the best possible items for the final measure, while addressing some of the problems encountered in Study 1, such as some of the issues of a small sample size as well as items with zero variance.

Specific Goals and Objectives of Study 2. Study 1 moved a PATC measure beyond face validity and tested the theory-driven measure through participants' performance and explanations. Social information processing theory (e.g., Dodge & Crick, 1990) was specifically supported as the theoretical basis for PATC by participants using cue-based reasoning and cues to guide their tease type evaluations and improve their accuracy in those tease type evaluations. These Study 1 findings support the first two hypotheses of interest:

Hypothesis 1: Participants use cues present in a tease scenario to guide their reasoning about prosocial-antisocial tease comprehension.

Hypothesis 2: The more cues an item has, the better the prosocial-antisocial tease comprehension and accuracy will be.

Study 2 will revisit these hypotheses by exploring items with different numbers of cues, as more cues would lead to more cue-based reasoning. The number of cues and their effects on accuracy will be tested using logistic regression and other statistical models to see if more cues translate to greater accuracy. This will essentially replicate many of the analyses completed in Study 1.

As explained with Study 1, previous research (Barnett et al., 2004; Barnett et al., 2013) suggests that the tease types may involve different cognitive strategies and

approaches and that different cue categories may contribute differentially to PATC, which is associated with the last two hypotheses of interest:

Hypothesis 3: There will be differences in item and general measure performance by tease type.

Hypothesis 4: Different cue categories will have different effects on prosocial-antisocial tease comprehension and accuracy.

The results of Study 1 supported both hypotheses, but specific item issues undermined this support. Study 2 should be better equipped to investigate differences by and within tease types and the roles of different cue categories.

Reevaluating these four hypotheses is one focus of Study 2, but the other primary focus is item selection. The first set of these should focus on general performance of the items. The purpose of Study 2 is to identify the best items for each tease scenario, tease type, cue category, and cue number combination for the final measure. Each item will be analyzed for how many of the participants identified the tease type correctly. Items that are underperforming (i.e., more incorrect than correct responses for prosocial or antisocial teases) and/or completely unrelated to relevant and comparable items will not be considered as options for the final measure. For the ambiguous teases, it is again expected that there will be more "Not Sure" responses to these items, but the important qualifier is that the responses should not be heavily skewed toward prosocial nor antisocial more than any other response. The most balanced ambiguous teases will be considered for the final measure. Item analyses will be performed on the items fitting this and other criteria to determine the strongest items for inclusion in the final measure.

Covariates, such as gender, age, and task duration (time spent on the measure), will also be considered. Past PATC measures and Study 1 have found no general effects of gender (e.g., Barnett et al., 2004; Jones et al., 2018) nor age (e.g., Barnett et al., 2013; Sonnentag et al., 2016) in PATC, but it is important to clarify this finding in a larger sample with improved items, especially when more recent research suggests an effect of age (Rothermich et al., 2021). Although Study 1 did not find a general effect of time on PATC, it is best to retest the possibility that those who spend more time considering the tease scenarios and the cues may have better PATC than others who spend less time on the measure in Study 2.

Overall, Study 2 will allow the expectations for the measure (i.e., the four hypotheses of interest) and the findings of Study 1 to be revisited, while enabling the investigation and selection of items for a more finalized measure of PATC.

Methods

Sample. For Study 2, a large sample was needed to provide each measure version and item with a sufficient sample. Study 1's sample of 20 participants for 16 scenarios/items was based on the sample approach used by Dodge and colleagues (1990) when they were developing a similar style of measure (50 participants for 58 vignettes), and that approach was maintained for Study 2. However, due to a couple items having zero variance in Study 1, the number of participants for a specific measure version was raised to a goal of about 30 or at least 25 participants per measure version.

The initial sample was 499 survey workers from Amazon Mechanical Turk with the goal of at least 25 participants per measure version. The participants (M_{age} = 40.97 years old; 51.05% women, 48.54% men, 0.42% another gender identity) received $0.53 or $0.70 for their participation in this study. The compensation was raised during data collection to increase participation and to better align with ethical wage standards based on task duration (i.e., maintain a payment close to the then-current federal minimum wage of $7.25/hour). Two participants were excluded from analyses for repeated responses across items (e.g., all measure items marked as "Slightly Mean") with incredibly short task durations (i.e., less than 2 minutes), and 17 participants were excluded for failing the attention check questions. Due to a survey design issue on two measure versions (Versions 5 & 16), 8 participants were not shown a few of the items, before the issue was recognized and fixed. Generally, this only resulted in not completing 1 or 2 items, but 2 of these participants were excluded for not being shown 3 or more items (1 from each affected measure version). The final sample for analyses was 478 participants.

Procedure and Measures. This study and its procedures were reviewed and approved by the Ohio State University Social and Behavioral Sciences Institutional Review Board.

The entire study will be completed online through the Qualtrics survey system. As with Study 1, the survey opens with the informed consent form that gives information about the study, their rights, and a place to choose to consent or not to consent to participate in the study. If they select the "No, I DO NOT consent to participate in the

study" option, they will be taken to the end-of-study debriefing and thanked. If they select the "Yes, I consent to participate in the study" option, they will begin the study.

Participants will first be asked to give their month and year of birth (to calculate age) and their gender, and then they start the primary measure.

As this study is designed to determine the best items (i.e., cue type and number of cue combinations) for each tease scenario, the measure each participant receives will have the same 16 tease scenarios, but the measure versions will differ with teases for each scenario demonstrating all possible combinations of cue type and number of cues (see Appendix C for item variations by scenario). A total of 16 sets of these combination items will comprise the measures that the participants will be randomly assigned to (see Appendix D for measure versions). This design decision and lack of a proper anchor test used across the measure versions limited the possible analyses that could be performed and specifically prevented the use of item response theory approaches at this stage.

The teases will also feature any changes that were found to be necessary after analysis of Study 1. The order of the tease scenarios will be randomized.

Each tease scenario will have one question. The first question asks the participant to identify what the tease type of the tease in the scenario was (prosocial ["Friendly"], antisocial ["Mean"], or ambiguous ["Not Sure"]) from their perspective, as in Study 1. However, the response options have increased: Very Friendly, Friendly, Slightly Friendly, Not Sure, Slightly Mean, Mean, and Very Mean. Some attention check items will also be included to safeguard against participants randomly responding.

Then, they will be debriefed on the full purpose of the study and thanked for their participation. This is expected to take 10 to 30 minutes at most, with 5 to 7 minutes or even less being typical in this sample source, but it depends on their pace.

Results

The analyses for Study 2 span across general analyses performed across measure versions and item analyses within specific measure versions. The general analyses thus provide more information and data to speak to the major hypotheses, but also must be considered with the knowledge that aggregation across different measure versions and items can lead to issues with the conclusions, especially as some items and versions may not perform as well as others during this measurement development process. As such, although these Study 2 general analyses provide more information and data to support, qualify, or fail to support findings from Study 1, their findings must also be revisited with the final measure. The item analyses likewise have topics that must be discussed and considered, and it would be best for the analyses and their findings to be revisited in future research, all to strengthen any conclusions and ensure the production of a measure that is useful and informative.

Demographics and Task Duration. The descriptive statistics for age and task duration can be found in Table 14. Information on gender was included in the sample description, and, due to only having two participants identify by a gender other than man or woman, they were excluded from analyses involving gender. Additionally, due to outliers in task duration that suggested segmented task completion, the task duration was

skewed. For analyses using task duration, the most extreme outlier (42.33 hours) was first removed and then any task duration 3 or more standard deviations from the mean (between -19.64 minutes and 35.55 minutes) was considered an outlier and withheld from analyses. The task durations of only 8 (1.7%) of 478 participants were excluded by this approach.

Table 14. Descriptive Statistics for Age and Task Duration in Study 2

Variable	Mean	Standard Deviation	Range
Age in Years	40.97	11.260	19 – 72
Task Duration in Minutes (excluding outliers)	6.62	5.26	1.15 – 34.47

When gender, age, and task duration were used in logistic regression models to predict accuracy for prosocial tease item accuracy, antisocial tease item accuracy, and ambiguous tease item accuracy, gender produced nonsignificant results (all $p > 0.05$) when with age and task duration. This again aligns with past research that found no gender difference for PATC in children and adolescents (e.g., Barnett et al., 2004; Jones et al., 2018). Although, on its own, gender significantly predicted accuracy in antisocial tease items, $\beta = 0.262$, SE = 0.093, Wald $\chi^2(1) = 7.982$, OR = 1.300, $p = 0.005$, 95% CI = [1.084, 1.559], the relationship between accuracy in antisocial tease items and gender is relatively weak. Gender only accounts for 0.4% of the variance in accuracy on antisocial

tease items (Nagelkerke $R^2 = 0.004$). By Cohen's (1988) benchmarks for R^2, this does not even reach the benchmark for a small effect, $R^2 = 0.02$. Like the gender variable, age and task duration were significant predictors of accuracy in different tease types, regardless of the inclusion of the other variables, but have similar issues.

Specifically, age was a significant predictor of antisocial tease item accuracy, $\beta = 0.020$, SE = 0.004, Wald $\chi^2(1) = 22.181$, OR = 1.021, $p < 0.001$, 95% CI = [1.012, 1.029]. As previously noted, no age differences were found for PATC in older adolescents (e.g., Barnett et al., 2013; Sonnentag et al., 2016), but Rothermich and her colleagues (2021) found tease comprehension differences between young (18-39), middle-aged (40-59), and older (60-79) adults. Specifically, they found that older adults tended to struggle more with tease comprehension than younger adults. Following the age groups used by Rothermich and colleagues (2021), there were no significant age group differences in accuracy for prosocial and ambiguous tease items (both *p*s>0.500), but there was a significant difference by age group in antisocial teases, $\chi^2(2, N = 3346) = 15.808$, $p < 0.001$. This is mainly driven by significantly increased accuracy on antisocial tease items from young adults to middle-aged adults, mean difference = 4.81%, $t(3127) = 3.511$, $p < 0.001$, and from young adults to older adults, mean difference = 6.81%, $t(2049) = 2.442$, $p = 0.015$. The increase between middle-aged adults and older adults was not significant, mean difference = 2.00%, $t(1510) = 0.78$, $p = 0.435$. This finding is in opposition with Rothermich and colleagues' finding that older adults had more trouble with tease comprehension (2021). However, as the results here are based on an aggregation across different measure versions and specific items and as Rothermich and colleagues were

focused on general comprehension and not necessarily tease type, this conflict will need to be clarified in future studies. Moreover, as with the gender variable, the relationship between accuracy in antisocial tease items and age is weak. On its own, age only accounts for 1.2% of the variance in accuracy on prosocial tease items (Nagelkerke $R^2 = 0.012$). Like with the gender variable, this does not even reach the benchmark for a small effect, $R^2 = 0.02$ (Cohen, 1988).

Similarly, task duration was significant predictor of accuracy in prosocial tease items, $\beta = -0.0004$, SE = 0.0001, Wald $\chi^2(1) = 8.356$, OR = 1.000, $p = 0.004$, 95% CI = [0.999, 1.000], and ambiguous tease items, $\beta = -0.001$, SE = 0.0004, Wald $\chi^2(1) = 4.925$, OR = 0.999, $p = 0.026$, 95% CI = [0.998, 1.000]. Yet, as with the gender and age variables and antisocial tease item accuracy, task duration's relationship to accuracy in prosocial and ambiguous tease items is not very strong. On its own, task duration only accounts for 0.4% of the variance in accuracy on prosocial tease items (Nagelkerke $R^2 = 0.004$) and for 1.2% of the variance in accuracy on ambiguous tease items (Nagelkerke $R^2 = 0.012$). Again, these values do not even reach the benchmark for a small effect, $R^2 = 0.02$ (Cohen, 1988).

In addition to these main effects, there were a handful of significant correlations between gender, age, and task duration and other variables (see Table 15), but they did not replicate similar correlations found in Study 1. Overall, there seem to be only weak effects of gender, age, and task duration in this study, which may be due to the aggregation of measure versions and specific items. Future studies are needed to better capture these possible relationships and their effects.

Table 15. Significant Correlations Between Gender, Age, and Task Duration and Other Variables in Study 2

Variable	Gender (1 = Man, 2 = Woman)	Age	Task Duration
Age	0.211**	*1*	0.050
Accuracy in Antisocial 1-Cue Tease Items	0.083	0.092*	0.016
Accuracy in Antisocial 2-Cue Tease Items	0.056	0.169**	0.007
Accuracy in Antisocial 3-Cue Tease Items	0.089	0.108*	0.035
Accuracy in Prosocial 2-Cue Tease Items	-0.015	-0.004	-0.112*
Accuracy in Ambiguous Tease Items	0.006	0.020	-0.097*

*$p < 0.05$ (two-tailed)
**$p < 0.01$ (two-tailed)

General Analyses

91.21% of participants demonstrated above chance performance on prosocial tease items (prosocial and antisocial tease items' chance level = 42.86%), and 93.10% of participants demonstrated above chance performance on antisocial tease items. Although the accuracy on ambiguous tease items is not as meaningful to this investigation, 24.90% of participants exhibited above chance performance (ambiguous tease items' chance level = 14.29%). The average accuracy was 79.47% for prosocial items and 83.05% for antisocial items. As accuracy on ambiguous items works differently from prosocial and

antisocial items, it is no surprise that the accuracy on ambiguous tease items was 14.5% or just slightly above the chance level of 14.29%.

Of the prosocial and antisocial tease items that did not have zero variance (only 6 items had zero variance in accuracy), all but 19 had accuracies that significantly differed from chance and were above chance (42.86%; all *p*s < 0.05). 17 of the 224 prosocial and antisocial tease items were not significantly different from chance (42.86%; all *p*s > 0.05), and, separately, 2 of the 224 prosocial and antisocial tease items were significantly different from chance (42.86%; both *p*s < 0.05) by being significantly below chance in terms of accuracy. Any items with zero variance or below-chance accuracy were considered poor-performing and not used when possible in item analyses to prevent well-performing items from being obscured. As expected, accuracy on most ambiguous tease items did not differ from chance (14.29%; all *p*s > 0.05), but four of the 16 ambiguous tease items did significantly differ from chance (14.29%; all *p*s < 0.05) with three having accuracies lower than chance and one having an accuracy higher than chance.

Additional general information on the measure and accuracy will be presented in relation to specific relevant hypotheses.

Hypothesis 1: Participants use cues present in a tease scenario to guide their reasoning about prosocial-antisocial tease comprehension.

Hypothesis 2: The more cues an item has, the better the prosocial-antisocial tease comprehension and accuracy will be.

Whether cues are used in PATC can also be clarified if different performance is found with different amounts of cues, and, as with Study 1, this pattern was found (see Figure 5 for accuracy by cue number). When ambiguous tease items are excluded because they have no specific cues, accuracy on prosocial and antisocial tease items are significantly positively correlated with number of cues in the item ($r = 0.179$, $p < 0.01$). Likewise, in a logistic regression model, cue number significantly predicts accuracy in prosocial and antisocial tease items, $\beta = 0.735$, SE = 0.051, Wald $\chi^2(1) = 206.151$, OR = 2.086, $p < 0.001$, 95% CI = [1.887, 2.306]. Table 16 shows the logistic regression coefficient, Wald test and p-values, and odds ratio for cue number.

This model was able to correctly classify 100% of those with accurate performance on prosocial and antisocial tease items and 0% of those with inaccurate performance on prosocial and antisocial tease items, producing an overall success rate of 81.3%. On its own, cue number explained 5.5% of the variance in accuracy on prosocial and antisocial tease items (Nagelkerke $R^2 = 0.055$), and the model with cue number was significantly more predictive over a model with only the intercept, $\chi^2(1, N = 6692) = 230.159$, $p < 0.001$.

Overall, the presence of more cues increases accuracy. Cue number explains a meaningful portion of the variance in accuracy on prosocial and antisocial tease items and produces a significantly predictive model of accuracy on prosocial and antisocial tease items and thus PATC.

Table 16. Logistic Regression Predicting Accuracy on Prosocial and Antisocial Tease Items from Cue Number in Study 2

Predictor	β	Wald χ^2	*p*-value	Odds Ratio
Cue Number	0.735	206.151	< 0.001	2.086
Intercept	0.287			

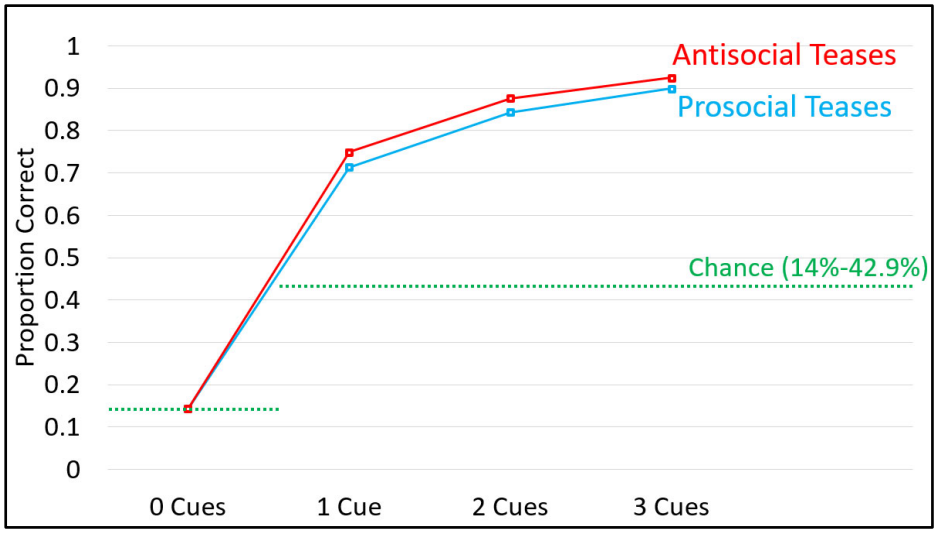

Figure 5. Accuracy for Prosocial and Antisocial Tease Items by Cue Number in Study 2

The shared point on the graph for 0 cues refers to accuracy on ambiguous tease items and is included on the lines for both prosocial and antisocial tease items for comparison.

Prosocial Teases. Accuracy on prosocial tease items are significantly positively correlated with the number of cues in the item ($r = 0.178$, $p < 0.001$). Likewise, in a logistic regression model, cue number significantly predicts accuracy in prosocial tease items, $\beta = 0.699$, SE = 0.069, Wald $\chi^2(1) = 102.405$, OR = 2.012, $p < 0.001$, 95% CI =

[1.757, 2.303]. Table 17 shows the logistic regression coefficient, Wald test and p-values, and odds ratio for cue number.

This model was able to correctly classify 100% of those with accurate performance on prosocial tease items and 0% of those with inaccurate performance on prosocial tease items, producing an overall success rate of 79.5%. On its own, cue number explained 5.2% of the variance in accuracy on prosocial tease items (Nagelkerke $R^2 = 0.052$), and the model with cue number was significantly more predictive over a model with only the intercept, $\chi^2(1, N = 3346) = 113.127$, $p < 0.001$.

Overall, the presence of more cues increases accuracy. Cue number explains a meaningful portion of the variance in accuracy on prosocial tease items and produce a significantly predictive model of accuracy on prosocial tease items and thus prosocial tease comprehension.

Table 17. Logistic Regression Predicting Accuracy on Prosocial Tease Items from Cue Number in Study 2

Predictor	β	Wald χ^2	*p*-value	Odds Ratio
Cue Number	0.699	102.405	< 0.001	2.012
Intercept	0.224			

Consistent with the regression models, prosocial tease accuracy increases with more cues within the overall model. Participants were more accurate on prosocial 2-cue

tease items than on prosocial 1-cue tease items, $t(2867) = 8.64$, $p < 0.001$. Participants were more accurate on the prosocial 3-cue tease item than on prosocial 2-cue tease items, $t(1911) = 2.91$, $p = 0.004$. Participants were more accurate on the prosocial 3-cue tease item than on prosocial 1-cue tease items, $t(1911) = 8.37$, $p < 0.001$.

Next, a persons-as-effect-sizes approach (Grice et al., 2020) is used to further clarify whether accuracy increases with more cues by identifying how many participants followed a general increasing trend in accuracy by number of cues within each tease type. For prosocial teases, 30.54% of all participants were at ceiling with all prosocial tease items correct. 1.05% were at floor with all prosocial tease items incorrect. Of those not at ceiling or floor, 76.76% (52.51% of the total sample) saw their accuracy increase from prosocial 1-cue tease items to the prosocial 3-cue tease item, whereas 11.93% had their accuracy decrease from prosocial 1-cue tease items to the prosocial 3-cue tease item and 11.93% had equal accuracy on prosocial 1-cue tease items and the prosocial 3-cue tease item.

When those who generally increased their accuracy from the least cues to the most cues are explored, 10.76% increased their accuracy with each additional cue, 19.92% had the same accuracy on 1-cue and 2-cue tease items and then increased with the 3-cue tease item, and 63.35% increased their accuracy from the 1-cue tease items to the 2-cue tease items and had the same accuracy on 2-cue and 3-cue tease items. The remaining 5.98% decreased their accuracy from the 1-cue tease items to the 2-cue tease items and then increased their accuracy with the 3-cue tease item (greater than their accuracy on the 1-cue tease items).

Antisocial Teases. Accuracy on antisocial tease items are significantly positively correlated with the number of cues in the item ($r = 0.181$, $p < 0.001$). Cue number also significantly predicted accuracy in antisocial tease items in a logistic regression model, $\beta = 0.783$, SE = 0.077, Wald $\chi^2(1) = 104.417$, OR = 2.187, $p < 0.001$, 95% CI = [1.882, 2.542]. Table 18 shows the logistic regression coefficient, Wald test and p-values, and odds ratio for cue number and cue-based reasoning.

This model was able to correctly classify 100% of those with accurate performance on antisocial tease items and 0% of those with inaccurate performance on antisocial tease items, producing an overall success rate of 83.1%. On its own, cue number explained 5.8% of the variance in accuracy on antisocial tease items (Nagelkerke $R^2 = 0.058$), and the model with cue number was not significantly more predictive over a model with only the intercept, $\chi^2(1, N = 3346) = 118.175$, $p < 0.001$.

Overall, more cues increase accuracy, unlike in Study 1 wherein there were specific issues with some of the antisocial tease items. Cue number explains a meaningful portion of the variance in accuracy on antisocial tease items and produces a significantly predictive model of accuracy on antisocial tease items and thus antisocial tease comprehension. These results do not support the possibility from Study 1 that the number of cues matters less in antisocial teases, despite cue-based reasoning still being predictive of accuracy on antisocial tease items. The results from Study 1 likely do reflect specific item issues.

Table 18. Logistic Regression Predicting Accuracy on Antisocial Tease Items from Cue Number in Study 2

Predictor	β	Wald χ^2	*p*-value	Odds Ratio
Cue Number	0.783	104.417	< 0.001	2.187
Intercept	0.343			

Consistent with the regression models, antisocial tease accuracy increases with more cues as in the overall model. Unlike in Study 1, participants were more accurate on antisocial 2-cue tease items than on antisocial 1-cue tease items, $t(2867) = 8.76$, $p < 0.001$. As in Study 1, the other t-tests support a general increasing trend in accuracy with more cues. Participants were more accurate on the antisocial 3-cue tease item than on antisocial 2-cue tease items, $t(1911) = 3.00$, $p = 0.003$. Participants were more accurate on the antisocial 3-cue tease item than on antisocial 1-cue tease items, $t(1911) = 8.36$, $p < 0.001$.

Using a person-as-effect-sizes approach and as with the prosocial tease items, 37.24% of participants were at ceiling for accuracy on antisocial tease items, answering all items correctly. 0.42% were at floor with all antisocial tease items incorrect. After excluding those at ceiling or floor, 76.17% saw their accuracy increase from antisocial 1-cue tease items to the antisocial 3-cue tease item, whereas 9.06% had their accuracy decrease from antisocial 1-cue tease items to the antisocial 3-cue tease item and 15.44% had equal accuracy on antisocial 1-cue tease items and the antisocial 3-cue tease item.

When those who generally increased their accuracy from the least cues to the most cues are explored, 10.57% increased their accuracy with each additional cue, 19.82% had the same accuracy on 1-cue and 2-cue tease items and then increased with the 3-cue tease item, and 66.52% increased their accuracy from the 1-cue tease items to the 2-cue tease items and had the same accuracy on 2-cue and 3-cue tease items. The remaining 3.08% decreased their accuracy from the 1-cue tease items to the 2-cue tease items and then increased their accuracy with the 3-cue tease item (greater than their accuracy on the 1-cue tease items).

Ambiguous Teases. Again, as stated, accuracy on ambiguous items is less meaningful as to be accurate necessitates choosing the middle or "Not Sure" option.

Hypothesis 3: There will be differences in item and general measure performance by tease type.

On average, participants had significantly different accuracy on prosocial items (M = 5.56 out of 7 total items, SD = 1.45) and antisocial items (M = 5.81 out of 7 total items, SD = 1.36), $t(477) = -2.866$, $p = 0.004$, 95% CI = [-0.423, -0.079].

As in Study 1, the accuracy on prosocial versus antisocial items was looked at on an individual level. 24.48% of participants were equally accurate on prosocial and antisocial items, getting the same exact number of items correct for each tease type. Including that group, 63.18% of participants were almost equally accurate on prosocial and antisocial items, only having at most a difference of 1 item between the number of correct items for each tease type. Despite the accuracies across prosocial and antisocial

tease items being significantly different, the participants largely performed similarly on prosocial and antisocial items in terms of accuracy.

When you compare the proportions of participants doing equally well on both tease types or doing better on one type over the other, 24.48% do equally well on both types, 32.64% do better on prosocial teases than on antisocial teases (14.23% did better on prosocial items by at least 2 points), and 42.89% do better on antisocial teases than on prosocial teases (22.59% did better on antisocial items by at least 2 points).

This hypothesis will be further explored in conjunction with Hypothesis 4.

Hypothesis 4: Different cue categories will have different effects on prosocial-antisocial tease comprehension and accuracy.

To investigate whether cues may produce different effects on PATC, correlations between the presence of different cue categories and accuracy on prosocial and antisocial items were examined. All cue categories were significantly correlated with accuracy on prosocial and antisocial tease items: facial expression cues, $r = 0.168$, $p < 0.001$; gesture cues, $r = -0.042$, $p = 0.001$; relationship information cues, $r = 0.128$, $p < 0.001$.

Adding the presence of the different cue categories as a predictor to a logistic regression model with cue number predicting accuracy in prosocial and antisocial tease items significantly improves predictive ability, compared to a model with just cue number and the intercept: facial expression cues, $\chi^2(1, N = 6692) = 56.769$, $p < 0.001$; gesture cues, $\chi^2(1, N = 6692) = 132.583$, $p < 0.001$; relationship information cues, $\chi^2(1, N$

= 6692) = 15.580, $p < 0.001$. Table 19 shows the logistic regression coefficients, Wald tests and p-values, and odds ratios for cue number and presence of different cues.

When all cue categories are included as predictors with cue number in a logistic regression model predicting accuracy in prosocial and antisocial tease items, one of the cue categories becomes nonsignificant and is not added into the model. This is due to cue categories like gesture cues having a larger effect. Table 20 shows this with the logistic regression coefficients, Wald tests and p-values, and odds ratios for cue number and presence of different cues.

Table 19. Logistic Regressions Predicting Accuracy on Prosocial and Antisocial Tease Items from Cue Number, and Presence of Different Cues. Part A: Facial Expression Cue; Part B: Gesture Cue; Part C: Relationship Information Cue.

Predictor	β	Wald χ^2	*p*-value	Odds Ratio
Cue Number	0.558	96.597	< 0.001	1.747
Facial Expression Cue	0.532	55.922	< 0.001	1.702
Intercept	0.304			

A

Predictor	β	Wald χ^2	*p*-value	Odds Ratio
Cue Number	1.007	311.186	< 0.001	2.737
Gesture Cue	-0.818	128.621	< 0.001	0.441
Intercept	0.326			

B

Continued

Table 19 Continued

Predictor	β	Wald χ^2	*p*-value	Odds Ratio
Cue Number	0.643	129.844	< 0.001	1.902
Relationship Information Cue	0.277	15.515	< 0.001	1.319
Intercept	0.291			

C

Table 20. Logistic Regressions Predicting Accuracy on Prosocial and Antisocial Tease Items from Cue Number, and Presence of Different Cues. Part A: Facial Expression and Gesture Cues; Part B: Facial Expression and Relationship Information Cues.

Predictor	β	Wald χ^2	*p*-value	Odds Ratio
Cue Number	0.922	172.394	< 0.001	2.513
Gesture Cue	-0.733	78.664	< 0.001	0.480
Facial Expression Cue	0.171	4.353	0.037	1.186
Intercept	0.328			

A

Predictor	β	Wald χ^2	*p*-value	Odds Ratio
Cue Number	0.188	7.028	0.008	1.207
Facial Expression Cue	0.904	117.864	< 0.001	2.471
Relationship Information Cue	0.733	4.353	< 0.001	2.082
Intercept	0.328			

B

Prosocial Teases. Facial expression cues were significantly and positively correlated with accuracy on prosocial tease items, $r = 0.279$, $p < 0.001$, while gesture cues were significantly and negatively correlated with accuracy on prosocial tease items, $r = -0.059$, $p = 0.001$.

Adding the presence of facial expression and gesture cues as predictors to a logistic regression model with cue number predicting accuracy in prosocial tease items significantly improves predictive ability, compared to a model with just cue number and the intercept, $\chi^2(1, N = 3346) = 285.065$, $p < 0.001$. Unlike in Study 1, these additions do not affect the status of cue number as a significant predictor of accuracy in prosocial tease items, $\beta = 0.471$, SE = 0.097, Wald $\chi^2(1) = 23.376$, OR = 1.602, $p < 0.001$, 95% CI = [1.323, 1.939]. Table 21 shows the logistic regression coefficient, Wald test and p-values, and odds ratio for cue number, presence of a facial expression cue, and presence of a gesture cue. Clearly, the presence of facial expression and gesture cues contributes something beyond an increase in cue number.

Table 21. Logistic Regression Predicting Accuracy on Prosocial Tease Items from Cue Number and Presence of a Facial Expression Cue and a Gesture Cue in Study 2

Predictor	β	Wald χ^2	*p*-value	Odds Ratio
Cue Number	0.471	23.376	< 0.001	1.602
Facial Expression Cue	1.066	82.783	< 0.001	2.903
Gesture Cue	-0.395	11.966	0.001	0.674
Intercept	0.326			

As facial expression cues were correlated with accuracy on prosocial tease items, prosocial tease items with facial expression cues will be compared in paired t-tests to items without facial expression cues (i.e., when facial expression cues are in addition to other cues or when items have the same number of cues, but one has a facial expression cue and the other has a different cue). When the 1-cue (facial expression) tease item is compared to both other 1-cue tease items, the benefit of a facial expression cue is clear with 25.10% increase in accuracy, $t(1432) = 10.238$, $p < 0.001$. The results of other such comparison analyses can be found in Table 22. The accuracy improvements of a prosocial tease item with a facial expression cue are clear across item type comparisons, but this will need to be revisited in future research with the final measure.

Table 22. Comparisons of Prosocial Tease Items With and Without the Facial Expression Cue in Study 2

Significant increases **bolded**.

Comparison Type	Prosocial Tease Item with Facial Expression Cue	Prosocial Tease Item without Facial Expression Cue	Mean Difference in Accuracy	95% CI		*t* (df = 477)	*p*-value
				Lower	Upper		
Addition of Facial Expression Cue	3-cue (Facial Expression, Gesture, and Relationship Information Cues)	2-cue (Gesture and Relationship Information Cues)	**0.159**	0.111	0.207	6.503	<0.001*
Facial Expression Cue Compared to Different Cues	2-cue (Facial Expression and Relationship Information Cues)	2-cue (Gesture and Relationship Information Cues)	**0.172**	0.124	0.219	7.145	<0.001*

Continued

Table 22 Continued

Comparison Type	Prosocial Tease Item with Facial Expression Cue	Prosocial Tease Item without Facial Expression Cue	Mean Difference in Accuracy	95% CI		*t* (df = 477)	*p*-value
				Lower	Upper		
Facial Expression Cue Compared to Different Cues	2-cue (Facial Expression and Gesture Cues)	2-cue (Gesture and Relationship Information Cues)	**0.144**	0.095	0.193	5.789	<0.001*
Addition of Facial Expression Cue	2-cue (Facial Expression and Gesture Cues)	1-cue (Gesture Cue)	**0.305**	0.252	0.358	11.317	<0.001*
Addition of Facial Expression Cue	2-cue (Facial Expression and Relationship Information Cues)	1-cue (Relationship Information Cue)	**0.232**	0.183	0.281	9.256	<0.001*
Facial Expression Cue Compared to Different Cues	1-cue (Facial Expression Cue)	2-cue (Gesture and Relationship Information Cues)	**0.140**	0.091	0.189	5.592	<0.001*
Facial Expression Cue Compared to Different Cues	1-cue (Facial Expression Cue)	1-cue (Relationship Information Cue)	**0.201**	0.150	0.252	7.694	<0.001*
Facial Expression Cue Compared to Different Cues	1-cue (Facial Expression Cue)	1-cue (Gesture Cue)	**0.301**	0.248	0.354	11.112	<0.001*

*$p < 0.05$

As gesture cues were correlated with accuracy on prosocial tease items, prosocial tease items with gesture cues were compared in paired t-tests to items without gesture cues (i.e., when gesture cues are in addition to other cues or when items have the same

number of cues, but one has a gesture cue and the other has a different cue). Unlike facial expression cues, gesture cues contribute minimally to increasing accuracy. The gesture cue increased item accuracy only when it was an addition, and this only occurred with the move from 1-cue to 2-cue teases. Even more, only one of those two increases was actually significant with a 6.1% increase in accuracy, $t(477) = 2.240$, $p = 0.026$. This suggests that any positive association between gesture cues and accuracy on prosocial tease items may be best attributed to the presence of any additional cue, not any inherent aspect of gesture cues. However, this was to be expected with gesture cues having an overall negative association with accuracy on prosocial tease items. In the paired t-tests, most of the significant decreases in accuracy associated with the presence of gesture cues compared to their absence can be explained by the presence of a facial expression cue, which has its own strong positive association with accuracy in prosocial tease items. As a result, a tease item with a gesture cue is generally going appear to produce lower accuracy in comparison to a tease item with a facial expression cue. Even when excluding comparisons to prosocial tease items with a facial expression cues, the presence of a gesture cue significantly decreases accuracy in comparison to the presence of a relationship information cue, $t(477) = -3.398$, $p = 0.001$. The results of these comparison analyses can be found in Table 23. Overall, the impact of gesture cues on prosocial tease comprehension is mixed, and this may be due to prosocial tease items with a gesture performing poorly in comparison to other prosocial tease items. Essentially, perhaps, the item construction itself was an issue, rather than gesture cues in themselves, and thus, aggregation across items may have exacerbated a systemic issue in

the use of gesture cues in items. Whether it was the item construction or something inherent to gesture cues will need to be explored in future research, but overall, it appears the impact of gesture cues on prosocial tease comprehension is uncertain but likely weaker than the impact of facial expression cues (see Figure 6 for accuracy by cue number and specific cue categories present).

Table 23. Comparisons of Prosocial Tease Items With and Without the Gesture Cue in Study 2

Significant increases **bolded**. Significant decreases *italicized*.

Comparison Type	Prosocial Tease Item with Gesture Cue	Prosocial Tease Item without Gesture Cue	Mean Difference in Accuracy	95% CI		*t* (df = 477)	*p*-value
				Lower	Upper		
Addition of Gesture Cue	3-cue (Facial Expression, Gesture, and Relationship Information Cues)	2-cue (Facial Expression and Relationship Information Cues)	-0.013	-0.045	0.020	-0.762	0.447
Gesture Cue Compared to Different Cues	2-cue (Gesture and Relationship Information Cues)	2-cue (Facial Expression and Relationship Information Cues)	*-0.172*	-0.214	-0.129	-7.885	<0.001*
Gesture Cue Compared to Different Cues	2-cue (Facial Expression and Gesture Cues)	2-cue (Facial Expression and Relationship Information Cues)	-0.027	-0.058	0.004	-1.725	0.085
Addition of Gesture Cue	2-cue (Facial Expression and Gesture Cues)	1-cue (Facial Expression Cue)	0.004	-0.032	0.040	0.229	0.819
Addition of Gesture Cue	2-cue (Gesture and Relationship Information Cues)	1-cue (Relationship Information Cue)	**0.061**	0.007	0.114	2.240	0.026*
Gesture Cue Compared to Different Cues	1-cue (Gesture Cue)	2-cue (Facial Expression and Relationship Information Cues)	*-0.333*	-0.383	-0.283	-13.051	<0.001*
Gesture Cue Compared to Different Cues	1-cue (Gesture Cue)	1-cue (Facial Expression Cue)	*-0.301*	-0.352	-0.250	-11.568	<0.001*
Gesture Cue Compared to Different Cues	1-cue (Gesture Cue)	1-cue (Relationship Information Cue)	*-0.100*	-0.158	-0.042	-3.398	0.001*

*$p < 0.05$

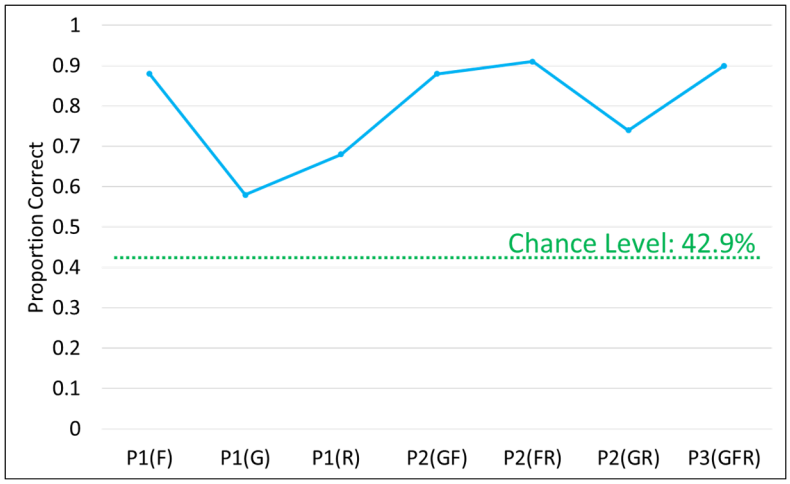

Figure 6. Accuracy of Prosocial Tease Items by Cue Number and Cue Categories Present in Study 2

Accuracy of prosocial (P) tease items is shown by number of cues (1, 2, or 3) and cue category (F = Facial Expression, G = Gesture, and R = Relationship Information). The order within the 1- and 2-cue items is arbitrary.

Antisocial Teases. Relationship information cues were significantly correlated with accuracy on antisocial tease items, $r = 0.232$, $p < 0.001$, as were facial expression cues, $r = 0.048$, $p = 0.005$. This is very much unlike Study 1, wherein gesture cues were significantly associated.

Adding the presence of relationship information and facial expression cues as predictors to a logistic regression model with cue number predicting accuracy in prosocial tease items significantly improves predictive ability, compared to a model with just cue number and the intercept, $\chi^2(1, N = 3346) = 96.467$, $p < 0.001$. Table 24 shows the logistic regression coefficient, Wald test and p-values, and odds ratio for cue number, presence of a relationship information cue, and presence of a facial expression cue.

Clearly, the presence of relationship information and facial expression cues contributes something beyond an increase in cue number.

Table 24. Logistic Regression Predicting Accuracy on Antisocial Tease Items from Cue Number and Presence of a Relationship Information Cue and a Facial Expression Cue in Study 2

Predictor	β	Wald χ^2	*p*-value	Odds Ratio
Cue Number	0.286	7.324	0.007	1.331
Relationship Information Cue	1.164	86.240	< 0.001	3.201
Facial Expression Cue	0.323	7.242	0.007	1.382
Intercept	0.413			

As relationship information and facial expression cues were correlated with and were a significant predictor of accuracy on antisocial tease items, antisocial tease items with relationship information and facial expression cues will be compared in paired t-tests to items with those cue categories.

As relationship information cues were correlated with accuracy on antisocial tease items, antisocial tease items with relationship information cues were compared in paired t-tests to items without relationship information cues (i.e., when relationship information cues are in addition to other cues or when items have the same number of cues, but one has a relationship information cue and the other has a different cue). Those comparisons revealed significant increases in accuracy due to the presence of a relationship

information cue in the antisocial tease items. When the 1-cue (relationship information) tease item is compared to both other 1-cue tease items, the tease item with the relationship information cue has a 18.41% accuracy advantage, $t(477) = 9.490$, $p < 0.001$. The results of other such comparison analyses can be found in Table 25. The benefit of an antisocial tease item with a relationship information cue is clear across comparisons, but this will need to be revisited in future research with the final measure.

Table 25. Comparisons of Antisocial Tease Items With and Without the Relationship Information Cue in Study 2

Significant increases **bolded**.

Comparison Type	Antisocial Tease Item with Relationship Information Cue	Antisocial Tease Item without Relationship Information Cue	Mean Difference in Accuracy	95% CI		*t* (df = 477)	*p*-value
				Lower	Upper		
Addition of Relationship Information Cue	3-cue (Facial Expression, Gesture, and Relationship Information Cues)	2-cue (Facial Expression and Gesture Cues)	**0.117**	0.078	0.156	5.915	<0.001*
Relationship Information Cue Compared to Different Cues	2-cue (Facial Expression and Relationship Information Cues)	2-cue (Facial Expression and Gesture Cues)	**0.107**	0.068	0.145	5.445	<0.001*
Relationship Information Cue Compared to Different Cues	2-cue (Gesture and Relationship Information Cues)	2-cue (Facial Expression and Gesture Cues)	**0.096**	0.054	0.139	4.430	<0.001*
Addition of Relationship Information Cue	2-cue (Facial Expression and Relationship Information Cues)	1-cue (Facial Expression Cue)	**0.184**	0.142	0.226	8.628	<0.001*
Addition of Relationship Information Cue	2-cue (Gesture and Relationship Information Cues)	1-cue (Gesture Cue)	**0.257**	0.209	0.306	10.409	<0.001*
Relationship Information Cue Compared to Different Cues	1-cue (Relationship Information Cue)	2-cue (Facial Expression and Gesture Cues)	**0.065**	0.022	0.108	2.966	0.003*
Relationship Information Cue Compared to Different Cues	1-cue (Relationship Information Cue)	1-cue (Facial Expression Cue)	**0.142**	0.098	0.187	6.298	<0.001*

Continued

Table 25 Continued

Comparison Type	Antisocial Tease Item with Relationship Information Cue	Antisocial Tease Item without Relationship Information Cue	Mean Difference in Accuracy	95% CI		*t* (df = 477)	*p*-value
				Lower	Upper		
Relationship Information Cue Compared to Different Cues	1-cue (Relationship Information Cue)	1-cue (Gesture Cue)	**0.226**	0.177	0.275	9.067	<0.001*

*$p < 0.05$

As facial expression cues were correlated with accuracy on antisocial tease items, antisocial tease items with facial expression cues will be compared in paired t-tests to items without facial expression cues (i.e., when facial expression cues are in addition to other cues or when items have the same number of cues, but one has a facial expression cue and the other has a different cue). The impact of facial expression cues in antisocial tease items seems similar to that of gesture cues in prosocial tease items, contributing minimally to increasing accuracy on their own and being overshadowed by a stronger association. The facial expression cue increased item accuracy when it was an addition, with a significant increase of 10.15% with the move from 1-cue to 2-cue teases, $t(477) = 6.371$, $p < 0.001$, and nonsignificant increase of 2.1% with the move from 2-cue to 3-cue teases, $t(477) = 1.337$, $p = 0.182$). All the significant decreases in accuracy associated with the presence of facial expression cues compared to their absence can be explained by the presence of a relationship information cue, which has its own strong positive association with accuracy in antisocial tease items. Unlike gesture cues in prosocial tease items which had an opposite association to that relationship cues in prosocial tease items,

relationship information cues and facial expression cues in antisocial tease items both have having a positive association with accuracy on antisocial tease items. As such, this appearance of lower antisocial tease item accuracy with facial expression cues when compared to that of items with relationship information cues is best explained by relationship information cues simply producing a stronger benefit for item accuracy in comparison. There is still a clear benefit for items with facial expression cues because, when excluding comparisons to antisocial tease items with a relationship information cue, the presence of a facial expression cue significantly increases accuracy in comparison to the presence of a gesture cue, $t(477) = 3.041$, $p = 0.002$. The results of other such comparison analyses can be found in Table 26. As aforementioned, this benefit of facial expression cues for antisocial tease comprehension will need to be followed up in future research with the final measure, but from this study, that benefit is clear although weaker than the accuracy advantage from a relationship information cue (see Figure 7 for accuracy by cue number and specific cue categories present).

Table 26. Comparisons of Antisocial Tease Items With and Without the Facial Expression Cue in Study 2

Significant increases **bolded**. Significant decreases *italicized*.

Comparison Type	Antisocial Tease Item with Facial Expression Cue	Antisocial Tease Item without Facial Expression Cue	Mean Difference in Accuracy	95% CI		*t* (df = 477)	*p*-value
				Lower	Upper		
Addition of Facial Expression Cue	3-cue (Facial Expression, Gesture, and Relationship Information Cues)	2-cue (Gesture and Relationship Information Cues)	0.021	-0.010	0.052	1.337	0.182
Facial Expression Cue Compared to Different Cues	2-cue (Facial Expression and Relationship Information Cues)	2-cue (Gesture and Relationship Information Cues)	0.010	-0.020	0.041	0.674	0.501
Facial Expression Cue Compared to Different Cues	2-cue (Facial Expression and Gesture Cues)	2-cue (Gesture and Relationship Information Cues)	*-0.096*	-0.139	-0.054	-4.430	<0.001*
Addition of Facial Expression Cue	2-cue (Facial Expression and Gesture Cues)	1-cue (Gesture Cue)	**0.161**	0.109	0.213	6.070	<0.001*
Addition of Facial Expression Cue	2-cue (Facial Expression and Relationship Information Cues)	1-cue (Relationship Information Cue)	**0.042**	0.009	0.075	2.475	0.014*
Facial Expression Cue Compared to Different Cues	1-cue (Facial Expression Cue)	2-cue (Gesture and Relationship Information Cues)	*-0.174*	-0.217	-0.130	-7.884	<0.001*
Facial Expression Cue Compared to Different Cues	1-cue (Facial Expression Cue)	1-cue (Relationship Information Cue)	*-0.142*	-0.187	-0.098	-6.298	<0.001*

Continued

Table 26 Continued

Comparison Type	Antisocial Tease Item with Facial Expression Cue	Antisocial Tease Item without Facial Expression Cue	Mean Difference in Accuracy	95% CI		*t* (df = 477)	*p*-value
				Lower	Upper		
Facial Expression Cue Compared to Different Cues	1-cue (Facial Expression Cue)	1-cue (Gesture Cue)	**0.084**	0.030	0.138	3.041	0.002*

*$p < 0.05$

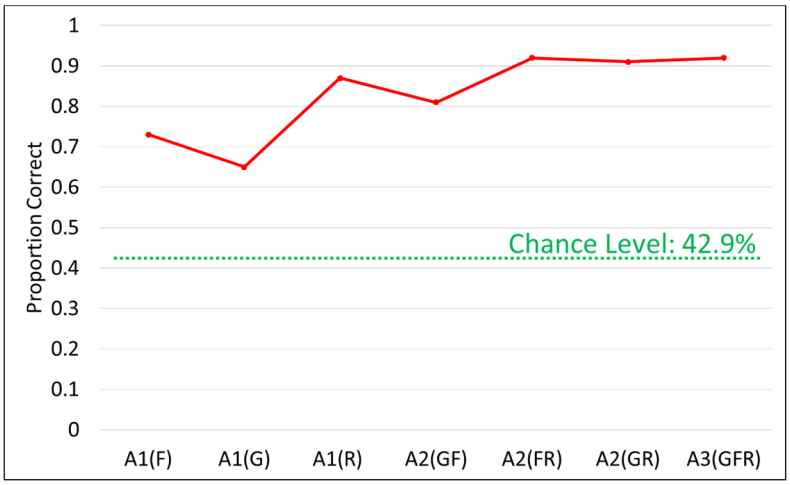

Figure 7. Accuracy of Antisocial Tease Items by Cue Number and Cue Categories Present in Study 2

Accuracy of antisocial (A) tease items is shown by number of cues (1, 2, or 3) and cue category (F = Facial Expression, G = Gesture, and R = Relationship Information). The order within the 1- and 2-cue items is arbitrary.

Ambiguous Teases. Due to ambiguous tease items having no specific assigned cues, no additional analyses will be performed for the ambiguous tease items in this aspect.

Prosocial and Antisocial Tease Item Analyses & Selection

Identifying the prosocial and antisocial tease items for the final measure relies on several analyses based in a classical test theory approach. However, as stated previously, the fundamental construction of prosocial and antisocial tease items with different cue category combinations limits the approaches that can be used. As cue categories can contribute in different ways, a classical test theory approach may not be able to best capture the relationship between items through inter-item and item-measure correlations. Items with different numbers and combinations of cue categories represented may not correlate as in another type of measure, although specific subsets of those correlations were still examined in this case. An item response theory approach would be able to deal with this issue, but, as explained previously, these measure versions and items were not tested in a manner that would allow for that approach to be taken.

To address these item analyses in the best way possible given the procedure used in Study 2, specific requirements and standards must be used to guide the item analyses and item selection process. There are some inherent expectations for the items of this PATC measure. Namely, the selected prosocial and antisocial tease items must a) have an accuracy above chance-level (42.857% for prosocial and antisocial tease items) in absolute terms, b) have an accuracy that is significantly different from chance-level accuracy, c) have a valence mean that is in the expected direction (positive/above zero for prosocial tease items and negative/below zero for antisocial tease items), d) have a valence mean that is significantly different from zero, e) relevant item-item correlations (i.e., matching cue categories and cue number) must be equivalent to or greater than a

small effect, and f) item-subscale (prosocial or antisocial subscale) must be equivalent to or greater than a large effect.

At least four out of the five main requirement standards must be met for consideration, and if a requirement is not fully met, the item must be qualitatively in line with the requirement and quantitatively close to meeting the requirement (i.e., meet a specific secondary requirement standard). As there are no formal standards available for this type of measure, arbitrary but meaningful standards were selected. For example, although item-item correlations are expected to be lower because of the measure construction, it can be expected that a well-performing item would still produce at least a small effect in terms of correlation with a relevant, well-performing item or a correlation coefficient of at least 0.10 (Cohen, 1988). As such, the main requirement standard for item-item correlations in these analyses is at least $r = 0.20$ with a secondary requirement standard of $r = 0.15$ and above considered if all other requirements are met. Similarly, it can be expected that a well-performing item would still produce at least a large effect in terms of correlation with the overall subscale or a correlation coefficient of at least 0.50 (Cohen, 1988). Thus, the main requirement standard for item-subscale correlations in these analyses is at least $r = 0.60$ with a secondary requirement standard of $r = 0.55$ and above considered if all other requirements are met. Additionally, accuracy on the items may be impacted by poor-performing items and the relatively small measure version sample sizes, so a p-value of 0.05 will be the main requirement standard, but those with a p-value of 0.10 and below will also be considered if all other requirements are met.

However, some requirements must have their main requirement standard met with no secondary standard considered. For example, the accuracy of an item must at minimum be greater than chance-level accuracy in absolute terms, even if the p-value is above 0.05 for the t-test. Likewise, the requirements related to item valence are specific and strict as they relate directly to tease type perception. The item valence must be in the expected direction: positive for prosocial tease items and negative for antisocial tease items. The item valence must also be significantly different from zero to be considered non-ambiguous.

The item analyses used to examine all items based on these requirement standards is described here:

a) Accuracy must be greater than chance level, 42.857% (see Tables 29 and 31).
b) T-tests with a null hypothesis of Mean (i.e., Accuracy) = 0.42857 were performed for each prosocial and antisocial tease item, wherein the null hypothesis should be rejected with a p-value of 0.05, but a p-value of 0.10 and below will be considered if all other requirements are met (see Tables 29 and 31).
c) Item valence must be in the expected direction: positive/above zero for prosocial tease items and negative/below zero for antisocial tease items (see Tables 29 and 31). This requirement must be fully met for consideration.
d) T-tests with a null hypothesis of Mean (i.e., Valence) = 0 were performed for each prosocial and antisocial tease item, wherein the null hypothesis should be rejected with a p-value of 0.05 (see Tables 29 and 31). This requirement must be fully met for consideration.

e) Correlations with relevant items must have a correlation coefficient of at least r = 0.20, but a correlation coefficient of r = 0.15 and above will be considered if all other requirements are met (see Tables 30 and 32). Relevant items are items with the same cue categories present and items with the same number of cues. For example, a 2-cue tease item with facial expression and relationship information cues should be compared to the other 2-cue tease items because they have the same number of cues, to the 3-cue tease item because it also has the facial expression and relationship information cues, to the 1-cue tease item with a facial expression cue because they share that cue, and to the 1-cue tease item with a relationship information cue because they share that cue. As the 2-cue tease item with facial expression and relationship information cues has no gesture cue, the 1-cue gesture item is not a relevant or meaningful comparison, so their correlation is not considered in this requirement.

f) Correlations with the related subscale (i.e., the correlation of a prosocial [antisocial] tease item with the prosocial [antisocial] subscale, an aggregation of the measure version's prosocial [antisocial] tease items) must have a correlation coefficient of at least r = 0.60, but a correlation coefficient of r = 0.55 and above will be considered if all other requirements are met (see Tables 30 and 32).

Applying these requirements led to most of the items being immediately removed from consideration for selection. Then, as all scenarios must be represented to ensure no scenarios repeat in the final measure, the item of a scenario that most strongly met the requirements was selected first, and then the item of a scenario that second most strongly

met the requirements was selected, and so on until all measure items and scenarios were selected.

Due to the process used here, the performance of the selected prosocial and antisocial tease items should be examined closely with all the applicable requirements above retested in future research. Any item that met most, if not all, requirements should also be considered in any further measure development and improvement, especially if an item response theory approach can be taken. After all, items chosen with a classical test theory approach to their analyses may not be as informative as intended or as suggested by the requirements and standards included here.

Ambiguous Tease Item Analyses & Selection

For this measure, ambiguous items need to be identified with different analyses as the goals for these items is inherently different. A valid ambiguous item needs to have a) a mean that is not significantly different from zero and b) minimal skew toward positive or negative evaluations. Additionally, as Study 2 was designed to allow ambiguous items to be measured in two different versions and samples, we also need the ambiguous items to work similarly in each sample. To clarify, the ambiguous items need to a) have a mean that is not significantly different from zero in both samples, separately and aggregated, b) have minimal skew toward positive or negative evaluations in both samples, and c) not be evaluated as significantly different when comparing the samples. All three of these requirements are vital because having only two out of three addressed leads to the selection of completely different items. For example, the ambiguous item for Scenario 1 has a mean that is not significantly different from zero in both samples, separately and

aggregated, and is not evaluated significantly different when comparing the samples, but it demonstrates a positive skew that would make the item uninformative and affect its interpretations and conclusions. The item analyses for these requirements are described next:

a) T-tests with a null hypothesis of Mean = 0 were performed for each ambiguous item, separately by sample or aggregation of the item's two samples, wherein the null hypothesis should not be rejected (see Table 33). Items with means significantly different from zero were not selected.

b) Skewness was measured with the ambiguous item's total sample (i.e., aggregation of the item's two samples). The skewness statistic was computed for each item (see Table 34). Items with higher skewness statistics (more than 0.50 or less than -0.50) were not selected. The selection restrictions for the skewness statistic here (between 0.50 and -0.50) are stricter than standard restrictions (between 2.00 and -2.00; Gravetter & Wallnau, 2014) because the skewness statistics for all ambiguous items were all between 0.90 and -0.90. Kurtosis and the Shapiro-Wilk test of normality were not informative because of the depression in the use of the middle answer option (i.e., "Not Sure," the "correct" answer for ambiguous tease items with a value of zero), which causes increased uniformness and inherent non-normality. The 95% confidence interval of the mean difference from zero computed in the aforementioned t-tests for the aggregated version of the item was included in Table 34 to provide a description of the skew. Items with mean differences from zero that had 95% confidence intervals that did not contain zero

were not selected. Additionally, a value for positive-negative skew in the 95% confidence interval was computed to further describe the skewness (see Table 28 for details). Items that did not have ranges with a positive-negative skew in the 95% confidence interval between 0.40 and -0.40 (a 70:30 ratio of the proportion of the range that was positive to the proportion of the range that was negative and vice versa) were not selected.

c) Independent samples t-tests with a null hypothesis of $Mean_{SampleA} = Mean_{SampleB}$ were performed for each ambiguous item, wherein the null hypothesis should not be rejected (see Table 27). Items that were evaluated as significantly different between samples were not selected.

These requirements led to the selection of the ambiguous items for Scenarios 8 and 15 for meeting all requirements (see Table 28). The ambiguous item for Scenario 13 also met all requirements, but its p-values in significance tests involving its separate samples or comparing those samples were much lower than those of Scenarios 8 and 15, suggesting that its evaluation was more sample dependent. To illustrate, for the Scenario 13 ambiguous item, Sample A produced a mean evaluation of -0.27, and Sample B produced a mean evaluation of 0.26. Each sample produced essentially equal but opposite evaluations. This can be most clearly observed when the positive-negative skew in the 95% confidence interval of the mean difference from zero (see Table 28 for details) is computed for each sample. When the samples were aggregated, the positive-negative skew was 0.00. When the samples are separate, the positive-negative skew was -0.49 in

Sample A and 0.56 in Sample B. With this information altogether, the Scenario 13 ambiguous item was not selected.

It should be noted that the selected ambiguous items from Scenarios 8 and 15 also demonstrate such tendencies to a far lesser extent (e.g., the positive-negative skew in the 95% confidence interval of the mean difference from zero for each sample for each ambiguous item is within -0.36 and 0.36). Thus, the performance of these selected ambiguous items should be examined closely in future research with all the applicable requirements above retested.

Table 27. Selection of Ambiguous Items by No Rejection of H_0: $Mean_{SampleA} = Mean_{SampleB}$ Requirement

Selected items are **bolded**.

Scenario of Ambiguous Item	Mean	Standard Deviation	Do not reject H_0: $Mean_{SampleA} = Mean_{SampleB}$		Requirements Met
			t statistic	*p*-value	
1	0.31	1.28	0.99	0.327 [R]	Yes
2	-0.28	1.21	-1.37	0.175 [R]	Yes
3	1.17	1.13	-0.29	0.775 [R]	Yes
4	-0.70	1.45	-0.85	0.397 [R]	Yes
5	-1.10	1.11	0.24	0.810 [R]	Yes
6	-1.18	1.26	-2.00	0.050	No
7	-0.02	1.49	2.26	0.028	No
8	**-0.02**	**1.57**	**0.91**	**0.365 [R]**	**Yes**
9	0.20	1.36	0.92	0.362 [R]	Yes
10	-0.41	1.26	-0.35	0.731 [R]	Yes
11	1.09	1.10	0.14	0.890 [R]	Yes
12	-0.58	1.44	-0.93	0.357 [R]	Yes
13	0.00	1.37	-1.52	0.135 [R]	Yes
14	-0.20	1.49	-1.58	0.121 [R]	Yes
15	**-0.05**	**1.45**	**0.62**	**0.539 [R]**	**Yes**
16	-0.59	1.23	0.67	0.504 [R]	Yes

[R] Requirement met: p-value > 0.05

Table 28. Total Requirements Met

Selected items are **bolded**.

Scenario of Ambiguous Item	Do not reject H_0: Mean = 0	Skewness	Do not reject H_0: $Mean_{SampleA} = Mean_{SampleB}$	All Requirements Met	Item Selection
1	Yes	No	Yes		
2	No	No	Yes		
3	No	No	Yes		
4	No	No	Yes		
5	No	No	Yes		
6	No	No	No		
7	Yes	Yes	No		
8	**Yes**	**Yes**	**Yes**	**Yes**	**Selected**
9	Yes	No	Yes		
10	No	No	Yes		
11	No	No	Yes		
12	No	No	Yes		
13	Yes	Yes	Yes	Yes	
14	Yes	No	Yes		
15	**Yes**	**Yes**	**Yes**	**Yes**	**Selected**
16	No	No	Yes		

[R] Requirement met: p-value > 0.05

Discussion

As in Study 1, the hypotheses of interest were all supported, which again provides support for the validity of the measure and its theoretical basis beyond face validity. In Study 2 and as postulated in Hypothesis 1, participants used cues present in a tease scenario to guide their reasoning about prosocial-antisocial tease comprehension. This replicates the findings of Study 1. Increasing the number of cues significantly increased accuracy in prosocial and antisocial tease items combined and prosocial tease items and antisocial tease items on their own. Unlike the antisocial tease items in Study 1, which had notable issues, participants across the measure versions in Study 2 demonstrated better accuracy with more cues on the antisocial tease items. Altogether, Hypothesis 2 is also supported: The more cues an item has, the better the prosocial-antisocial tease comprehension and accuracy were.

As hypothesized with Hypothesis 3, there were also differences in item and general measure performance by tease type. In terms of general performance, participants demonstrated strong, above-chance PATC performance with relatively large proportions of participants at ceiling for prosocial tease items (30.54%) and/or antisocial tease items (37.24%). Unlike in Study 1, participants did significantly better on antisocial tease items than on prosocial tease items. However, 63.18% of participants still performed similarly on prosocial and antisocial tease items with their performances differing by one item or less. Fewer participants had their performances on prosocial and antisocial tease items differ by two or more items with 22.59% with better performance on antisocial items and 14.23% with better performance on prosocial items.

When focusing on item performance and the effects of specific cue categories, there were distinct effects on PATC and accuracy by tease type, just as expected by Hypotheses 3 and 4. In addition to the effect of cue number, the performance on prosocial tease items benefited from the presence of a facial expression cue and harmed by the presence of a gesture cue, whereas the performance on antisocial tease items benefited from the presence of a relationship information cue and the presence of a facial expression cue. Study 1 also found the benefit of facial expression cues with the prosocial tease items, so this benefit has been replicated in Study 2. This benefit of facial expression cues was also stronger than the harm of the presence of gesture cues. With antisocial tease items, relationship information cues produced a greater benefit in terms of performance than facial expression cues. Apart from the replicated benefit of facial expression cues with the prosocial tease items, these other tease-type-specific benefits are new to Study 2 and should be revisited in the future.

The possible covariates of gender, age, and task duration were explored once again, and, despite the larger sample and wider representation of ages, gender, age, and task duration had significant but weak effects on accuracy. Gender and age were significant predictors of antisocial tease comprehension, wherein women and older adults had greater accuracy on antisocial tease items. This may in part be connected to a significant positive correlation suggesting women in the overall sample tended to be older. Task duration was a significant predictor of prosocial tease comprehension and accuracy on ambiguous tease items, such that greater task duration was associated with worse accuracy on prosocial and ambiguous tease items. However, as stated, these were

all very small effects, each explaining 1.2% or less of the variance in accuracy and all not meeting the standards for a small effect (Cohen, 1988). Thus, replicating Study 1, no major differences in PATC by gender, age, and task duration were found, but these covariates should remain part of future PATC analyses, until it is clear whether they play any role in PATC. Studies 1 and 2 suggest that role may be negligible at best.

The final and most vital analyses of Study 2 were focused on item analysis and item selection. These analyses resulted in 16 items being selected for the current "final" PATC measure. These 16 items were selected using requirements specific to the needs of the item's construction and of the overall measure. An important note on these selected items is that they were based on standards that, although meaningful and carefully chosen, could be considered overly arbitrary, and, as such, this PATC measure may benefit from future improvement and improvement in the development process. Namely, being able to fully utilize the methods of item response theory would greatly benefit this measure and its development, but the procedure and design of Study 2 made that option inaccessible at this time. So, while these selected items, just like any measure, need further testing, especially in terms of validity and reliability, it would be best to use this measure as a starting point for further and better PATC measure development. Any research completed with this measure may benefit from later replication with an improved version.

As with PATC measure development, the findings of Study 2 should be reassessed in future studies with the current "final" PATC measure of the selected items altogether and potentially with any other items that met the requirements and standards in Study 2.

Chapter 3. Discussion

In both Studies 1 and 2, the hypotheses of interest were all supported, marking that a) this PATC measure has surpassed most of the past PATC measures in terms of validation efforts and that b) the theoretical basis aligns with PATC and its measurement. Most general findings in Study 1 were replicated in Study 2, even after changes that improved the PATC measure under development. Participants across both studies were also quite successful in their PATC with strong performance and accuracy on both prosocial tease items (80.71% in Study 1; 79.43% in Study 2) and antisocial tease items (80% in Study 1; 83% in Study 2), even despite any item-specific issues. Large proportions of participants were also at ceiling for prosocial tease items (35% in Study 1; 30.54% in Study 2) and/or for antisocial tease items (20% in Study 1; 37.24% in Study 2). Overall, the adult participants had great success with PATC.

As for the hypotheses, Hypothesis 1 was supported. Participants did use cues present in a tease scenario to guide their reasoning about PATC, using cue-based reasoning significantly more than other reasoning for prosocial and antisocial tease items in Study 1. As the number of cues present in a tease item increased, participants used cue-based reasoning more and mentioned more cues in their evaluation explanations. In addition, PATC accuracy increased with more cues, which further supports that participants used cues present in a tease scenario to guide their reasoning about PATC.

Both studies thus also support Hypothesis 2 that more cues are associated with increased accuracy on PATC. In logistic regression models, cue number was a significant predictor of accuracy for prosocial tease items in Study 1. Although problems with some antisocial tease items in Study 1 prevented cue number from being a significant predictor of accuracy on Study 1 antisocial tease items, 68.75% of the participants not at ceiling still had higher accuracy on the antisocial 3-cue tease item than the antisocial 1-cue tease items. Following the improvements on the items for Study 2, cue number was a significant predictor of accuracy for both prosocial and antisocial tease items in Study 2.

Hypothesis 3, postulating that there will be differences in item and general performance by tease type, was largely supported, but Study 1 did not find a different in general performance by tease type. As with Hypothesis 2, this may reflect item-specific issues with the antisocial tease items in Study 1, and once those issues were addressed, participants did significantly better on antisocial tease items than on prosocial tease items in Study 2. However, the majority of participants (75% in Study 1; 63.18% in Study 2) still performed similarly on prosocial and antisocial tease items with their performances differing by one item or less.

The part of Hypothesis 3 centered on item performance differences by tease type was also supported along with Hypothesis 4, which focuses on the effects of different cue categories on PATC and thus also on the differences in item performance. Different cue categories did have different effects on PATC and accuracy in both studies. Notably, facial expression cues seem to provide an accuracy advantage on prosocial tease items as found in Study 1 and replicated by Study 2. Study 1 additionally found antisocial tease

items received an accuracy advantage from gesture cues, but that was not replicated in Study 2. New associations prosocial or antisocial tease item accuracy and specific cue categories were found in Study 2, specifically accuracy on prosocial tease items was disadvantaged by gesture cues and accuracy on antisocial tease items was advantaged by relationship information and facial expression cues with relationship information cues producing a stronger accuracy advantage. These association should undergo replication and investigation to understand how they are differentially contributing to PATC.

The possible PATC differences by gender, age, and task duration were less supported. In line with previous research (e.g., Barnett et al., 2004; Jones et al., 2018), Study 1 found no major differences in PATC by gender, age, and task duration. Although gender, age, and task duration did have significant associations with PATC in Study 2, all the associations were quite weak and explained 1.2% or less of the variance in accuracy. Because these associations were still significant, these variables and their associations with PATC will need to be reassessed in future research.

Finally, as a result of the item analyses in Study 2 and specific expectations and requirements, a "final" PATC measure has been created that can proceed to further measure development and improvement. However, this measure and its item analysis and selection processes have several flaws from nonstandard and potentially arbitrary requirements for item selection to the inability to use item response theory approach due study design. Overall, this PATC measure will continue its development with some of these issues being addressable in subsequent studies, but it may be beneficial to rerun Study 2 in the future with the necessary study design and only the best and well-

performing items identified in Study 2. This will make more and more well-suited methods and statistical analyses available and may enable the development of a much-improved PATC measure in the future.

The findings overall include the expected and unexpected. As expected, the hypotheses of interest were supported and meet the expectations from social information processing theory (e.g., Dodge & Crick, 1990). The results also align with findings from irony comprehension research that also uses a cue-based framework to conceptualize comprehension (e.g., Rivière & Champagne-Lavau, 2020). Yet, even though differences in item performance by specific cue categories and by tease type were anticipated, the specific cue categories and the nature of their associations with the tease types were not predicted because no research is currently on how much a specific cue category contributes to different types of teasing. A similar question is gaining interest in irony comprehension research (e.g., Pexman, in press) and likewise currently does not have the research to properly answer how much a specific cue category contributes to comprehension. Studies 1 and 2, for example, point to a strong comprehension contribution by facial expression cues for prosocial tease comprehension. There is no current research to explain the specifics of why this is what both studies found, but research from other areas of psychology and social cognition may suggest some possible explanations.

First, considering and adapting from the Social Interaction Sequence (e.g., Darley & Fazio, 1980), which is normally used in expectancy models of prejudice, can clarify what type of information facial expression cues may be providing that benefits PATC

accuracy (see Figure 8). The Social Interaction Sequence can be simply understood as describing the different parts and sequence of a social interaction between two people. Before the recipient, teasee, or "target" of the interaction is fully involved, the sender, teaser, or "perceiver" already brings specific expectations about the target to the interaction. This can be understood as context, and, in terms of the construction of the tease scenarios for the PATC measure, relationship information cues tend to be provided first and give contextual information about the social interactions and teasing the teaser and teasee may have engaged in. After all, it may be expected that friends are more likely to engage in prosocial teasing. Then, the perceiver performs their part of the interaction and delivers their message to the target, which is shaped by the context. In the tease scenarios, this includes the tease itself, the gesture cue, and the teaser's part of the facial expression cue. Finally, the target responds to the social interaction and message. In the tease scenarios, only the teasee's part of the facial expression cue gives response information and the facial expression cue overall tends to be the last cue provided in the scenario. Altogether, this highlights that facial expression cues seem to provide more information than the other cue categories with facial expression information for both individuals in the tease scenario and that the facial expression cue is the only cue category to describe how the person being teased is responding and possibly interpreting the tease themselves. From this, it may be inferred that the accuracy advantage of facial expression cues, especially in prosocial teases, is simply due to the cue actually serving as two cues. However, it may also be that prosocial tease comprehension is harder than antisocial tease comprehension as suggested by their different developmental trajectories

(e.g., Barnett et al., 2004; Jones et al., 2018) and accuracy on antisocial teases being significantly higher than accuracy on prosocial teases in Study 2. As such, prosocial tease comprehension may need and depend upon the additional information facial expression cues provide and specifically knowing how the other person perceives the tease, whereas antisocial teases that benefit most strongly from relationship information cues may be easier and thus the first piece of information provided and the contextual information is enough for comprehension. However, even antisocial teases still benefit from facial expression cues to some extent. This all may also speak to people viewing the impact as more important to classifying a tease than the intent. These possibilities can and should be investigated in future PATC research.

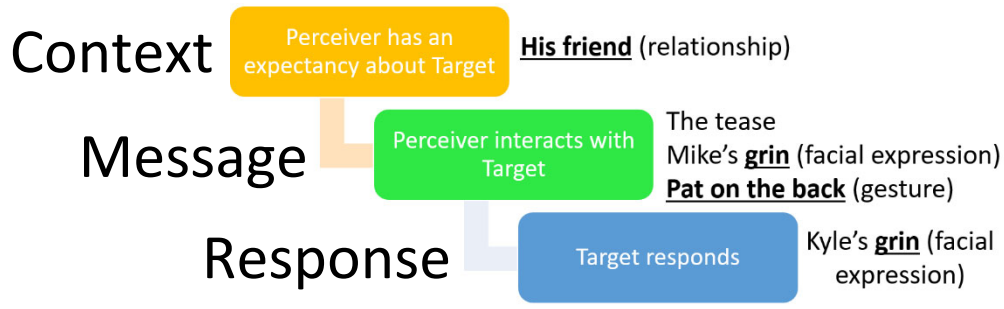

Figure 8. Social Interaction Sequence Applied to Figure 1 Tease Scenario

Another perspective considers the order that the cues are presented and why prosocial teases benefit strongly from facial expression cues and antisocial teases benefit strongly from relationship information cues. Essentially, Trope and Liberman (1996)

described how we engage in social hypothesis-testing and make conclusions based on different goals and situations. On one hand, there is the cost of information model, wherein the goal is to reach some conclusion quickly and thus information acquisition efforts are "costly." For example, if a person is fatigued or under some time pressure and needs to make a conclusion, they need to do so quickly and may gravitate to information that they can find quickly, such as the first pieces of information provided on a document (Trope & Liberman, 1996). In teasing comprehension and experiences, antisocial teasing may be more likely to follow the cost of information model. To explain, consider how an antisocial tease may be more likely to lead to an escalation in hostility and, at worst, a physical confrontation. Avoiding such engagements may necessitate quick comprehension and thus make information gathering costly. In line with all of these expectations, antisocial teases are more benefitted by the first cue provided, the relationship information cue. On the other hand, Trope and Liberman (1996) also describe the cost of error model, wherein the goal is to avoid inaccuracy in our conclusions and thus errors are costly. For example, if a person is experiencing evaluation apprehension and faces accountability for their conclusions, they work to come to an accurate conclusion and are likely to dedicate their time to gathering as much information as possible, increasing their chances of considering even the very last pieces of information provided on a document (Trope & Liberman, 1996). In teasing comprehension and experiences, prosocial teasing may be more likely to follow the cost of error model. Unlike antisocial teases, prosocial teases may provide the first piece of information that the interaction is between friends, and that immediately makes it

important to not misinterpret the tease and potentially harm the relationship in personal experiences. As such, using all the available information including the last pieces of information like facial expression cues would be the chosen approach to avoid inaccuracy and misunderstanding. From another point of view, prosocial teasing may be more likely to follow the cost of error model because prosocial tease comprehension is harder than antisocial tease comprehension and thus may involve more evaluation apprehension inherently. Again, this would lead to the high-information facial expression cue that is shared last in the scenario to being considered and used more in an effort to make an accurate conclusion. These models seem to align with the results, but these models and their goals can be induced by different factors and thus tested in future research to determine whether they can explain these results (Trope & Liberman, 1996).

However, a final possibility to consider is that prosocial facial expression cues are more diagnostic of prosocial teasing than the other prosocial cues and that antisocial relationship information cues are more diagnostic of antisocial teasing than other antisocial cues (Skowronski & Carlston, 1989). Another way to describe this is that prosocial facial expression cues have high cue validity for prosocial teasing and that antisocial relationship information cues have high cue validity for antisocial teasing. A cue with high cue validity would be specifically something that is present in at least most examples of a specific category (e.g., prosocial teases) and is absent in at least most examples of a separate category (e.g., antisocial teases; Sloutsky, 2015). In essence, a facial expression cue of both of those involved smiling or laughing may be more diagnostic of prosocial teasing than a relationship information cue of those involved

being friends and a gesture cue of a high-five or pat on the back, and a facial expression cue of both of those involved smiling or laughing may also have high cue validity because it is present in most cases of prosocial teasing and absent in most cases of antisocial teasing. The ideas of cue diagnosticity and cue validity seem highly applicable to the results of Studies 1 and 2 and to the cue-based theoretical basis of the measure and seem to be strong explanatory candidates to test in future research.

Along with these three perspectives, there may be other theories, models, and approaches that can explain these unexpected results, but Studies 1 and 2 on their own cannot determine which is best. Future studies are needed to identify and disqualify any of these potential explanations.

Limitations

A major limitation was the samples. First, the sample is Study 1 was quite small and may limit generalization. Second and likely exacerbated by the small sample size, the sample of Study 1 came from a college student convenience sampling pool. Such a sampling pool is expected to produce results that may be less representative of and less generalizable to other adults (e.g., Boehnke, Lietz, Schreier, & Wilhelm, 2011). As with Study 1, the sample of Study 2 is possibly not very representative because, although the sample was drawn from a much wider age range than Study 1 and may have included a more diverse and representative group than Study 1 (cf. Casler, Bickel, & Hackett, 2013), the sample was still taken from an online survey worker platform. Survey workers may be distinctly different from other adults (e.g., Paolacci & Chandler, 2014). Additionally,

the students complete studies for class credit, whereas survey workers complete studies for money, either as extra money or income. This collection of studies treats these differently-sourced participants as all experiencing their studies similarly when that may not be accurate. Conclusions from Studies 1 and 2 and their potentially less representative samples should be reevaluated with the current "final" measure or any future iterations and more representative samples in future research.

There were also study-specific limitations to consider. One possible limitation involves the evaluation explanations in Study 1, wherein participants were asked to explain why they made their tease type evaluation for every item. At least since Wilhelm Wundt's research with introspection and his criticisms of its use and misuse (e.g., Blumenthal, 1975), psychological researchers have discussed whether people can truly describe their cognitive processes and reasoning for their decisions and conclusions, and, as Wundt argued more than a century ago, the general consensus is that people are largely limited in their ability to accurately and thoroughly introspect on their own higher-level cognitive processes (Nisbett & Wilson, 1977). Nisbett and Wilson (1977) explained that study participants are often unaware of influential stimuli, of how these stimuli may influence their response, and of their response itself, and this seems to be affected by the individual's own biases and expectations. For example, if a participant reading a tease scenario does not consider a gesture cue as an important and causal cue of the teaser's intent, they may neglect to include it in their evaluation explanations, even if the gesture cue is actually influential in their tease type evaluation. So, on its own, the evaluation explanations and their associated content analysis may not be as informative as intended,

but Studies 1 and 2 are able to clarify the content analysis findings with more objective comparisons of item performance and can support the hypotheses of interest without relying on just the content analysis findings. Additionally, even if a participant failed to mention or be aware of every cue present in the tease scenarios when writing their evaluation explanations, simply including at least one cue can still inform on some of what may be occurring in the participant's cognitive processes.

A limitation specific to Study 2 was already mentioned alongside the item analyses: the aggregation across measure versions and different items. Many of the Study 2 findings rely on the aggregation across the 16 measure versions and 240 items. This is not necessarily a poor approach to these analyses, given the procedure and design of Study 2. However, the item analysis and item selection process revealed that several items did not perform well and may have obscured better items or affected performance on other items in the same measure version. Although this was not a majority of the items, this could still impact the results and thus the conclusions. Luckily, the aggregation of so many items can minimize the effect of a few poorly-performing items, but this should not discount the inherent problems of aggregating across different measure versions, items, and samples.

As noted, an item response theory approach would be the most beneficial and informative approach to use in developing a PATC measure. However, with another limitation specific to Study 2, the design of Study 2 was flawed and prevented the use of various analyses and the specific use of item response theory. Recall that equating across measure version requires an anchor test (Crocker & Algina, 2008). Study 2 had anchor

items in that ambiguous tease items were each included in two measure versions, but one item shared in only two of the 16 measure versions does not make for an informative anchor item, especially in a measure wherein ambiguous tease items do not represent a majority of the items. One item is also far less than the general standard of about 20% of the number of anchor items on either of the measure versions, which would need to be about three items for this measure (Angoff, 1971, as cited in Crocker & Algina, 2008, p. 466). Furthermore, item response theory anchor tests must measure the same latent traits assessed across measure versions (Crocker & Algina, 2008), so, again, just using one ambiguous tease item will not measure all the latent traits expected to be associated with this PATC measure (e.g., prosocial tease comprehension, antisocial tease comprehension, cue-based reasoning).

The final limitations to mention are general measure concerns. First, the tease scenarios and the tease items using them do not consider the effects of the content of tease itself. For example, the tease scenarios can involve people being teased based on their ability, appearance, things in their control versus out of their control, and a temporary or permanent feature or situation, and it could be expected that teases directed at abilities and more permanent features and situations may garner more antisocial evaluations. After all, these may be interpreted as personal attacks and inherently more antisocial. Some of the Study 1 evaluation explanations speak to the content mattering, with one participant describing, “a jab at someone’s appearance for which they can't completely control is more hurtful than just saying something about their clothing or a mistake they made.” For an ambiguous tease item that could be considered a tease about

ability (i.e., grades or academic ability), a participant wrote, "[…] bad grades aren't typically good material to tease someone about." Another participant responded to the same tease, "[T]his is mean even if they were friends." So, the content of tease might even supersede the cues in some cases, but this effect may be very individualized with another participant responding to the same tease with "Joking about lower than average test scores is a common tease between my friends and me." Future research may want to explore this situation more thoroughly.

Second, future research and the next studies with this PATC measure should revise the middle response option. In Studies 1 and 2, the "ambiguous" answer option was named "Not Sure." There is nothing inherently wrong with this serving as the middle option, especially as the instructions do explain that it should be used as the in-between option if they cannot decide between friendly and mean response options (i.e., when they are uncertain or "not sure"). Participants are also informed that it may be the right option in some cases (i.e., ambiguous tease items). However, observing the histograms for item responses reveal general normal distributions with friendly and mean response options but also less responses to the "Not Sure" option than should be expected by those distributions. The suppression of the middle option can cause issues with interpretations and statistical assumptions, so the middle option needs revision. A simple option would be to eliminate the middle option, but that would leave no information about uncertainty and ambiguity as well as no "correct" answer for the ambiguous tease items. Instead, the option that would be best is to change the wording of the middle option to make its purpose clearer, such as with "Neither Friendly nor Mean" or something similar. This

would be an excellent change for the next study with this PATC measure, especially before the next steps in validity and reliability testing take place.

The third and final limitation is that the present studies really are only the first steps in this measure development process. Although this PATC measure has a great start with these studies and has undergone much more testing and validation efforts than previous PATC measures, more work is still needed to fully address validity and reliability of the measure.

Next Steps

The next steps for this measure and its development first include what can be considered Studies 3 and 4. Study 3 further examines the validity of this PATC measure. It will be able to overcome some, but not all, the limitations mentioned above. For example, the middle option will be revised, and aggregation across measure versions will not be a concern. At this stage, the measure may be explored using at least some item response theory methods and analyses and will at least be examined using more traditional classical test theory approaches, such as a random sample of the participants being a part of an exploratory factor analysis to identify the factors in the measure and then the remaining participants being included in a confirmatory factor analysis to delve into the validity and nature of the measure. Other future studies should explore validity by comparing performance on this PATC measure directly to that of previous PATC measures and comparable measures like the teasing-specific materials from the Relational Inference in Social Communication database (Rothermich & Pell, 2015) used to study

nonliteral communication comprehension in recent studies (e.g., Rothermich et al., 2021). Study 3 will also investigate reliability more deeply than could be done in Studies 1 and 2, but Study 4 will evaluate the reliability of this PATC measure even more with a main focus on test-retest reliability.

The research following Studies 3 and 4 should first focus on revisiting the development of this PATC measure with item response theory approaches, if possible. Then, another main emphasis should be toward adapting the measure for children and adolescents because understanding PATC development was a major purpose for developing a better PATC measure. As those concentrations are being addressed, other studies can explore whether other theories can explain the findings of Studies 1 and 2, such as the differences in accuracy advantages by specific cue categories. One study could use experimental methods to test whether the presence of response information in the tease scenario explains why PATC, especially prosocial tease comprehension, benefits from facial expression cues, as well as why such accuracy advantages may occur.

Future research should also aim to learn more about PATC and teasing using the more finalized version of the PATC measure. As mentioned, the content of the teases may influence item performance in the PATC measure, so a future study could examine what that impact might be and the individual differences associated with that, which may lead to further improvements in the measure. The measure could additionally be used to explore potential mechanisms of PATC and its development, including whether theory of mind or social perspective-taking impacts PATC as it does with irony comprehension

(e.g., Pexman, in press). Other research questions, such as whether the gender of the teaser and teasee in the scenarios affects evaluations or if different modalities of the measure (i.e., text, pictures, videos) affect PATC, could be asked as well.

Developing this PATC measure has created the foundation for future research on PATC and teasing more generally, while illuminating aspects in need of further study that were not found using past PATC measures. These studies are a good start on a long road of measure development and improvement, and there is much more work ahead to truly capture and understand this complex form of communication.

References

Barnett, M. A., Barlett, N. D., Livengood, J. L., Murphy, D. L., & Brewton, K. E. (2010). Factors associated with children's anticipated responses to ambiguous teases. *Journal of Genetic Psychology, 171*(1), 54-72.

Barnett, M. A., Burns, S. R., Sanborn, F. W., Bartel, J. S., & Wilds, S. J. (2004). Antisocial and prosocial teasing among children: Perceptions and individual differences. *Social Development, 13*(2), 292-310.

Barnett, M. A., Nichols, M. B., Sonnentag, T. L., & Wadian, T. W. (2013). Factors associated with early adolescents' anticipated emotional and behavioral responses to ambiguous teases on Facebook. *Computers in Human Behavior*, *29*(6), 2225-2229.

Bland, J. M., & Altman, D. G. (1997). Statistics notes: Cronbach's alpha. *BMJ*, *314*, 572.

Blumenthal, A. L. (1975). A reappraisal of Wilhelm Wundt. *American Psychologist*, *30*(11), 1081-1088.

Boehnke, K., Lietz, P., Schreier, M., & Wilhelm, A. (2011). Sampling: The selection of cases for culturally comparative psychological research. In D. Matsumoto and F. J. R. Van de Vijver (Eds). *Cross-Cultural Research Methods in Psychology* (p. 101-129). Cambridge University Press.

Brunner, M., & Süβ, H. M. (2005). Analyzing the reliability of multidimensional measures: An example from intelligence research. *Educational and Psychological Measurement*, *65*(2), 227-240.

Capelli, C. A., Nakagawa, N., & Madden, C. M. (1990). How children understand sarcasm: The role of context and intonation. *Child Development*, *61*(6), 1824-1841.

Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior*, *29*(6), 2156-2160.

Climie, E. A., & Pexman, P. M. (2008). Eye gaze provides a window on children's understanding of verbal irony. *Journal of Cognition and Development*, *9*(3), 257-285.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Laurence Erlbaum Associates.

Cornell, D., Gregory, A., Huang, F., & Fan, X. (2013). Perceived prevalence of teasing and bullying predicts high school dropout rates. *Journal of Educational Psychology*, *105*(1), 138-149.

Crocker, L., & Algina, J. (2008). *Introduction to Classical and Modern Test Theory*. Cengage Learning.

Darley, J. M., & Fazio, R. H. (1980). Expectancy confirmation processes arising in the social interaction sequence. *American Psychologist, 35*, 867-881.

Demorest, A., Meyer, C., Phelps, E., Gardner, H., & Winner, E. (1984). Words speak louder than actions: Understanding deliberately false remarks. *Child Development, 55*(4), 1527-1534.

Dews, S., Winner, E., Kaplan, J., Rosenblatt, E., Hunt, M., Lim, K., McGovern, A., Qualter, A., & Smarsh, B. (1996). Children's understanding of the meaning and functions of verbal irony. *Child Development*, *67*(6), 3071-3085.

Dodge, K. A., & Crick, N. R. (1990). Social information-processing bases of aggressive behavior in children. *Personality and Social Psychology Bulletin*, *16*(1), 8-22.

Dodge, K. A., Price, J. M., Bachorowski, J. A., & Newman, J. P. (1990). Hostile attributional biases in severely aggressive adolescents. *Journal of Abnormal Psychology*, *99*(4), 385-392.

Eisenberg, A. R. (1986). Teasing: Verbal play in two Mexican homes. In B. B. Schieffelin & E. Ochs (Eds.), *Language socialization across cultures* (p. 182-198). Cambridge University Press.

Embretson, S. E., & Poggio, J. (2012). The impact of scaling and measurement methods on individual differences in growth. In B. Laursen, T. D. Little, & N. A. Card (Eds). *Handbook of Developmental Research Methods* (p. 82-108). Guilford Press.

Gibbs, R. J. (2000). Irony in talk among friends. *Metaphor and Symbol*, *15*(1-2), 5-27.

Glenwright, M., Tapley, B., Rano, J. K., & Pexman, P. M. (2017). Developing appreciation for sarcasm and sarcastic gossip: It depends on perspective. *Journal of Speech, Language, and Hearing Research*, *60*(11), 3295-3309.

Gravetter, F., & Wallnau, L. (2014). *Essentials of statistics for the behavioral sciences* (8th ed.). Wadsworth.

Grice, J. W., Medellin, E., Jones, I., Horvath, S., McDaniel, H., O'lansen, C., & Baker, M. (2020). Persons as effect sizes. *Advances in Methods and Practices in Psychological Science*, *3*(4), 443-455.

Happé, F. G. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders, 24*(2), 129-154.

Harris, M., & Pexman, P. M. (2003). Children's perceptions of the social functions of verbal irony. *Discourse Processes*, *36*(3), 147-165.

Huang, F. L., & Cornell, D. G. (2019). School teasing and bullying after the presidential election. *Educational Researcher*, *48*(2), 69-83.

Ivanko, S. L., & Pexman, P. M. (2003). Context incongruity and irony processing. *Discourse Processes*, *35*(3), 241-279.

Jones, T. L., Wadian, T. W., Barnett, M. A., Hellmer, M. K., & Pino, L. N. (2018). Young children's ability to discriminate between antisocial and prosocial teases. *Journal of Genetic Psychology, 179*(1), 1-8.

Keltner, D. (2008, December 5). In defense of teasing. *The New York Times*. http://www.nytimes.com/2008/12/07/magazine/07teasing-t.html

Keltner, D., Capps, L., Kring, A. M., Young, R. C., & Heerey, E. A. (2001). Just teasing: A conceptual analysis and empirical review. *Psychological Bulletin, 127*(2), 229-248.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174.

Miller, P. (1986). Teasing as language socialization and verbal play in a white working class community. In B. B. Schieffelin & E. Ochs (Eds.), *Language socialization across cultures* (p. 199-212). Cambridge University Press.

Mills, C. B., & Carwile, A. M. (2009). The good, the bad, and the borderline: Separating teasing from bullying. *Communication Education*, *58*(2), 276-301.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*(3), 231-259.

Paolacci, G., & Chandler, J. (2014). Inside the Turk: Understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science*, *23*(3), 184-188.

Pexman, P. M. (2008). It's fascinating research: The cognition of verbal irony. *Current Directions in Psychological Science*, *17*(4), 286-290.

Pexman, P. M. (in press). Irony and thought: Developmental insights. In R. Gibbs and H. Colston (Eds.), *The Cambridge Handbook of Irony and Thought*. Cambridge University Press. https://www.researchgate.net/publication/350740252_Irony_and_Thought_Developmental_Insights

Pexman, P. M., Glenwright, M., Krol, A., & James, T. (2005). An acquired taste: Children's perceptions of humor and teasing in verbal irony. *Discourse Processes*, *40*(3), 259-288.

Pexman, P. M., & Olineck, K. M. (2002). Understanding irony: How do stereotypes cue speaker intent?. *Journal of Language and Social Psychology*, *21*(3), 245-274.

Pexman, P. M., Rostad, K. R., McMorris, C. A., Climie, E. A., Stowkowy, J., & Glenwright, M. R. (2011). Processing of ironic language in children with high-functioning autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *41*(8), 1097-1112.

Pexman, P. M., Zdrazilova, L., McConnachie, D., Deater-Deckard, K., & Petrill, S. A. (2009). “That was smooth, mom”: Children's production of verbal and gestural irony. *Metaphor and Symbol*, *24*(4), 237-248.

Pexman, P. M., & Zvaigzne, M. T. (2004). Does irony go better with friends?. *Metaphor and Symbol*, *19*(2), 143-163.

Potter, G. B., Gibbs, J. C., Robbins, M., & Langdon, P. E. (2015). *A comprehensive cognitive behavioral program for offenders: Responsible adult culture.* Springer International.

Rivière, E., & Champagne-Lavau, M. (2020). Which contextual and sociocultural information predict irony perception?. *Discourse Processes*, *57*(3), 259-277.

Rothermich, K., Giorio, C., Falkins, S., Leonard, L., & Roberts, A. (2021). Nonliteral language processing across the lifespan. *Acta Psychologica*, *212*, 103213.

Rothermich, K., & Pell, M. D. (2015). Introducing RISC: A new video inventory for testing social perception. *PloS one*, *10*(7), e0133902.

Schvey, N. A., Marwitz, S. E., Mi, S. J., Galescu, O. A., Broadney, M. M., Young-Hyman, D., ... & Yanovski, J. A. (2019). Weight-based teasing is associated with gain in BMI and fat mass among children and adolescents at-risk for obesity: a longitudinal study. *Pediatric Obesity*, *14*(10), e12538.

Schieffelin, B. B. (1986). Teasing and shaming in Kaluli children's interactions. In B. B. Schieffelin & E. Ochs (Eds.), *Language socialization across cultures* (pp. 165-181). Cambridge University Press.

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, *74*(1), 107-120.

Skowronski, J. J., & Carlston, D. E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin*, *105*(1), 131-142.

Sloutsky, V. M. (2015) Conceptual development. In R. M. Lerner, L. S. Liben, & U. Mueller (Eds.), *Handbook of Child Psychology and Developmental Science* (7th ed., Vol. 2, p. 469-518). Wiley.

Sonnentag, T. L., Barnett, M. A., Wadian, T. W., & Nichols, M. B. (2016). The role of public self-consciousness and other factors in adolescents' anticipated responses to ambiguous teases on Facebook. *North American Journal of Psychology*, *18*(1), 59-69.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, *2*, 53-55.

Trope, Y., & Liberman, A. (1996). Social hypothesis testing: Cognitive and motivational mechanisms. In E. T. Higgins and A. W. Kruglanski (Eds). *Social Psychology: Handbook of Basic Principles*. Guilford Press.

Vessey, J. A., & O'Neill, K. M. (2011). Helping students with disabilities better address teasing and bullying situations: A MASNRN study. *The Journal of School Nursing*, *27*(2), 139-148.

Warm, T. R. (1997). The role of teasing in development and vice versa. *Journal of Developmental and Behavioral Pediatrics, 18*(2), 97-101.

Whalen, J. M., Doyle, A., & Pexman, P. M. (2020). Sarcasm between siblings: Children's use of relationship information in processing ironic remarks. *Journal of Pragmatics*, *156*, 149-159.

Appendix A. Tease Scenarios in Study 1

p3	Kyle accidentally got his soda all over his shirt and pants. His friend, Mike, says, “You’re always so smooth.” They grin at each other as Mike gives him a pat on the back.
p2gr	Ron got hit in the face by the ball while playing basketball. His friend, Cody, pats him on the shoulder and says, “Next time, you might want to try catching the ball with your hands instead of with your face.”
p2fr	Gary pulls off his winter hat, and his hair is messy. His buddy, Adam, grins and says, “Well, it looks like it is Wild Hair Day and someone forgot to tell me.” Gary smiles, too.
p2gf	Alex finished his lunch quickly and ended up going to class with some food on his face. John grins and passes Alex the tissue box in the room. John then asks, “Did any of your lunch make it into your mouth?” Alex smiles back.
p1r	Dan tries to push open a door, but the door is not moving. His friend, Steve, successfully pulls open the door instead and says, “Maybe you should learn to read the signs.”
p1g	Ethan came last in a footrace. Max high-fives Ethan as he finishes the race and says, “My grandmother could give you a head start and still beat you in a race.”
p1f	Rick stayed up late last night and has dark bags under his eyes. Paul says, “I didn’t know you were a raccoon.” They both laugh.
a3	Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his ears look bigger than they are. Nick, who is always rude to him, sees him and says, “Did your ears grow?” Nick pulls at his own ears and laughs, which makes Sam frown.
a2gr	Carter came to class with his shirt on backwards. Will, who always gets on his nerves, points at the tag sticking out and says, “Maybe you should get dressed with the lights on next time.”
a2fr	Greg accidentally wore his little brother’s shirt to class, and it has Mickey Mouse on the front. Brian, a classmate Greg hates, smirks and says, “I had a shirt just like that when I was in preschool.” Greg scowls.
a2gf	On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew sneers at Josh and asks, “Did you swim in the ocean on your way here?” Drew blocks the empty seat next to him with his feet, and Josh frowns.
a1r	Peter arrived late to work and got scolded by his boss. Isaac, a coworker he dislikes, asks, “So, when are you getting your employee of the month award?”
a1g	Mark trips and drops his books. Zack walks by without stopping and asks, “Did you have a nice trip?”
a1f	Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony scrunches his face up and says, “Thanks for sharing.” Matt glares.
amb-g	Luke got a bad grade on a test. Adrian sees and says, “That looks like a new personal best for you.”
amb-t	David ate a popsicle that made his teeth bright red. Brad sees him and says, “I think you need to meet some friends of mine, Toothbrush and Toothpaste.”

Appendix B. Raters' Coding Instructions for Study 1

Study 1 requests that the participants explain their reasons for their chosen response option (Friendly, Mean, or Not Sure). The categories/coding elements used to organize their reasoning are as follows:

Cues (C)

Code a response as **C**, if the participant uses the cues present in the tease scenario as their reasoning for their interpretation. This can involve listing the cues that are present or restating the exact scenario in their own words.

Examples

Well, it said they were friends.
They were laughing afterward, so it must have been friendly.
He didn't stop to help him, so it was a mean tease.

Personal Knowledge/Experience (P)

Code a response as **P**, if the participant uses their own personal experiences with teasing. This can involve citing similar teases that they have experienced or discussed a person they know who teases a lot.

Examples

My friends always tease me like this, so I know it is friendly.
A bully did something similar to me, so it must be mean.

Generalization of the Nature of Teasing (G)

Code a response as **G**, if the participant use general statements about teasing that are meant to transcend the specific scenario and apply to all teasing instances. This can involve describing teasing as having a specific, inherent characteristic that is unaffected by situational factors.

Examples

Teasing is not something you should do, so it's always mean.
Teasing is always situational.
All teasing is lighthearted.

Self-Generated Information about the Scenario (SG)

Code a response as **SG**, if the participant uses created/assumed information that is not actually in the scenario as part of their reasoning. This can involve the participant taking a present cue and using it to make further assumptions about the situation, such as adding self-generated details about the two boys' relationship or their behaviors/actions within/outside of this situation.

Examples

They hang out all the time, so they probably always joke around like this.

He didn't stop to help him, so he must be an awful person, who always bullies people.

Maybe he is saving that seat for a different person, so it's not necessarily mean.

Other Reasons (O)

Code a response as **O**, if the participant's response does not include any of the elements described in other categories. This can involve no response being given or the participant writing a non-answer to the question.

Examples

Just because.

I don't know.

Appendix C. Item Variations by Scenario in Study 2

Bolded items mark the sequence of items selected for each measure version with the bolded items making up Version 1, the items right below the bolded items making up Version 2, and so on.

Base Tease: Kyle accidentally got his soda all over his shirt and pants. Mike says, "You're always so smooth."

p3	**Kyle accidentally got his soda all over his shirt and pants. His friend, Mike, says, "You're always so smooth." They grin at each other as Mike gives him a pat on the back.**
p2gr	Kyle accidentally got his soda all over his shirt and pants. His friend, Mike, says, "You're always so smooth." Mike gives him a pat on the back.
p2fr	Kyle accidentally got his soda all over his shirt and pants. His friend, Mike, says, "You're always so smooth." They grin at each other.
p2gf	Kyle accidentally got his soda all over his shirt and pants. Mike says, "You're always so smooth." They grin at each other as Mike gives him a pat on the back.
p1r	Kyle accidentally got his soda all over his shirt and pants. His friend, Mike, says, "You're always so smooth."
p1g	Kyle accidentally got his soda all over his shirt and pants. Mike says, "You're always so smooth." Mike gives him a pat on the back.
p1f	Kyle accidentally got his soda all over his shirt and pants. Mike says, "You're always so smooth." They grin at each other.
a3	Kyle accidentally got his soda all over his shirt and pants. His bully, Mike says, "You're always so smooth." They frown at each other as Mike gives him a shove on the back.
a2gr	Kyle accidentally got his soda all over his shirt and pants. His bully, Mike gives him a shove on the back and says, "You're always so smooth."
a2fr	Kyle accidentally got his soda all over his shirt and pants. His bully, Mike says, "You're always so smooth." They frown at each other.
a2gf	Kyle accidentally got his soda all over his shirt and pants. Mike says, "You're always so smooth." They frown at each other as Mike gives him a shove on the back.
a1r	Kyle accidentally got his soda all over his shirt and pants. His bully, Mike says, "You're always so smooth."
a1g	Kyle accidentally got his soda all over his shirt and pants. Mike gives him a shove on the back and says, "You're always so smooth."
a1f	Kyle accidentally got his soda all over his shirt and pants. Mike says, "You're always so smooth." They frown at each other.
amb	Kyle accidentally got his soda all over his shirt and pants. Mike says, "You're always so smooth."

Base Tease: Alex finished his lunch quickly and ended up going to class with some food on his face. John asks, "Did any of your lunch make it into your mouth?"

p3	Alex finished his lunch quickly and ended up going to class with some food on his face. His friend, John, grins and passes Alex the tissue box in the room. John then asks, "Did any of your lunch make it into your mouth?" Alex smiles back.
p2gr	**Alex finished his lunch quickly and ended up going to class with some food on his face. His friend, John, passes Alex the tissue box in the room. John then asks, "Did any of your lunch make it into your mouth?"**
p2fr	Alex finished his lunch quickly and ended up going to class with some food on his face. His friend, John, grins and asks, "Did any of your lunch make it into your mouth?" Alex smiles back.
p2gf	Alex finished his lunch quickly and ended up going to class with some food on his face. John grins and passes Alex the tissue box in the room. John then asks, "Did any of your lunch make it into your mouth?" Alex smiles back.
p1r	Alex finished his lunch quickly and ended up going to class with some food on his face. His friend, John, asks, "Did any of your lunch make it into your mouth?"
p1g	Alex finished his lunch quickly and ended up going to class with some food on his face. John passes Alex the tissue box in the room and asks, "Did any of your lunch make it into your mouth?"
p1f	Alex finished his lunch quickly and ended up going to class with some food on his face. John grins and asks, "Did any of your lunch make it into your mouth?" Alex smiles back.
a3	Alex finished his lunch quickly and ended up going to class with some food on his face. John, who despises him, grins and points at Alex's face. John then asks, "Did any of your lunch make it into your mouth?" Alex glares back.
a2gr	Alex finished his lunch quickly and ended up going to class with some food on his face. John, who despises him, points at Alex's face and asks, "Did any of your lunch make it into your mouth?"
a2fr	Alex finished his lunch quickly and ended up going to class with some food on his face. John, who despises him, grins and asks, "Did any of your lunch make it into your mouth?" Alex glares back.
a2gf	Alex finished his lunch quickly and ended up going to class with some food on his face. John grins and points at Alex's face. John then asks, "Did any of your lunch make it into your mouth?" Alex glares back.
a1r	Alex finished his lunch quickly and ended up going to class with some food on his face. John, who despises him, asks, "Did any of your lunch make it into your mouth?"
a1g	Alex finished his lunch quickly and ended up going to class with some food on his face. John points at Alex's face and asks, "Did any of your lunch make it into your mouth?"
a1f	Alex finished his lunch quickly and ended up going to class with some food on his face. John grins and asks, "Did any of your lunch make it into your mouth?" Alex glares back.
amb	Alex finished his lunch quickly and ended up going to class with some food on his face. John asks, "Did any of your lunch make it into your mouth?"

Base Tease: Ron got hit in the face by the ball while playing basketball. Cody says, "Next time, you might want to try catching the ball with your hands instead of with your face."

p3	Ron got hit in the face by the ball while playing basketball. His friend, Cody, pats him on the shoulder and says, "Next time, you might want to try catching the ball with your hands instead of with your face." They share a laugh.
p2gr	Ron got hit in the face by the ball while playing basketball. His friend, Cody, pats him on the shoulder and says, "Next time, you might want to try catching the ball with your hands instead of with your face."
p2fr	Ron got hit in the face by the ball while playing basketball. His friend, Cody, says, "Next time, you might want to try catching the ball with your hands instead of with your face." They share a laugh.
p2gf	**Ron got hit in the face by the ball while playing basketball. Cody pats him on the shoulder and says, "Next time, you might want to try catching the ball with your hands instead of with your face." They share a laugh.**
p1r	Ron got hit in the face by the ball while playing basketball. His friend, Cody, says, "Next time, you might want to try catching the ball with your hands instead of with your face."
p1g	Ron got hit in the face by the ball while playing basketball. Cody pats him on the shoulder and says, "Next time, you might want to try catching the ball with your hands instead of with your face."
p1f	Ron got hit in the face by the ball while playing basketball. Cody says, "Next time, you might want to try catching the ball with your hands instead of with your face." They share a laugh.
a3	Ron got hit in the face by the ball while playing basketball. Cody, who doesn't like him, purposefully knocks into his shoulder and says, "Next time, you might want to try catching the ball with your hands instead of with your face." Cody laughs, and Ron does not.
a2gr	Ron got hit in the face by the ball while playing basketball. Cody, who doesn't like him, purposefully knocks into his shoulder and says, "Next time, you might want to try catching the ball with your hands instead of with your face."
a2fr	Ron got hit in the face by the ball while playing basketball. Cody, who doesn't like him, says, "Next time, you might want to try catching the ball with your hands instead of with your face." Cody laughs, and Ron does not.
a2gf	Ron got hit in the face by the ball while playing basketball. Cody purposefully knocks into his shoulder and says, "Next time, you might want to try catching the ball with your hands instead of with your face." Cody laughs, and Ron does not.
a1r	Ron got hit in the face by the ball while playing basketball. Cody, who doesn't like him, says, "Next time, you might want to try catching the ball with your hands instead of with your face."
a1g	Ron got hit in the face by the ball while playing basketball. Cody purposefully knocks into his shoulder and says, "Next time, you might want to try catching the ball with your hands instead of with your face."
a1f	Ron got hit in the face by the ball while playing basketball. Cody says, "Next time, you might want to try catching the ball with your hands instead of with your face." Cody laughs, and Ron does not.
amb	Ron got hit in the face by the ball while playing basketball. Cody says, "Next time, you might want to try catching the ball with your hands instead of with your face."

Base Tease: Ethan came last in a footrace. Max says, “My mom could beat you in a race.”

p3	Ethan came last in a footrace. His friend, Max, high-fives Ethan as he finishes the race and says, “My mom could beat you in a race.” They grin at each other.
p2gr	Ethan came last in a footrace. His friend, Max, high-fives Ethan as he finishes the race and says, “My mom could beat you in a race.”
p2fr	Ethan came last in a footrace. His friend, Max, says, “My mom could beat you in a race.” They grin at each other.
p2gf	Ethan came last in a footrace. Max high-fives Ethan as he finishes the race and says, “My mom could beat you in a race.” They grin at each other.
p1r	Ethan came last in a footrace. His friend, Max, says, “My mom could beat you in a race.”
p1g	**Ethan came last in a footrace. Max high-fives Ethan as he finishes the race and says, “My mom could beat you in a race.”**
p1f	Ethan came last in a footrace. Max says, “My mom could beat you in a race.” They grin at each other.
a3	Ethan came last in a footrace. Max, who is rarely nice to him, tries to trip Ethan as he finishes the race and says, “My mom could beat you in a race.” Max smirks at Ethan.
a2gr	Ethan came last in a footrace. Max, who is rarely nice to him, tries to trip Ethan as he finishes the race and says, “My mom could beat you in a race.”
a2fr	Ethan came last in a footrace. Max, who is rarely nice to him, says, “My mom could beat you in a race.” Max smirks at Ethan.
a2gf	Ethan came last in a footrace. Max tries to trip Ethan as he finishes the race and says, “My mom could beat you in a race.” Max smirks at Ethan.
a1r	Ethan came last in a footrace. Max, who is rarely nice to him, says, “My mom could beat you in a race.”
a1g	Ethan came last in a footrace. Max tries to trip Ethan as he finishes the race and says, “My mom could beat you in a race.”
a1f	Ethan came last in a footrace. Max says, “My mom could beat you in a race.” Max smirks at Ethan.
amb	Ethan came last in a footrace. Max says, “My mom could beat you in a race.”

Base Tease: Gary pulls off his winter hat, and his hair is messy. Adam says, "Well, it looks like it is Wild Hair Day and someone forgot to tell me."

p3	Gary pulls off his winter hat, and his hair is messy. His buddy, Adam, grins and says, "Well, it looks like it is Wild Hair Day and someone forgot to tell me." Adam ruffles his hair, and Gary smiles.
p2gr	Gary pulls off his winter hat, and his hair is messy. His buddy, Adam, ruffles his hair and says, "Well, it looks like it is Wild Hair Day and someone forgot to tell me."
p2fr	**Gary pulls off his winter hat, and his hair is messy. His buddy, Adam, grins and says, "Well, it looks like it is Wild Hair Day and someone forgot to tell me." Gary smiles.**
p2gf	Gary pulls off his winter hat, and his hair is messy. Adam grins and says, "Well, it looks like it is Wild Hair Day and someone forgot to tell me." Adam ruffles his hair, and Gary smiles.
p1r	Gary pulls off his winter hat, and his hair is messy. His buddy, Adam, says, "Well, it looks like it is Wild Hair Day and someone forgot to tell me."
p1g	Gary pulls off his winter hat, and his hair is messy. Adam ruffles his hair and says, "Well, it looks like it is Wild Hair Day and someone forgot to tell me."
p1f	Gary pulls off his winter hat, and his hair is messy. Adam grins and says, "Well, it looks like it is Wild Hair Day and someone forgot to tell me." Gary smiles.
a3	Gary pulls off his winter hat, and his hair is messy. His bully, Adam, smirks and says, "Well, it looks like it is Wild Hair Day and someone forgot to tell me." Adam steals his hat, and Gary frowns.
a2gr	Gary pulls off his winter hat, and his hair is messy. His bully, Adam, steals his hat and says, "Well, it looks like it is Wild Hair Day and someone forgot to tell me."
a2fr	Gary pulls off his winter hat, and his hair is messy. His bully, Adam, smirks and says, "Well, it looks like it is Wild Hair Day and someone forgot to tell me." Gary frowns.
a2gf	Gary pulls off his winter hat, and his hair is messy. Adam smirks and says, "Well, it looks like it is Wild Hair Day and someone forgot to tell me." Adam steals his hat, and Gary frowns.
a1r	Gary pulls off his winter hat, and his hair is messy. His bully, Adam, says, "Well, it looks like it is Wild Hair Day and someone forgot to tell me."
a1g	Gary pulls off his winter hat, and his hair is messy. Adam steals his hat and says, "Well, it looks like it is Wild Hair Day and someone forgot to tell me."
a1f	Gary pulls off his winter hat, and his hair is messy. Adam smirks and says, "Well, it looks like it is Wild Hair Day and someone forgot to tell me." Gary frowns.
amb	Gary pulls off his winter hat, and his hair is messy. Adam says, "Well, it looks like it is Wild Hair Day and someone forgot to tell me."

Base Tease: Dan tries to push open a door, but the door is not moving. Steve successfully pulls open the door instead and says, "Maybe you should learn to read the signs."

p3	Dan tries to push open a door, but the door is not moving. His friend, Steve, successfully pulls open the door instead, holds the door open for him, and says, "Maybe you should learn to read the signs." They snicker over it.
p2gr	Dan tries to push open a door, but the door is not moving. His friend, Steve, successfully pulls open the door instead, holds the door open for him, and says, "Maybe you should learn to read the signs."
p2fr	Dan tries to push open a door, but the door is not moving. His friend, Steve, successfully pulls open the door instead and says, "Maybe you should learn to read the signs." They snicker over it.
p2gf	Dan tries to push open a door, but the door is not moving. Steve successfully pulls open the door instead, holds the door open for him, and says, "Maybe you should learn to read the signs." They snicker over it.
p1r	**Dan tries to push open a door, but the door is not moving. His friend, Steve, successfully pulls open the door instead and says, "Maybe you should learn to read the signs."**
p1g	Dan tries to push open a door, but the door is not moving. Steve successfully pulls open the door instead, holds the door open for him, and says, "Maybe you should learn to read the signs."
p1f	Dan tries to push open a door, but the door is not moving. Steve successfully pulls open the door instead and says, "Maybe you should learn to read the signs." They snicker over it.
a3	Dan tries to push open a door, but the door is not moving. Steve, a guy who doesn't like Dan, successfully pulls open the door instead and says, "Maybe you should learn to read the signs." Steve shuts the door in Dan's face as Dan glares.
a2gr	Dan tries to push open a door, but the door is not moving. Steve, a guy who doesn't like Dan, successfully pulls open the door instead and says, "Maybe you should learn to read the signs." Steve shuts the door in Dan's face.
a2fr	Dan tries to push open a door, but the door is not moving. Steve, a guy who doesn't like Dan, successfully pulls open the door instead and says, "Maybe you should learn to read the signs." Dan glares.
a2gf	Dan tries to push open a door, but the door is not moving. Steve successfully pulls open the door instead and says, "Maybe you should learn to read the signs." Steve shuts the door in Dan's face as Dan glares.
a1r	Dan tries to push open a door, but the door is not moving. Steve, a guy who doesn't like Dan, successfully pulls open the door instead and says, "Maybe you should learn to read the signs."
a1g	Dan tries to push open a door, but the door is not moving. Steve successfully pulls open the door instead and says, "Maybe you should learn to read the signs." Steve shuts the door in Dan's face.
a1f	Dan tries to push open a door, but the door is not moving. Steve successfully pulls open the door instead and says, "Maybe you should learn to read the signs." Dan glares.
amb	Dan tries to push open a door, but the door is not moving. Steve successfully pulls open the door instead and says, "Maybe you should learn to read the signs."

Base Tease: Rick stayed up late last night and has dark bags under his eyes. Paul says, "I didn't know you were a raccoon."

p3	Rick stayed up late last night and has dark bags under his eyes. Paul, who always has his back, hands him a cup of coffee and says, "I didn't know you were a raccoon." They both laugh.
p2gr	Rick stayed up late last night and has dark bags under his eyes. Paul, who always has his back, hands him a cup of coffee and says, "I didn't know you were a raccoon."
p2fr	Rick stayed up late last night and has dark bags under his eyes. Paul, who always has his back, says, "I didn't know you were a raccoon." They both laugh.
p2gf	Rick stayed up late last night and has dark bags under his eyes. Paul hands him a cup of coffee and says, "I didn't know you were a raccoon." They both laugh.
p1r	Rick stayed up late last night and has dark bags under his eyes. Paul, who always has his back, says, "I didn't know you were a raccoon."
p1g	Rick stayed up late last night and has dark bags under his eyes. Paul hands him a cup of coffee and says, "I didn't know you were a raccoon."
p1f	**Rick stayed up late last night and has dark bags under his eyes. Paul says, "I didn't know you were a raccoon." They both laugh.**
a3	Rick stayed up late last night and has dark bags under his eyes. Paul, who always tries to pick a fight with Rick, pokes his face and says, "I didn't know you were a raccoon." Paul laughs, and Rick scowls.
a2gr	Rick stayed up late last night and has dark bags under his eyes. Paul, who always tries to pick a fight with Rick, pokes his face and says, "I didn't know you were a raccoon."
a2fr	Rick stayed up late last night and has dark bags under his eyes. Paul, who always tries to pick a fight with Rick, says, "I didn't know you were a raccoon." Paul laughs, and Rick scowls.
a2gf	Rick stayed up late last night and has dark bags under his eyes. Paul pokes his face and says, "I didn't know you were a raccoon." Paul laughs, and Rick scowls.
a1r	Rick stayed up late last night and has dark bags under his eyes. Paul, who always tries to pick a fight with Rick, says, "I didn't know you were a raccoon."
a1g	Rick stayed up late last night and has dark bags under his eyes. Paul pokes his face and says, "I didn't know you were a raccoon."
a1f	Rick stayed up late last night and has dark bags under his eyes. Paul says, "I didn't know you were a raccoon." Paul laughs, and Rick scowls.
amb	Rick stayed up late last night and has dark bags under his eyes. Paul says, "I didn't know you were a raccoon."

Base Tease: Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his ears look bigger than they are. Nick sees him and says, “Did your ears grow?”

p3	Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his ears look bigger than they are. Nick, who is always nice to him, sees him and says, “Did your ears grow?” Nick pats him on the back, and they share a smile.
p2gr	Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his ears look bigger than they are. Nick, who is always nice to him, sees him and says, “Did your ears grow?” Nick pats him on the back.
p2fr	Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his ears look bigger than they are. Nick, who is always nice to him, sees him and says, “Did your ears grow?” They share a smile.
p2gf	Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his ears look bigger than they are. Nick sees him and says, “Did your ears grow?” Nick pats him on the back, and they share a smile.
p1r	Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his ears look bigger than they are. Nick, who is always nice to him, sees him and says, “Did your ears grow?”
p1g	Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his ears look bigger than they are. Nick sees him and says, “Did your ears grow?” Nick pats him on the back.
p1f	Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his ears look bigger than they are. Nick sees him and says, “Did your ears grow?” They share a smile.
a3	**Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his ears look bigger than they are. Nick, who is always rude to him, sees him and says, “Did your ears grow?” Nick pulls at his own ears and laughs, which makes Sam frown.**
a2gr	Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his ears look bigger than they are. Nick, who is always rude to him, sees him and says, “Did your ears grow?” Nick pulls at his own ears.
a2fr	Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his ears look bigger than they are. Nick, who is always rude to him, sees him and says, “Did your ears grow?” Nick laughs, which makes Sam frown.
a2gf	Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his ears look bigger than they are. Nick sees him and says, “Did your ears grow?” Nick pulls at his own ears and laughs, which makes Sam frown.
a1r	Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his ears look bigger than they are. Nick, who is always rude to him, sees him and says, “Did your ears grow?”
a1g	Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his ears look bigger than they are. Nick sees him and says, “Did your ears grow?” Nick pulls at his own ears.
a1f	Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his ears look bigger than they are. Nick sees him and says, “Did your ears grow?” Nick laughs, which makes Sam frown.
amb	Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his ears look bigger than they are. Nick sees him and says, “Did your ears grow?”

Base Tease: Greg accidentally wore his little brother's shirt to class, and it has Mickey Mouse on the front. Brian says, "I had a shirt just like that when I was in preschool."

p3	Greg accidentally wore his little brother's shirt to class, and it has Mickey Mouse on the front. Brian, a classmate who respects Greg, grins and says, "I had a shirt just like that when I was in preschool." Brian fist-bumps him, making Greg smile.
p2gr	Greg accidentally wore his little brother's shirt to class, and it has Mickey Mouse on the front. Brian, a classmate who respects Greg, fist-bumps him and says, "I had a shirt just like that when I was in preschool."
p2fr	Greg accidentally wore his little brother's shirt to class, and it has Mickey Mouse on the front. Brian, a classmate who respects Greg, grins and says, "I had a shirt just like that when I was in preschool." Greg smiles back.
p2gf	Greg accidentally wore his little brother's shirt to class, and it has Mickey Mouse on the front. Brian, grins and says, "I had a shirt just like that when I was in preschool." Brian fist-bumps him, making Greg smile.
p1r	Greg accidentally wore his little brother's shirt to class, and it has Mickey Mouse on the front. Brian, a classmate who respects Greg, says, "I had a shirt just like that when I was in preschool."
p1g	Greg accidentally wore his little brother's shirt to class, and it has Mickey Mouse on the front. Brian fist-bumps him and says, "I had a shirt just like that when I was in preschool."
p1f	Greg accidentally wore his little brother's shirt to class, and it has Mickey Mouse on the front. Brian grins and says, "I had a shirt just like that when I was in preschool." Greg smiles back.
a3	Greg accidentally wore his little brother's shirt to class, and it has Mickey Mouse on the front. Brian, a classmate who hates Greg, smirks and says, "I had a shirt just like that when I was in preschool." Brian points at the shirt, making Greg scowl.
a2gr	Greg accidentally wore his little brother's shirt to class, and it has Mickey Mouse on the front. Brian, a classmate who hates Greg, points at the shirt and says, "I had a shirt just like that when I was in preschool."
a2fr	**Greg accidentally wore his little brother's shirt to class, and it has Mickey Mouse on the front. Brian, a classmate who hates Greg, smirks and says, "I had a shirt just like that when I was in preschool." Greg scowls.**
a2gf	Greg accidentally wore his little brother's shirt to class, and it has Mickey Mouse on the front. Brian smirks and says, "I had a shirt just like that when I was in preschool." Brian points at the shirt, making Greg scowl.
a1r	Greg accidentally wore his little brother's shirt to class, and it has Mickey Mouse on the front. Brian, a classmate who hates Greg, says, "I had a shirt just like that when I was in preschool."
a1g	Greg accidentally wore his little brother's shirt to class, and it has Mickey Mouse on the front. Brian points at the shirt and says, "I had a shirt just like that when I was in preschool."
a1f	Greg accidentally wore his little brother's shirt to class, and it has Mickey Mouse on the front. Brian smirks and says, "I had a shirt just like that when I was in preschool." Greg scowls.
amb	Greg accidentally wore his little brother's shirt to class, and it has Mickey Mouse on the front. Brian says, "I had a shirt just like that when I was in preschool."

Base Tease: Carter came to class with his shirt on backwards. Will says, "Maybe you should get dressed with the lights on next time."

p3	Carter came to class with his shirt on backwards. Will, who always jokes around with him, knocks on the front of the backwards shirt and says, "Maybe you should get dressed with the lights on next time." Carter and Will chuckle.
p2gr	Carter came to class with his shirt on backwards. Will, who always jokes around with him, knocks on the front of the backwards shirt and says, "Maybe you should get dressed with the lights on next time."
p2fr	Carter came to class with his shirt on backwards. Will, who always jokes around with him, says, "Maybe you should get dressed with the lights on next time." Carter and Will chuckle.
p2gf	Carter came to class with his shirt on backwards. Will knocks on the front of the backwards shirt and says, "Maybe you should get dressed with the lights on next time." Carter and Will chuckle.
p1r	Carter came to class with his shirt on backwards. Will, who always jokes around with him, says, "Maybe you should get dressed with the lights on next time."
p1g	Carter came to class with his shirt on backwards. Will knocks on the front of the backwards shirt and says, "Maybe you should get dressed with the lights on next time."
p1f	Carter came to class with his shirt on backwards. Will says, "Maybe you should get dressed with the lights on next time." Carter and Will chuckle.
a3	Carter came to class with his shirt on backwards. Will, who purposefully gets on his nerves, laughs and points at the tag sticking out. Will then says, "Maybe you should get dressed with the lights on next time." Carter glares.
a2gr	**Carter came to class with his shirt on backwards. Will, who purposefully gets on his nerves, points at the tag sticking out and says, "Maybe you should get dressed with the lights on next time."**
a2fr	Carter came to class with his shirt on backwards. Will, who purposefully gets on his nerves, laughs and says, "Maybe you should get dressed with the lights on next time." Carter glares.
a2gf	Carter came to class with his shirt on backwards. Will laughs and points at the tag sticking out. Will then says, "Maybe you should get dressed with the lights on next time." Carter glares.
a1r	Carter came to class with his shirt on backwards. Will, who purposefully gets on his nerves, says, "Maybe you should get dressed with the lights on next time."
a1g	Carter came to class with his shirt on backwards. Will points at the tag sticking out and says, "Maybe you should get dressed with the lights on next time."
a1f	Carter came to class with his shirt on backwards. Will laughs and says, "Maybe you should get dressed with the lights on next time." Carter glares.
amb	Carter came to class with his shirt on backwards. Will says, "Maybe you should get dressed with the lights on next time."

Base Tease: Mark trips and drops his books. Zack asks, “Did you have a nice trip?”

p3	Mark trips and drops his books. His pal, Zack picks up some of the books and asks, “Did you have a nice trip?” They both laugh.
p2gr	Mark trips and drops his books. His pal, Zack picks up some of the books and asks, “Did you have a nice trip?”
p2fr	Mark trips and drops his books. His pal, Zack asks, “Did you have a nice trip?” They both laugh.
p2gf	Mark trips and drops his books. Zack picks up some of the books and asks, “Did you have a nice trip?” They both laugh.
p1r	Mark trips and drops his books. His pal, Zack asks, “Did you have a nice trip?”
p1g	Mark trips and drops his books. Zack picks up some of the books and asks, “Did you have a nice trip?”
p1f	Mark trips and drops his books. Zack asks, “Did you have a nice trip?” They both laugh.
a3	Mark trips and drops his books. Zack, who always tries to upset Mark, walks by without stopping and asks, “Did you have a nice trip?” They both glare.
a2gr	Mark trips and drops his books. Zack, who always tries to upset Mark, walks by without stopping and asks, “Did you have a nice trip?”
a2fr	Mark trips and drops his books. Zack, who always tries to upset Mark, asks, “Did you have a nice trip?” They both glare.
a2gf	Mark trips and drops his books. Zack, who always tries to upset Mark, asks, “Did you have a nice trip?” They both glare.
a1r	Mark trips and drops his books. Zack, who always tries to upset Mark, asks, “Did you have a nice trip?”
a1g	**Mark trips and drops his books. Zack walks by without stopping and asks, “Did you have a nice trip?”**
a1f	Mark trips and drops his books. Zack asks, “Did you have a nice trip?” They both glare.
amb	Mark trips and drops his books. Zack asks, “Did you have a nice trip?”

Base Tease: On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew asks, "Did you swim in the ocean on your way here?"

p3	On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew, who gets along with him, taps the empty seat next to him and asks, "Did you swim in the ocean on your way here?" Josh and Drew grin.
p2gr	On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew, who gets along with him, taps the empty seat next to him and asks, "Did you swim in the ocean on your way here?"
p2fr	On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew, who gets along with him, asks, "Did you swim in the ocean on your way here?" Josh and Drew grin.
p2gf	On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew taps the empty seat next to him and asks, "Did you swim in the ocean on your way here?" Josh and Drew grin.
p1r	On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew, who gets along with him, asks, "Did you swim in the ocean on your way here?"
p1g	On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew taps the empty seat next to him and asks, "Did you swim in the ocean on your way here?"
p1f	On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew asks, "Did you swim in the ocean on your way here?" Josh and Drew grin.
a3	On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew, who doesn't get along with him, sneers at Josh and asks, "Did you swim in the ocean on your way here?" Drew blocks the empty seat next to him with his feet, and Josh frowns.
a2gr	On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew, who doesn't get along with him, asks, "Did you swim in the ocean on your way here?" Drew blocks the empty seat next to him with his feet.
a2fr	On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew, who doesn't get along with him, sneers at Josh and asks, "Did you swim in the ocean on your way here?" Josh frowns.
a2gf	**On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew sneers at Josh and asks, "Did you swim in the ocean on your way here?" Drew blocks the empty seat next to him with his feet, and Josh frowns.**
a1r	On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew, who doesn't get along with him, asks, "Did you swim in the ocean on your way here?"
a1g	On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew asks, "Did you swim in the ocean on your way here?" Drew blocks the empty seat next to him with his feet.
a1f	On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew sneers at Josh and asks, "Did you swim in the ocean on your way here?" Josh frowns.
amb	On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew asks, "Did you swim in the ocean on your way here?"

Base Tease: Peter arrived late to work and got scolded by his boss. Isaac asks, "So, when are you getting your employee of the month award?"

p3	Peter arrived late to work and got scolded by his boss. Isaac, a coworker who likes Peter, claps Peter's arm and asks, "So, when are you getting your employee of the month award?" They both smile.
p2gr	Peter arrived late to work and got scolded by his boss. Isaac, a coworker who likes Peter, claps Peter's arm and asks, "So, when are you getting your employee of the month award?"
p2fr	Peter arrived late to work and got scolded by his boss. Isaac, a coworker who likes Peter, asks, "So, when are you getting your employee of the month award?" They both smile.
p2gf	Peter arrived late to work and got scolded by his boss. Isaac claps Peter's arm and asks, "So, when are you getting your employee of the month award?" They both smile.
p1r	Peter arrived late to work and got scolded by his boss. Isaac, a coworker who likes Peter, asks, "So, when are you getting your employee of the month award?"
p1g	Peter arrived late to work and got scolded by his boss. Isaac claps Peter's arm and asks, "So, when are you getting your employee of the month award?"
p1f	Peter arrived late to work and got scolded by his boss. Isaac asks, "So, when are you getting your employee of the month award?" They both smile.
a3	Peter arrived late to work and got scolded by his boss. Isaac, a coworker who dislikes Peter, pushes past him and asks, "So, when are you getting your employee of the month award?" Isaac smiles, and Peter does not.
a2gr	Peter arrived late to work and got scolded by his boss. Isaac, a coworker who dislikes Peter, pushes past him and asks, "So, when are you getting your employee of the month award?"
a2fr	Peter arrived late to work and got scolded by his boss. Isaac, a coworker who dislikes Peter, asks, "So, when are you getting your employee of the month award?" Isaac smiles, and Peter does not.
a2gf	Peter arrived late to work and got scolded by his boss. Isaac pushes past him and asks, "So, when are you getting your employee of the month award?" Isaac smiles, and Peter does not.
a1r	**Peter arrived late to work and got scolded by his boss. Isaac, a coworker who dislikes Peter, asks, "So, when are you getting your employee of the month award?"**
a1g	Peter arrived late to work and got scolded by his boss. Isaac past him and asks, "So, when are you getting your employee of the month award?"
a1f	Peter arrived late to work and got scolded by his boss. Isaac asks, "So, when are you getting your employee of the month award?" Isaac smiles, and Peter does not.
amb	Peter arrived late to work and got scolded by his boss. Isaac asks, "So, when are you getting your employee of the month award?"

Base Tease: Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony says, "Thanks for sharing."

p3	Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony, who is always kind to him, chuckles and says, "Thanks for sharing." Tony pats Matt's shoulder, and Matt smiles.
p2gr	Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony, who is always kind to him, pats Matt's shoulder and says, "Thanks for sharing."
p2fr	Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony, who is always kind to him, chuckles and says, "Thanks for sharing." Matt smiles.
p2gf	Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony chuckles and says, "Thanks for sharing." Tony pats Matt's shoulder, and Matt smiles.
p1r	Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony, who is always kind to him, says, "Thanks for sharing."
p1g	Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony pats Matt's shoulder and says, "Thanks for sharing."
p1f	Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony chuckles and says, "Thanks for sharing." Matt smiles.
a3	Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony, who is always horrible to him, scrunches his face up in disgust and says, "Thanks for sharing." Tony moves his chair farther away, and Matt glares.
a2gr	Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony, who is always horrible to him, moves his chair farther away and says, "Thanks for sharing."
a2fr	Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony, who is always horrible to him, scrunches his face up in disgust and says, "Thanks for sharing." Matt glares.
a2gf	Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony scrunches his face up in disgust and says, "Thanks for sharing." Tony moves his chair farther away, and Matt glares.
a1r	Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony, who is always horrible to him, says, "Thanks for sharing."
a1g	Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony moves his chair farther away and says, "Thanks for sharing."
a1f	**Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony scrunches his face up in disgust and says, "Thanks for sharing." Matt glares.**
amb	Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony says, "Thanks for sharing."

Base Tease: David ate a popsicle that made his teeth bright red. Brad sees him and says, "I think you need to meet some friends of mine, Toothbrush and Toothpaste."

p3	David ate a popsicle that made his teeth bright red. His friend, Brad sees him, gestures at his own clean teeth, and says, "I think you need to meet some friends of mine, Toothbrush and Toothpaste." David and Brad laugh.
p2gr	David ate a popsicle that made his teeth bright red. His friend, Brad sees him, gestures at his own clean teeth, and says, "I think you need to meet some friends of mine, Toothbrush and Toothpaste."
p2fr	David ate a popsicle that made his teeth bright red. His friend, Brad sees him and says, "I think you need to meet some friends of mine, Toothbrush and Toothpaste." David and Brad laugh.
p2gf	David ate a popsicle that made his teeth bright red. Brad sees him, gestures at his own clean teeth, and says, "I think you need to meet some friends of mine, Toothbrush and Toothpaste." David and Brad laugh.
p1r	David ate a popsicle that made his teeth bright red. His friend, Brad sees him and says, "I think you need to meet some friends of mine, Toothbrush and Toothpaste."
p1g	David ate a popsicle that made his teeth bright red. Brad sees him, gestures at his own clean teeth, and says, "I think you need to meet some friends of mine, Toothbrush and Toothpaste."
p1f	David ate a popsicle that made his teeth bright red. Brad sees him and says, "I think you need to meet some friends of mine, Toothbrush and Toothpaste." David and Brad laugh.
a3	David ate a popsicle that made his teeth bright red. Brad, who treats him badly, sees him, gestures at David's colored teeth, and says, "I think you need to meet some friends of mine, Toothbrush and Toothpaste." Brad smirks, and David frowns.
a2gr	David ate a popsicle that made his teeth bright red. Brad, who treats him badly, sees him, gestures at David's colored teeth, and says, "I think you need to meet some friends of mine, Toothbrush and Toothpaste."
a2fr	David ate a popsicle that made his teeth bright red. Brad, who treats him badly, sees him and says, "I think you need to meet some friends of mine, Toothbrush and Toothpaste." Brad smirks, and David frowns.
a2gf	David ate a popsicle that made his teeth bright red. Brad sees him, gestures at David's colored teeth, and says, "I think you need to meet some friends of mine, Toothbrush and Toothpaste." Brad smirks, and David frowns.
a1r	David ate a popsicle that made his teeth bright red. Brad, who treats him badly, sees him and says, "I think you need to meet some friends of mine, Toothbrush and Toothpaste."
a1g	David ate a popsicle that made his teeth bright red. Brad sees him, gestures at David's colored teeth, and says, "I think you need to meet some friends of mine, Toothbrush and Toothpaste."
a1f	David ate a popsicle that made his teeth bright red. Brad sees him and says, "I think you need to meet some friends of mine, Toothbrush and Toothpaste." Brad smirks, and David frowns.
amb	**David ate a popsicle that made his teeth bright red. Brad sees him and says, "I think you need to meet some friends of mine, Toothbrush and Toothpaste."**

Base Tease: Luke got a bad grade on a test. Adrian sees and says, "It can't get worse than that."

p3	Luke got a bad grade on a test. Adrian, a classmate who thinks highly of him, sees, smiles, and says, "It can't get worse than that." Adrian hands Luke his own test as Luke laughs.
p2gr	Luke got a bad grade on a test. Adrian, a classmate who thinks highly of him, sees, hands Luke his own test, and says, "It can't get worse than that."
p2fr	Luke got a bad grade on a test. Adrian, a classmate who thinks highly of him, sees, smiles, and says, "It can't get worse than that." Luke laughs.
p2gf	Luke got a bad grade on a test. Adrian sees, smiles, and says, "It can't get worse than that." Adrian hands Luke his own test as Luke laughs.
p1r	Luke got a bad grade on a test. Adrian, a classmate who thinks highly of him, sees and says, "It can't get worse than that."
p1g	Luke got a bad grade on a test. Adrian sees, hands Luke his own test, and says, "It can't get worse than that."
p1f	Luke got a bad grade on a test. Adrian sees, smiles, and says, "It can't get worse than that." Luke laughs.
a3	Luke got a bad grade on a test. Adrian, a classmate who criticizes him, sees, smiles, and says, "It can't get worse than that." Adrian takes the test from Luke's hands as Luke frowns.
a2gr	Luke got a bad grade on a test. Adrian, a classmate who criticizes him, sees, takes the test from Luke's hands, and says, "It can't get worse than that."
a2fr	Luke got a bad grade on a test. Adrian, a classmate who criticizes him, sees, smiles, and says, "It can't get worse than that." Luke frowns.
a2gf	Luke got a bad grade on a test. Adrian sees, smiles, and says, "It can't get worse than that." Adrian takes the test from Luke's hands as Luke frowns.
a1r	Luke got a bad grade on a test. Adrian, a classmate who criticizes him, sees and says, "It can't get worse than that."
a1g	Luke got a bad grade on a test. Adrian sees, takes the test from Luke's hands, and says, "It can't get worse than that."
a1f	Luke got a bad grade on a test. Adrian sees, smiles, and says, "It can't get worse than that." Luke frowns.
amb	**Luke got a bad grade on a test. Adrian sees and says, "It can't get worse than that."**

Appendix D. Measure Versions for Study 2

Version 1 Items

Kyle accidentally got his soda all over his shirt and pants. His friend, Mike, says, “You’re always so smooth.” They grin at each other as Mike gives him a pat on the back.
Alex finished his lunch quickly and ended up going to class with some food on his face. His friend, John, passes Alex the tissue box in the room. John then asks, “Did any of your lunch make it into your mouth?”
Gary pulls off his winter hat, and his hair is messy. His buddy, Adam, grins and says, “Well, it looks like it is Wild Hair Day and someone forgot to tell me.” Gary smiles, too.
Ron got hit in the face by the ball while playing basketball. Cody pats him on the shoulder and says, “Next time, you might want to try catching the ball with your hands instead of with your face.” They share a laugh.
Dan tries to push open a door, but the door is not moving. His friend, Steve, successfully pulls open the door instead and says, “Maybe you should learn to read the signs.”
Ethan came last in a footrace. Max high-fives Ethan as he finishes the race and says, “My mom could beat you in a race.”
Rick stayed up late last night and has dark bags under his eyes. Paul says, “I didn’t know you were a raccoon.” They both laugh.
Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his ears look bigger than they are. Nick, who is always rude to him, sees him and says, “Did your ears grow?”Nick pulls at his own ears and laughs, which makes Sam frown.
Carter came to class with his shirt on backwards. Will, who purposefully gets on his nerves, points at the tag sticking out and says, “Maybe you should get dressed with the lights on next time.”
Greg accidentally wore his little brother’s shirt to class, and it has Mickey Mouse on the front. Brian, a classmate who hates Greg, smirks and says, “I had a shirt just like that when I was in preschool.” Greg scowls.
On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew sneers at Josh and asks, “Did you swim in the ocean on your way here?” Drew blocks the empty seat next to him with his feet, and Josh frowns.
Peter arrived late to work and got scolded by his boss. Isaac, a coworker who dislikes Peter, asks, “So, when are you getting your employee of the month award?”
Mark trips and drops his books. Zack walks by without stopping and asks, “Did you have a nice trip?”
Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony scrunches his face up in disgust and says, “Thanks for sharing.” Matt glares.
David ate a popsicle that made his teeth bright red. Brad sees him and says, “I think you need to meet some friends of mine, Toothbrush and Toothpaste.”
Luke got a bad grade on a test. Adrian sees and says, “It can’t get worse than that.”

Version 2 Items

Luke got a bad grade on a test. Adrian, a classmate who thinks highly of him, sees, smiles, and says, “It can’t get worse than that.” Adrian hands Luke his own test as Luke laughs.
Kyle accidentally got his soda all over his shirt and pants. His friend, Mike, says, “You’re always so smooth.” Mike gives him a pat on the back.
Alex finished his lunch quickly and ended up going to class with some food on his face. His friend, John, grins and asks, “Did any of your lunch make it into your mouth?” Alex smiles back.
Gary pulls off his winter hat, and his hair is messy. Adam grins and says, “Well, it looks like it is Wild Hair Day and someone forgot to tell me.” Adam ruffles his hair, and Gary smiles.
Ron got hit in the face by the ball while playing basketball. His friend, Cody, says, “Next time, you might want to try catching the ball with your hands instead of with your face.”
Dan tries to push open a door, but the door is not moving. Steve successfully pulls open the door instead, holds the door open for him, and says, “Maybe you should learn to read the signs.”
Ethan came last in a footrace. Max says, “My mom could beat you in a race.” They grin at each other.
Rick stayed up late last night and has dark bags under his eyes. Paul, who always tries to pick a fight with Rick, pokes his face and says, “I didn’t know you were a raccoon.” Paul laughs, and Rick scowls.
Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his ears look bigger than they are. Nick, who is always rude to him, sees him and says, “Did your ears grow?” Nick pulls at his own ears.
Carter came to class with his shirt on backwards. Will, who purposefully gets on his nerves, laughs and says, “Maybe you should get dressed with the lights on next time.” Carter glares.
Greg accidentally wore his little brother’s shirt to class, and it has Mickey Mouse on the front. Brian smirks and says, “I had a shirt just like that when I was in preschool.” Brian points at the shirt, making Greg scowl.
On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew, who doesn’t get along with him, asks, “Did you swim in the ocean on your way here?”
Peter arrived late to work and got scolded by his boss. Isaac past him and asks, “So, when are you getting your employee of the month award?”
Mark trips and drops his books. Zack asks, “Did you have a nice trip?” They both glare.
Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony says, “Thanks for sharing.”
David ate a popsicle that made his teeth bright red. Brad sees him and says, “I think you need to meet some friends of mine, Toothbrush and Toothpaste.”

Version 3 Items

David ate a popsicle that made his teeth bright red. His friend, Brad sees him, gestures at his own clean teeth, and says, “I think you need to meet some friends of mine, Toothbrush and Toothpaste.” David and Brad laugh.
Luke got a bad grade on a test. Adrian, a classmate who thinks highly of him, sees, hands Luke his own test, and says, “It can’t get worse than that.”
Kyle accidentally got his soda all over his shirt and pants. His friend, Mike, says, “You’re always so smooth.” They grin at each other.
Alex finished his lunch quickly and ended up going to class with some food on his face. John grins andpasses Alex the tissue box in the room. John then asks, “Did any of your lunch make it into your mouth?” Alex smiles back.
Gary pulls off his winter hat, and his hair is messy. His buddy, Adam, says, “Well, it looks like it is Wild Hair Day and someone forgot to tell me.”
Ron got hit in the face by the ball while playing basketball. Cody pats him on the shoulder and says, “Next time, you might want to try catching the ball with your hands instead of with your face.”
Dan tries to push open a door, but the door is not moving. Steve successfully pulls open the door instead and says, “Maybe you should learn to read the signs.” They snicker over it.
Ethan came last in a footrace. Max, who is rarely nice to him, tries to trip Ethan as he finishes the race and says, “My mom could beat you in a race.” Max smirks at Ethan.
Rick stayed up late last night and has dark bags under his eyes. Paul, who always tries to pick a fight with Rick, pokes his face and says, “I didn’t know you were a raccoon.”
Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his ears look bigger than they are. Nick, who is always rude to him, sees him and says, “Did your ears grow?”Nick laughs, which makes Sam frown.
Carter came to class with his shirt on backwards. Will laughs and points at the tag sticking out. Will then says, “Maybe you should get dressed with the lights on next time.” Carter glares.
Greg accidentally wore his little brother’s shirt to class, and it has Mickey Mouse on the front. Brian, a classmate who hates Greg, says, “I had a shirt just like that when I was in preschool.”
On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew asks, “Did you swim in the ocean on your way here?” Drew blocks the empty seat next to him with his feet.
Peter arrived late to work and got scolded by his boss. Isaac asks, “So, when are you getting your employee of the month award?” Isaac smiles, and Peter does not.
Mark trips and drops his books. Zack asks, “Did you have a nice trip?”
Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony says, “Thanks for sharing.”

Version 4 Items

Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony,who is always kind to him, chuckles and says, “Thanks for sharing.” Tony pats Matt’s shoulder, and Matt smiles.
David ate a popsicle that made his teeth bright red. His friend, Brad sees him, gestures at his own clean teeth, and says, “I think you need to meet some friends of mine, Toothbrush and Toothpaste.”
Luke got a bad grade on a test. Adrian, a classmate who thinks highly of him, sees, smiles, and says, “It can’t get worse than that.” Luke laughs.
Kyle accidentally got his soda all over his shirt and pants. Mike says, “You’re always so smooth.” They grin at each other as Mike gives him a pat on the back.
Alex finished his lunch quickly and ended up going to class with some food on his face. His friend, John, asks, “Did any of your lunch make it into your mouth?”
Gary pulls off his winter hat, and his hair is messy. Adam ruffles his hair and says, “Well, it looks like it is Wild Hair Day and someone forgot to tell me.”
Ron got hit in the face by the ball while playing basketball. Cody says, “Next time, you might want to try catching the ball with your hands instead of with your face.” They share a laugh.
Dan tries to push open a door, but the door is not moving. Steve, a guy who doesn’t like Dan, successfully pulls open the door instead and says, “Maybe you should learn to read the signs.” Steve shuts the door in Dan’s face as Dan glares.
Ethan came last in a footrace. Max, who is rarely nice to him, tries to trip Ethan as he finishes the race and says, “My mom could beat you in a race.”
Rick stayed up late last night and has dark bags under his eyes. Paul, who always tries to pick a fight with Rick, says, “I didn’t know you were a raccoon.” Paul laughs, and Rick scowls.
Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his earslook bigger than they are. Nick sees him and says, “Did your ears grow?” Nick pulls at his own ears and laughs, which makes Sam frown.
Carter came to class with his shirt on backwards. Will, who purposefully gets on his nerves, says, “Maybe you should get dressed with the lights on next time.”
Greg accidentally wore his little brother’s shirt to class, and it has Mickey Mouse on the front. Brian points at the shirt and says, “I had a shirt just like that when I was in preschool.”
On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew sneers at Josh and asks, “Did you swim in the ocean on your way here?” Josh frowns.
Peter arrived late to work and got scolded by his boss. Isaac asks, “So, when are you getting your employee of the month award?”
Mark trips and drops his books. Zack asks, “Did you have a nice trip?”

Version 5 Items

Mark trips and drops his books. His pal, Zack picks up some of the books and asks, “Did you have a nice trip?” They both laugh.
Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony, who is always kind to him, pats Matt’s shoulder and says, “Thanks for sharing.”
David ate a popsicle that made his teeth bright red. His friend, Brad sees him and says, “I think you need to meet some friends of mine, Toothbrush and Toothpaste.” David and Brad laugh.
Luke got a bad grade on a test. Adrian sees, smiles, and says, “It can’t get worse than that.” Adrian hands Luke his own test as Luke laughs.
Kyle accidentally got his soda all over his shirt and pants. His friend, Mike, says, “You’re always so smooth.”
Alex finished his lunch quickly and ended up going to class with some food on his face. John passes Alex the tissue box in the room and asks, “Did any of your lunch make it into your mouth?”
Gary pulls off his winter hat, and his hair is messy. Adam grins and says, “Well, it looks like it is Wild Hair Day and someone forgot to tell me.” Gary smiles.
Ron got hit in the face by the ball while playing basketball. Cody, who doesn’t like him, purposefully knocks into his shoulder and says, “Next time, you might want to try catching the ball with your hands instead of with your face.” Cody laughs, and Ron does not.
Dan tries to push open a door, but the door is not moving. Steve, a guy who doesn’t like Dan, successfully pulls open the door instead and says, “Maybe you should learn to read the signs.” Steve shuts the door in Dan’s face.
Ethan came last in a footrace. Max, who is rarely nice to him, says, “My mom could beat you in a race.” Max smirks at Ethan.
Rick stayed up late last night and has dark bags under his eyes. Paul pokes his face and says, “I didn’t know you were a raccoon.” Paul laughs, and Rick scowls.
Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his ears look bigger than they are. Nick, who is always rude to him, sees him and says, “Did your ears grow?”
Carter came to class with his shirt on backwards. Will points at the tag sticking out and says, “Maybe you should get dressed with the lights on next time.”
Greg accidentally wore his little brother’s shirt to class, and it has Mickey Mouse on the front. Brian smirks and says, “I had a shirt just like that when I was in preschool.” Greg scowls.
On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew asks, “Did you swim in the ocean on your way here?”
Peter arrived late to work and got scolded by his boss. Isaac asks, “So, when are you getting your employee of the month award?”

Version 6 Items

Peter arrived late to work and got scolded by his boss. Isaac, a coworker who likes Peter, claps Peter's arm and asks, "So, when are you getting your employee of the month award?" They both smile.
Mark trips and drops his books. His pal, Zack picks up some of the books and asks, "Did you have a nice trip?"
Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony, who is always kind to him, chuckles and says, "Thanks for sharing." Matt smiles.
David ate a popsicle that made his teeth bright red. Brad sees him, gestures at his own clean teeth,and says, "I think you need to meet some friends of mine, Toothbrush and Toothpaste." David and Brad laugh.
Luke got a bad grade on a test. Adrian, a classmate who thinks highly of him, sees and says, "It can't get worse than that."
Kyle accidentally got his soda all over his shirt and pants. Mike says, "You're always so smooth." Mike gives him a pat on the back.
Alex finished his lunch quickly and ended up going to class with some food on his face. John grins and asks, "Did any of your lunch make it into your mouth?" Alex smiles back.
Gary pulls off his winter hat, and his hair is messy. His bully, Adam, smirks and says, "Well, it looks like it is Wild Hair Day and someone forgot to tell me." Adam steals his hat, and Gary frowns.
Ron got hit in the face by the ball while playing basketball. Cody, who doesn't like him, purposefully knocks into his shoulder and says, "Next time, you might want to try catching the ball with your hands instead of with your face."
Dan tries to push open a door, but the door is not moving. Steve, a guy who doesn't like Dan, successfully pulls open the door instead and says, "Maybe you should learn to read the signs." Dan glares.
Ethan came last in a footrace. Max tries to trip Ethan as he finishes the race and says, "My mom could beat you in a race." Max smirks at Ethan.
Rick stayed up late last night and has dark bags under his eyes. Paul, who always tries to pick a fight with Rick, says, "I didn't know you were a raccoon."
Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his ears look bigger than they are. Nick sees him and says, "Did your ears grow?" Nick pulls at his own ears.
Carter came to class with his shirt on backwards. Will laughs and says, "Maybe you should get dressed with the lights on next time." Carter glares.
Greg accidentally wore his little brother's shirt to class, and it has Mickey Mouse on the front. Brian says, "I had a shirt just like that when I was in preschool."
On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew asks, "Did you swim in the ocean on your way here?"

Version 7 Items

On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew,who gets along with him, taps the empty seat next to him and asks, "Did you swim in the ocean on your way here?" Josh and Drew grin.
Peter arrived late to work and got scolded by his boss. Isaac, a coworker who likes Peter, claps Peter's arm and asks, "So, when are you getting your employee of the month award?"
Mark trips and drops his books. His pal, Zack asks, "Did you have a nice trip?" They both laugh.
Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony chuckles and says, "Thanks for sharing." Tony pats Matt's shoulder, and Matt smiles.
David ate a popsicle that made his teeth bright red. His friend, Brad sees him and says, "I think you need to meet some friends of mine, Toothbrush and Toothpaste."
Luke got a bad grade on a test. Adrian sees, hands Luke his own test, and says, "It can't get worse than that."
Kyle accidentally got his soda all over his shirt and pants. Mike says, "You're always so smooth." They grin at each other.
Alex finished his lunch quickly and ended up going to class with some food on his face. John, who despises him, grins and points at Alex's face. John then asks, "Did any of your lunch make it into your mouth?" Alex glares back.
Gary pulls off his winter hat, and his hair is messy. His bully, Adam, steals his hat and says, "Well, it looks like it is Wild Hair Day and someone forgot to tell me."
Ron got hit in the face by the ball while playing basketball. Cody, who doesn't like him, says, "Next time, you might want to try catching the ball with your hands instead of with your face." Cody laughs, and Ron does not.
Dan tries to push open a door, but the door is not moving. Steve successfully pulls open the door instead and says, "Maybe you should learn to read the signs." Steve shuts the door in Dan's face as Dan glares.
Ethan came last in a footrace. Max, who is rarely nice to him, says, "My mom could beat you in a race."
Rick stayed up late last night and has dark bags under his eyes. Paul pokes his face and says, "I didn't know you were a raccoon."
Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his earslook bigger than they are. Nick sees him and says, "Did your ears grow?" Nick laughs, which makes Sam frown.
Carter came to class with his shirt on backwards. Will says, "Maybe you should get dressed with the lights on next time."
Greg accidentally wore his little brother's shirt to class, and it has Mickey Mouse on the front. Brian says, "I had a shirt just like that when I was in preschool."

Version 8 Items

Greg accidentally wore his little brother's shirt to class, and it has Mickey Mouse on the front. Brian, aclassmate who respects Greg, grins and says, "I had a shirt just like that when I was in preschool." Brian fist-bumps him, making Greg smile.
On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew,who gets along with him, taps the empty seat next to him and asks, "Did you swim in the ocean on your way here?"
Peter arrived late to work and got scolded by his boss. Isaac, a coworker who likes Peter, asks, "So, when are you getting your employee of the month award?" They both smile.
Mark trips and drops his books. Zack picks up some of the books and asks, "Did you have a nice trip?" They both laugh.
Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony, who is always kind to him, says, "Thanks for sharing."
David ate a popsicle that made his teeth bright red. Brad sees him, gestures at his own clean teeth, and says, "I think you need to meet some friends of mine, Toothbrush and Toothpaste."
Luke got a bad grade on a test. Adrian sees, smiles, and says, "It can't get worse than that." Luke laughs.
Kyle accidentally got his soda all over his shirt and pants. His bully, Mike says, "You're always so smooth." They frown at each other as Mike gives him a shove on the back.
Alex finished his lunch quickly and ended up going to class with some food on his face. John, who despises him, points at Alex's face and asks, "Did any of your lunch make it into your mouth?"
Gary pulls off his winter hat, and his hair is messy. His bully, Adam, smirks and says, "Well, it looks like it is Wild Hair Day and someone forgot to tell me." Gary frowns.
Ron got hit in the face by the ball while playing basketball. Cody purposefully knocks into his shoulderand says, "Next time, you might want to try catching the ball with your hands instead of with your face." Cody laughs, and Ron does not.
Dan tries to push open a door, but the door is not moving. Steve, a guy who doesn't like Dan, successfully pulls open the door instead and says, "Maybe you should learn to read the signs."
Ethan came last in a footrace. Max tries to trip Ethan as he finishes the race and says, "My mom could beat you in a race."
Rick stayed up late last night and has dark bags under his eyes. Paul says, "I didn't know you were a raccoon." Paul laughs, and Rick scowls.
Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his ears look bigger than they are. Nick sees him and says, "Did your ears grow?"
Carter came to class with his shirt on backwards. Will says, "Maybe you should get dressed with the lights on next time."

Version 9 Items

Carter came to class with his shirt on backwards. Will, who always jokes around with him, knocks on the front of the backwards shirt and says, “Maybe you should get dressed with the lights on next time.” Carter and Will chuckle.
Greg accidentally wore his little brother’s shirt to class, and it has Mickey Mouse on the front. Brian, aclassmate who respects Greg, fist-bumps him and says, “I had a shirt just like that when I was in preschool.”
On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew, who gets along with him, asks, “Did you swim in the ocean on your way here?” Josh and Drew grin.
Peter arrived late to work and got scolded by his boss. Isaac claps Peter’s arm and asks, “So, when are you getting your employee of the month award?” They both smile.
Mark trips and drops his books. His pal, Zack asks, “Did you have a nice trip?”
Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony pats Matt’s shoulder and says, “Thanks for sharing.”
David ate a popsicle that made his teeth bright red. Brad sees him and says, “I think you need to meet some friends of mine, Toothbrush and Toothpaste.” David and Brad laugh.
Luke got a bad grade on a test. Adrian, a classmate who criticizes him, sees, smiles, and says, “It can’t get worse than that.” Adrian takes the test from Luke’s hands as Luke frowns.
Kyle accidentally got his soda all over his shirt and pants. His bully, Mike gives him a shove on the back and says, “You’re always so smooth.”
Alex finished his lunch quickly and ended up going to class with some food on his face. John, who despises him, grins and asks, “Did any of your lunch make it into your mouth?” Alex glares back.
Gary pulls off his winter hat, and his hair is messy. Adam smirks and says, “Well, it looks like it is Wild Hair Day and someone forgot to tell me.” Adam steals his hat, and Gary frowns.
Ron got hit in the face by the ball while playing basketball. Cody, who doesn’t like him, says, “Next time, you might want to try catching the ball with your hands instead of with your face.”
Dan tries to push open a door, but the door is not moving. Steve successfully pulls open the door instead and says, “Maybe you should learn to read the signs.” Steve shuts the door in Dan’s face.
Ethan came last in a footrace. Max says, “My mom could beat you in a race.” Max smirks at Ethan.
Rick stayed up late last night and has dark bags under his eyes. Paul says, “I didn’t know you were a raccoon.”
Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his ears look bigger than they are. Nick sees him and says, “Did your ears grow?”

Version 10 Items

Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his earslook bigger than they are. Nick, who is always nice to him, sees him and says, “Did your ears grow?” Nick pats him on the back, and they share a smile.
Carter came to class with his shirt on backwards. Will, who always jokes around with him, knocks onthe front of the backwards shirt and says, “Maybe you should get dressed with the lights on next time.”
Greg accidentally wore his little brother’s shirt to class, and it has Mickey Mouse on the front. Brian, aclassmate who respects Greg, grins and says, “I had a shirt just like that when I was in preschool.” Greg smiles back.
On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drewtaps the empty seat next to him and asks, “Did you swim in the ocean on your way here?” Josh and Drew grin.
Peter arrived late to work and got scolded by his boss. Isaac, a coworker who likes Peter, asks, “So, when are you getting your employee of the month award?”
Mark trips and drops his books. Zack picks up some of the books and asks, “Did you have a nice trip?”
Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony chuckles and says, “Thanks for sharing.” Matt smiles.
David ate a popsicle that made his teeth bright red. Brad, who treats him badly, sees him, gestures at David’s colored teeth, and says, “I think you need to meet some friends of mine, Toothbrush andToothpaste.” Brad smirks, and David frowns.
Luke got a bad grade on a test. Adrian, a classmate who criticizes him, sees, takes the test from Luke’s hands, and says, “It can’t get worse than that.”
Kyle accidentally got his soda all over his shirt and pants. His bully, Mike says, “You’re always so smooth.” They frown at each other.
Alex finished his lunch quickly and ended up going to class with some food on his face. John grins andpoints at Alex’s face. John then asks, “Did any of your lunch make it into your mouth?” Alex glares back.
Gary pulls off his winter hat, and his hair is messy. His bully, Adam, says, “Well, it looks like it is Wild Hair Day and someone forgot to tell me.”
Ron got hit in the face by the ball while playing basketball. Cody purposefully knocks into his shoulderand says, “Next time, you might want to try catching the ball with your hands instead of with your face.”
Dan tries to push open a door, but the door is not moving. Steve successfully pulls open the door instead and says, “Maybe you should learn to read the signs.” Dan glares.
Ethan came last in a footrace. Max says, “My mom could beat you in a race.”
Rick stayed up late last night and has dark bags under his eyes. Paul says, “I didn’t know you were a raccoon.”

Version 11 Items

Rick stayed up late last night and has dark bags under his eyes. Paul, who always has his back, hands him a cup of coffee and says, “I didn’t know you were a raccoon.” They both laugh.
Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his earslook bigger than they are. Nick, who is always nice to him, sees him and says, “Did your ears grow?” Nick pats him on the back.
Carter came to class with his shirt on backwards. Will, who always jokes around with him, says, “Maybe you should get dressed with the lights on next time.” Carter and Will chuckle.
Greg accidentally wore his little brother’s shirt to class, and it has Mickey Mouse on the front. Brian,grins and says, “I had a shirt just like that when I was in preschool.” Brian fist-bumps him, making Greg smile.
On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew, who gets along with him, asks, “Did you swim in the ocean on your way here?”
Peter arrived late to work and got scolded by his boss. Isaac claps Peter’s arm and asks, “So, when are you getting your employee of the month award?”
Mark trips and drops his books. Zack asks, “Did you have a nice trip?” They both laugh.
Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony,who is always horrible to him, scrunches his face up in disgust and says, “Thanks for sharing.” Tony moves his chair farther away, and Matt glares.
David ate a popsicle that made his teeth bright red. Brad, who treats him badly, sees him, gestures at David’s colored teeth, and says, “I think you need to meet some friends of mine, Toothbrush and Toothpaste.”
Luke got a bad grade on a test. Adrian, a classmate who criticizes him, sees, smiles, and says, “It can’t get worse than that.” Luke frowns.
Kyle accidentally got his soda all over his shirt and pants. Mike says, “You’re always so smooth.” They frown at each other as Mike gives him a shove on the back.
Alex finished his lunch quickly and ended up going to class with some food on his face. John, who despises him, asks, “Did any of your lunch make it into your mouth?”
Gary pulls off his winter hat, and his hair is messy. Adam steals his hat and says, “Well, it looks like it is Wild Hair Day and someone forgot to tell me.”
Ron got hit in the face by the ball while playing basketball. Cody says, “Next time, you might want to try catching the ball with your hands instead of with your face.” Cody laughs, and Ron does not.
Dan tries to push open a door, but the door is not moving. Steve successfully pulls open the door instead and says, “Maybe you should learn to read the signs.”
Ethan came last in a footrace. Max says, “My mom could beat you in a race.”

Version 12 Items

Ethan came last in a footrace. His friend, Max, high-fives Ethan as he finishes the race and says, “My mom could beat you in a race.” They grin at each other.
Rick stayed up late last night and has dark bags under his eyes. Paul, who always has his back, hands him a cup of coffee and says, “I didn’t know you were a raccoon.”
Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his earslook bigger than they are. Nick, who is always nice to him, sees him and says, “Did your ears grow?” They share a smile.
Carter came to class with his shirt on backwards. Will knocks on the front of the backwards shirt and says, “Maybe you should get dressed with the lights on next time.” Carter and Will chuckle.
Greg accidentally wore his little brother’s shirt to class, and it has Mickey Mouse on the front. Brian, a classmate who respects Greg, says, “I had a shirt just like that when I was in preschool.”
On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew taps the empty seat next to him and asks, “Did you swim in the ocean on your way here?”
Peter arrived late to work and got scolded by his boss. Isaac asks, “So, when are you getting your employee of the month award?” They both smile.
Mark trips and drops his books. Zack, who always tries to upset Mark, walks by without stopping and asks, “Did you have a nice trip?” They both glare.
Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony, who is always horrible to him, moves his chair farther away and says, “Thanks for sharing.”
David ate a popsicle that made his teeth bright red. Brad, who treats him badly, sees him and says, “I think you need to meet some friends of mine, Toothbrush and Toothpaste.” Brad smirks, and David frowns.
Luke got a bad grade on a test. Adrian sees, smiles, and says, “It can’t get worse than that.” Adrian takes the test from Luke’s hands as Luke frowns.
Kyle accidentally got his soda all over his shirt and pants. His bully, Mike says, “You’re always so smooth.”
Alex finished his lunch quickly and ended up going to class with some food on his face. John points at Alex’s face and asks, “Did any of your lunch make it into your mouth?”
Gary pulls off his winter hat, and his hair is messy. Adam smirks and says, “Well, it looks like it is Wild Hair Day and someone forgot to tell me.” Gary frowns.
Ron got hit in the face by the ball while playing basketball. Cody says, “Next time, you might want to try catching the ball with your hands instead of with your face.”
Dan tries to push open a door, but the door is not moving. Steve successfully pulls open the door instead and says, “Maybe you should learn to read the signs.”

Version 13 Items

Dan tries to push open a door, but the door is not moving. His friend, Steve, successfully pulls open the door instead, holds the door open for him, and says, "Maybe you should learn to read the signs." They snicker over it.
Ethan came last in a footrace. His friend, Max, high-fives Ethan as he finishes the race and says, "My mom could beat you in a race."
Rick stayed up late last night and has dark bags under his eyes. Paul, who always has his back, says, "I didn't know you were a raccoon." They both laugh.
Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his earslook bigger than they are. Nick sees him and says, "Did your ears grow?" Nick pats him on the back, and they share a smile.
Carter came to class with his shirt on backwards. Will, who always jokes around with him, says, "Maybe you should get dressed with the lights on next time."
Greg accidentally wore his little brother's shirt to class, and it has Mickey Mouse on the front. Brian fist-bumps him and says, "I had a shirt just like that when I was in preschool."
On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew asks, "Did you swim in the ocean on your way here?" Josh and Drew grin.
Peter arrived late to work and got scolded by his boss. Isaac, a coworker who dislikes Peter, pushes past him and asks, "So, when are you getting your employee of the month award?" Isaac smiles, and Peter does not.
Mark trips and drops his books. Zack, who always tries to upset Mark, walks by without stopping and asks, "Did you have a nice trip?"
Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony,who is always horrible to him, scrunches his face up in disgust and says, "Thanks for sharing." Matt glares.
David ate a popsicle that made his teeth bright red. Brad sees him, gestures at David's colored teeth,and says, "I think you need to meet some friends of mine, Toothbrush and Toothpaste." Brad smirks, and David frowns.
Luke got a bad grade on a test. Adrian, a classmate who criticizes him, sees and says, "It can't get worse than that."
Kyle accidentally got his soda all over his shirt and pants. Mike gives him a shove on the back and says, "You're always so smooth."
Alex finished his lunch quickly and ended up going to class with some food on his face. John grins and asks, "Did any of your lunch make it into your mouth?" Alex glares back.
Gary pulls off his winter hat, and his hair is messy. Adam says, "Well, it looks like it is Wild Hair Day and someone forgot to tell me."
Ron got hit in the face by the ball while playing basketball. Cody says, "Next time, you might want to try catching the ball with your hands instead of with your face."

Version 14 Items

Ron got hit in the face by the ball while playing basketball. His friend, Cody, pats him on the shoulderand says, “Next time, you might want to try catching the ball with your hands instead of with your face.” They share a laugh.
Dan tries to push open a door, but the door is not moving. His friend, Steve, successfully pulls open the door instead, holds the door open for him, and says, “Maybe you should learn to read the signs.”
Ethan came last in a footrace. His friend, Max, says, “My mom could beat you in a race.” They grin at each other.
Rick stayed up late last night and has dark bags under his eyes. Paul hands him a cup of coffee and says, “I didn’t know you were a raccoon.” They both laugh.
Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his ears look bigger than they are. Nick, who is always nice to him, sees him and says, “Did your ears grow?”
Carter came to class with his shirt on backwards. Will knocks on the front of the backwards shirt and says, “Maybe you should get dressed with the lights on next time.”
Greg accidentally wore his little brother’s shirt to class, and it has Mickey Mouse on the front. Brian grins and says, “I had a shirt just like that when I was in preschool.” Greg smiles back.
On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew,who doesn’t get along with him, sneers at Josh and asks, “Did you swim in the ocean on your way here?” Drew blocks the empty seat next to him with his feet, and Josh frowns.
Peter arrived late to work and got scolded by his boss. Isaac, a coworker who dislikes Peter, pushes past him and asks, “So, when are you getting your employee of the month award?”
Mark trips and drops his books. Zack, who always tries to upset Mark, asks, “Did you have a nice trip?” They both glare.
Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony scrunches his face up in disgust and says, “Thanks for sharing.” Tony moves his chair farther away, and Matt glares.
David ate a popsicle that made his teeth bright red. Brad, who treats him badly, sees him and says, “I think you need to meet some friends of mine, Toothbrush and Toothpaste.”
Luke got a bad grade on a test. Adrian sees, takes the test from Luke’s hands, and says, “It can’t get worse than that.”
Kyle accidentally got his soda all over his shirt and pants. Mike says, “You’re always so smooth.” They frown at each other.
Alex finished his lunch quickly and ended up going to class with some food on his face. John asks, “Did any of your lunch make it into your mouth?”
Gary pulls off his winter hat, and his hair is messy. Adam says, “Well, it looks like it is Wild Hair Day and someone forgot to tell me.”

Version 15 Items

Gary pulls off his winter hat, and his hair is messy. His buddy, Adam, grins and says, "Well, it looks like it is Wild Hair Day and someone forgot to tell me." Adam ruffles his hair, and Gary smiles.
Ron got hit in the face by the ball while playing basketball. His friend, Cody, pats him on the shoulderand says, "Next time, you might want to try catching the ball with your hands instead of with your face."
Dan tries to push open a door, but the door is not moving. His friend, Steve, successfully pulls open the door instead and says, "Maybe you should learn to read the signs." They snicker over it.
Ethan came last in a footrace. Max high-fives Ethan as he finishes the race and says, "My mom could beat you in a race." They grin at each other.
Rick stayed up late last night and has dark bags under his eyes. Paul, who always has his back, says, "I didn't know you were a raccoon."
Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his ears look bigger than they are. Nick sees him and says, "Did your ears grow?" Nick pats him on the back.
Carter came to class with his shirt on backwards. Will says, "Maybe you should get dressed with the lights on next time." Carter and Will chuckle.
Greg accidentally wore his little brother's shirt to class, and it has Mickey Mouse on the front. Brian, aclassmate who hates Greg, smirks and says, "I had a shirt just like that when I was in preschool." Brian points at the shirt, making Greg scowl.
On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew,who doesn't get along with him, asks, "Did you swim in the ocean on your way here?" Drew blocks the empty seat next to him with his feet.
Peter arrived late to work and got scolded by his boss. Isaac, a coworker who dislikes Peter, asks, "So, when are you getting your employee of the month award?" Isaac smiles, and Peter does not.
Mark trips and drops his books. Zack, who always tries to upset Mark, asks, "Did you have a nice trip?" They both glare.
Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony, who is always horrible to him, says, "Thanks for sharing."
David ate a popsicle that made his teeth bright red. Brad sees him, gestures at David's colored teeth, and says, "I think you need to meet some friends of mine, Toothbrush and Toothpaste."
Luke got a bad grade on a test. Adrian sees, smiles, and says, "It can't get worse than that." Luke frowns.
Kyle accidentally got his soda all over his shirt and pants. Mike says, "You're always so smooth."
Alex finished his lunch quickly and ended up going to class with some food on his face. John asks, "Did any of your lunch make it into your mouth?"

Version 16 Items

Alex finished his lunch quickly and ended up going to class with some food on his face. His friend, John, grins and passes Alex the tissue box in the room. John then asks, “Did any of your lunch make it into your mouth?” Alex smiles back.
Gary pulls off his winter hat, and his hair is messy. His buddy, Adam, ruffles his hair and says, “Well, it looks like it is Wild Hair Day and someone forgot to tell me.”
Ron got hit in the face by the ball while playing basketball. His friend, Cody, says, “Next time, you might want to try catching the ball with your hands instead of with your face.” They share a laugh.
Dan tries to push open a door, but the door is not moving. Steve successfully pulls open the doorinstead, holds the door open for him, and says, “Maybe you should learn to read the signs.” They snicker over it.
Ethan came last in a footrace. His friend, Max, says, “My mom could beat you in a race.”
Rick stayed up late last night and has dark bags under his eyes. Paul hands him a cup of coffee and says, “I didn’t know you were a raccoon.”
Sam got his hair cut yesterday, but the hair was cut too short around his ears, which makes his ears look bigger than they are. Nick sees him and says, “Did your ears grow?” They share a smile.
Carter came to class with his shirt on backwards. Will, who purposefully gets on his nerves, laughs andpoints at the tag sticking out. Will then says, “Maybe you should get dressed with the lights on next time.” Carter glares.
Greg accidentally wore his little brother’s shirt to class, and it has Mickey Mouse on the front. Brian, aclassmate who hates Greg, points at the shirt and says, “I had a shirt just like that when I was in preschool.”
On a rainy day, Josh forgot his umbrella, so he was soaking wet by the time he made it to class. Drew,who doesn’t get along with him, sneers at Josh and asks, “Did you swim in the ocean on your way here?” Josh frowns.
Peter arrived late to work and got scolded by his boss. Isaac pushes past him and asks, “So, when are you getting your employee of the month award?” Isaac smiles, and Peter does not.
Mark trips and drops his books. Zack, who always tries to upset Mark, asks, “Did you have a nice trip?”
Matt is getting over the flu and is still sneezing and coughing a lot. After Matt sneezes again, Tony moves his chair farther away and says, “Thanks for sharing.”
David ate a popsicle that made his teeth bright red. Brad sees him and says, “I think you need to meet some friends of mine, Toothbrush and Toothpaste.” Brad smirks, and David frowns.
Luke got a bad grade on a test. Adrian sees and says, “It can’t get worse than that.”
Kyle accidentally got his soda all over his shirt and pants. Mike says, “You’re always so smooth.”

Table 29. Selected Prosocial Items with Valence and Accuracy

Scenario of Item	Measure Version	Cue Categories			Mean	Standard Deviation	t-Test: Mean = 0		Accuracy	Standard Deviation	t-Test: Mean = 0.42857	
		Facial Expression	Gesture	Relationship Information			*t*	*p*-value			*t*	*p*-value
5	3	Present			1.03	1.377	4.11	<0.001***	0.73	0.450	3.711	0.001***
11	12		Present		0.87	1.358	3.50	0.002**	0.67	0.479	2.720	0.011*
7	15			Present	1.40	0.932	8.23	<0.001***	0.90	0.305	8.462	<0.001***
2	3	Present	Present		1.43	1.223	6.42	<0.001***	0.87	0.346	6.940	<0.001***
1	3	Present		Present	1.80	0.805	12.25	<0.001***	0.93	0.254	10.897	<0.001***
14	5		Present	Present	0.89	1.595	2.96	0.006**	0.61	0.497	1.900	0.068 [R]
16	2	Present	Present	Present	1.60	1.354	6.47	<0.001***	0.87	0.346	6.940	<0.001***

[R]. Main requirement standard not met, but secondary requirement standard was met (Significant at the 0.10 level)
*. Significant at the 0.05 level
**. Significant at the 0.01 level
***. Significant at the 0.001 level

Table 30. Selected Prosocial Items and Correlations

Scenario of Item	Measure Version	Cue Categories			Item-Item Correlations within Measure Versions							Item-Prosocial Subscale Correlations
					1 Cue			*2 Cues*			*3 Cues*	
		Facial Expression	Gesture	Relationship Information	Facial Expression	Gesture	Relationship Information	Facial Expression and Gesture	Facial Expression and Relationship Information	Gesture and Relationship Information	Facial Expression, Gesture, and Relationship Information	
5	3	Present				0.342	0.553**	0.650***	0.443*		0.337	0.669***
11	12		Present		*PPI*		*PPI*	0.279		0.347	0.530**	0.595** [R]
7	15			Present	0.259	0.302			0.630***	0.267	**C**	0.691***
2	3	Present	Present		0.650***	0.419*			0.681***	*PPI*	0.614***	0.884***
1	3	Present		Present	0.443*		0.802***	0.681***		*PPI*	0.598***	0.710***
14	5		Present	Present		0.424*	0.380*	0.389*	0.345		0.299	0.803***
16	2	Present	Present	Present	0.423*	*PPI*	0.251	0.681***	0.288	0.423*		0.670***

Gray cells represent irrelevant item-item correlations (e.g., the item has no gesture cue, so the 1-cue gesture item is not a meaningful comparison)

*. Significant at the 0.05 level

**. Significant at the 0.01 level

***. Significant at the 0.001 level

PPI. This item performed poorly (i.e., doesn't meet at least two of the requirements: correlations in item-item and item-subscale comparisons below main and secondary standards, below chance-level accuracy, accuracy not significantly different from chance-level accuracy, and valence significantly different from zero and in the expected direction, positive for prosocial and negative for antisocial) and thus was ignored in these analyses/comparisons to not obscure well-performing items.

C. Cannot be computed because this comparison item had zero variance.

Table 31. Selected Antisocial Items with Valence and Accuracy

Scenario of Item	Measure Version	Cue Categories			Mean	Standard Deviation	t-Test: Mean = 0		Accuracy	Standard Deviation	t-Test: Mean = 0.42857	
		Facial Expression	Gesture	Relationship Information			*t*	*p*-value			*t*	*p*-value
9	6	Present			-0.63	1.426	-2.43	0.021*	0.73	0.450	3.711	0.001***
13	1		Present		-1.47	1.008	-7.97	<0.001***	0.87	0.346	6.940	<0.001***
6	7			Present	-1.74	1.125	-8.63	<0.001***	0.90	0.301	8.793	<0.001***
3	9	Present	Present		-0.69	1.561	-2.38	0.024*	0.69	0.471	2.986	0.006**
4	7	Present		Present	-1.84	1.036	-9.88	<0.001***	0.90	0.301	8.793	<0.001***
12	14		Present	Present	-1.96	0.999	-10.40	<0.001***	0.93	0.262	10.088	<0.001***
10	15	Present	Present	Present	-1.83	0.913	-11.00	<0.001***	0.97	0.183	16.143	<0.001***

*. Significant at the 0.05 level
**. Significant at the 0.01 level
***. Significant at the 0.001 level

Table 32. Selected Antisocial Items and Correlations

Scenario of Item	Measure Version	Cue Categories			Item-Item Correlations within Measure Versions							Item-Antisocial Subscale Correlations
					1 Cue			*2 Cues*			*3 Cues*	
		Facial Expression	Gesture	Relationship Information	Facial Expression	Gesture	Relationship Information	Facial Expression and Gesture	Facial Expression and Relationship Information	Gesture and Relationship Information	Facial Expression, Gesture, and Relationship Information	
9	6	Present				0.267	0.641***	*PPI*	0.302		0.429*	0.855***
13	1		Present		0.423*		0.523**	0.473**		0.294	0.523**	0.781***
6	7			Present	*PPI*	0.278			0.631***	0.525**	0.850***	0.764***
3	9	Present	Present		*PPI*	**C**			0.282	0.262	0.406*	0.687***
4	7	Present		Present	0.450*		0.631***	*PPI*		0.850***	0.525**	0.838***
12	14		Present	Present		*PPI*	0.679***	0.283	0.694***		0.694***	0.687***
10	15	Present	Present	Present	0.308	*PPI*	*PPI*	1.000***	1.000***	0.695***		0.796***

Gray cells represent irrelevant item-item correlations (e.g., the item has no gesture cue, so the 1-cue gesture item is not a meaningful comparison)

*. Significant at the 0.05 level

**. Significant at the 0.01 level

***. Significant at the 0.001 level

PPI. This item performed poorly (i.e., doesn't meet at least two of the requirements: correlations in item-item and item-subscale comparisons below main and secondary standards, below chance-level accuracy, accuracy not significantly different from chance-level accuracy, and valence significantly different from zero and in the expected direction, positive for prosocial and negative for antisocial) and thus was ignored in these analyses/comparisons to not obscure well-performing items.

C. Cannot be computed because at least one of the variables is constant.

Table 33. Selection of Ambiguous Items by No Rejection of H_0: Mean = 0 (in Aggregate and Separately)
Selected items are **bolded**.

Scenario of Ambiguous Item	Mean	Standard Deviation	Do not reject H_0: Mean = 0						Requirements Met
			Aggregate		Separate				
					Sample A		Sample B		
			t statistic	*p*-value	*t* statistic	*p*-value	*t* statistic	*p*-value	
1	0.31	1.28	1.84	0.072 [R]	1.88	0.070 [R]	0.63	0.537 [R]	Yes
2	-0.28	1.21	-1.74	0.088 [R]	-2.05	0.050	-0.33	0.745 [R]	No
3	1.17	1.13	7.94	<0.001	5.62	<0.001	5.51	<0.001	No
4	-0.70	1.45	-3.79	<0.001	-3.43	0.002	-2.00	0.054 [R]	No
5	-1.10	1.11	-7.76	<0.001	-5.58	<0.001	-5.32	<0.001	No
6	-1.18	1.26	-7.33	<0.001	-7.04	<0.001	-3.78	0.001	No
7	-0.02	1.49	-0.09	0.931 [R]	1.51	0.142 [R]	-1.69	0.102 [R]	Yes
8	**-0.02**	**1.57**	**-0.08**	**0.934 [R]**	**0.57**	**0.573 [R]**	**-0.72**	**0.477 [R]**	**Yes**
9	0.20	1.36	1.13	0.264 [R]	1.38	0.176 [R]	0.14	0.889 [R]	Yes
10	-0.41	1.26	-2.55	0.013	-2.00	0.055 [R]	-1.58	0.125 [R]	No
11	1.09	1.10	7.54	<0.001	5.33	<0.001	5.25	<0.001	No
12	-0.58	1.44	-3.07	0.003	-3.20	0.003	-1.30	0.203 [R]	No
13	0.00	1.37	0.00	1.000 [R]	-1.00	0.326 [R]	1.16	0.255 [R]	Yes
14	-0.20	1.49	-1.04	0.304 [R]	-1.70	0.100 [R]	0.41	0.682 [R]	Yes
15	**-0.05**	**1.45**	**-0.27**	0.791 [R]	**0.26**	**0.798 [R]**	**-0.60**	**0.550 [R]**	**Yes**
16	-0.59	1.23	-3.69	<0.001	-2.25	0.032	-2.91	0.007	No

[R]. Requirement met: p-value>0.05

Table 34. Selection of Ambiguous Items by Skewness Requirement

The positive-negative skew in the 95% confidence interval value was computed by comparing the proportion of the range of the 95% confidence interval of the mean difference from zero (when an item's two samples were aggregated) that was positive (above zero) to the proportion of the range that was negative (below zero). This was computed to be a helpful description of the skewness in the item's evaluation, wherein, if the value was closer to zero, then the evaluation of the item demonstrated minimal skew. If the value was closer to 1 (-1), then the evaluation of the item tended to be more positive and prosocial (more negative and antisocial). Selected items are **bolded**.

Scenario of Ambiguous Item	Mean	Standard Deviation	Skewness					Requirements Met
			Statistic	Standard Error	95% Confidence Interval of Difference of Mean from Zero (Aggregate)		Positive-Negative Skew in 95% Confidence Interval	
					Lower Limit	Upper Limit		
1	0.31	1.28	0.119 [R1]	0.311	-0.03 [R2]	0.64 [R2]	0.91	No
2	-0.28	1.21	0.065 [R1]	0.314	-0.59 [R2]	0.04 [R2]	-0.87	No
3	1.17	1.13	-0.789	0.311	0.87	1.46	1.00	No
4	-0.70	1.45	0.541	0.306	-1.08	-0.33	-1.00	No
5	-1.10	1.11	0.811	0.306	-1.38	-0.82	-1.00	No
6	-1.18	1.26	0.871	0.306	-1.5	-0.86	-1.00	No
7	-0.02	1.49	-0.164 [R1]	0.311	-0.41 [R2]	0.37 [R2]	-0.05 [R3]	Yes
8	**-0.02**	**1.57**	**0.195 [R1]**	**0.311**	**-0.43 [R2]**	**0.39 [R2]**	**-0.05 [R3]**	**Yes**
9	0.20	1.36	-0.206 [R1]	0.306	-0.15 [R2]	0.55 [R2]	0.57	No
10	-0.41	1.26	0.154 [R1]	0.306	-0.73	-0.09	-1.00	No
11	1.09	1.10	-0.837	0.314	0.80	1.37	1.00	No
12	-0.58	1.44	0.355 [R1]	0.311	-0.95	-0.20	-1.00	No
13	0.00	1.37	-0.081 [R1]	0.306	-0.35 [R2]	0.35 [R2]	0.00 [R3]	Yes
14	-0.20	1.49	0.071 [R1]	0.309	-0.59 [R2]	0.19 [R2]	-0.51	No
15	**-0.05**	**1.45**	**0.021 [R1]**	**0.309**	**-0.43 [R2]**	**0.33 [R2]**	**-0.13 [R3]**	**Yes**
16	-0.59	1.23	0.022 [R1]	0.311	-0.91	-0.27	-1.00	No

[R1]. Requirement 1 met: Skewness statistic is between 0.50 and -0.50.

[R2]. Requirement 2 met: 95% confidence interval of the mean difference from zero contains zero.

[R3]. Requirement 3 met: Positive-negative skew in the 95% confidence interval between 0.40 and -0.40.