IRONY COMPREHENSION IN CHILDREN WITH ASD:

DOES CONVENTIONALITY MAKE A DIFFERENCE?

Hannah Carlgren

School of Family Studies and Human Services
Program in Communication Sciences and Disorders
Kansas State University

December 7, 2017

University Honors Program Project

Debra L. Burnett

Dr. Debra L. Burnett, Supervisor

Abstract

Irony, a non-literal language convention in which the meaning of the statement and what is said are contradictory, is employed and practiced in the English language regularly. To interpret an ironic statement accurately, an individual must be able to infer the speaker's meaning, attitude and intent. This type of interpretation is potentially one of the most difficult tasks to develop for children diagnosed with Autism Spectrum Disorder (ASD), as a core deficit of individuals with ASD is perspective taking skills as well as deficits in pragmatic communication. The current study explored the acquisition of irony for children ages eight to eleven years with ASD in comparison to their typically developing same-aged peers, and their peers with language impairments. It will also examine how the conventionality of the ironic statement influences the child's comprehension skills. The study presented 22 short vignettes to participants via auditory or written modalities. These vignettes contained remarks that were either ironic or literal in nature, and utilized both conventional and situation-specific remarks. Following presentation of the vignette, participants were asked a series of questions to assess their comprehension of speaker meaning, attitude, and intent. Results from the sample of 5 participants indicated that there were no significant differences between the two types of remarks. Implications for those findings are discussed.

Introduction

Individuals with Autism Spectrum Disorder (ASD) are diagnosed based on deficits and excesses in three main areas; communication, social interaction, and behavior (National Institute of Mental Health, 2016; American Psychiatric Association, 2013). Although ASD presents itself differently with each individual, some common behaviors that are evident include rigidity, limited interests, and repetitive behaviors (National Institute of Mental Health, 2016; American Psychiatric Association, 2013). In addition, research on children with ASD indicates distinct deficits in pragmatic communication and perspective taking abilities that are necessary to comprehend verbal irony (Pexman et al., 2011). Previous studies have examined the deficits of non-literal language comprehension for children with ASD, however few have studied how these children best understand and process non-literal remarks. The purpose of the current study is to investigate how children with Autism Spectrum Disorder (ASD) are able to comprehend irony as compared to prior studies of typically developing, same-aged peers and children with language impairments. Of particular interest is how the conventionality of the ironic statement used affects comprehension when comparing performance among these three groups. Previous studies have found that typically developing children begin to possess the skills necessary to comprehend irony by the age of five or six (Glenwright & Pexman, 2010; Pexman et al., 2011). The conventionality of a statement will have an impact on the participant's ability to correctly comprehend the utterance. The graded salience hypothesis states that more salient meanings are processed initially before less salient meanings are activated (Giora, 1997). Saliency refers to when the meaning of the word/expression can be retrieved directly from the lexicon (Giora, Fein, & Schwartz, 1998); and the lexicon is referring to one's mental dictionary, including word definitions and associations (Burnett, 2015). Previous studies offer support of the graded salience

hypothesis and have analyzed the processing required for both ironic and literal statements. A study conducted by Giora found that when interpreting irony, one must process both the literal meaning and its non-literal counterpart. The literal meaning is processed initially, whereas the ironic (less salient) meaning is processed only after the initial processing is completed. Non-literal processing requires the use of sequential processing, which makes ironic interpretation more difficult and time-consuming than literal interpretation (Giora, 1997). Thus, if a phrase's meaning is salient, it is possible that it could decrease the challenge of comprehension for the listener (Giora, Fein, & Schwartz, 1998).

The current study on irony comprehension for children with ASD utilized both conventional (more salient) and novel/ situation-specific remarks (less salient). Keeping in mind prior findings, researchers expected that participants would be able to comprehend conventional remarks with more ease than the novel/situation-specific remarks (Giora, Fein, & Schwartz, 1998). Due to the lack of information in the field, this research may offer insight for practitioners to better understand these individual's difficulty with irony in discourse. It also may allow clinicians to tailor intervention techniques and formulate/improve communication strategies for children with ASD. That being said, more research must be done before deriving conclusions to apply this knowledge to clinical populations. In the current study, participants were presented with short vignettes through an auditory modality or through a written modality. The vignettes featured a character, Pat, who made remarks that were either literal or non-literal in nature. At completion of the vignette, the children responded to a series of audio-recorded questions in order to assess their understanding of the events that occurred as well as speaker meaning, attitude, and intent.

Methods

Participants

Five English-speaking children ages eight to eleven years (8;2 - 11;3 years) (*M*=10;0 years, *SD*=16 months) with a diagnosis of Autism participated in this study. All of the participants for the current study were males. This age-range was the target of the current study due to previous findings suggesting this was when children were likely to have emerging skills for inferring speaker meaning and intent for ironic remarks (Burnett, 2015). The participants were recruited via advertisements in internal university news announcements at Kansas State. Informational flyers were also provided to professionals who provide services to children with ASD including autism specialists, speech-language pathologists, the USD 383 Autism Coordinator, and in public places around the greater Manhattan, Kansas area. At no time did the professionals provide contact information for any parent/family to whom they shared information, all participants contacted the researchers if interested. The socioeconomic status (SES) of the participants, which was measured using the Hollingshead scale (Hollingshead, 1975), averaged Upper Middle Class (M=61.6, SD=17.7, range=38-77). Eligibility requirements for inclusion in the study were as follows: must have a diagnosis of ASD, be between ages 7;0-11;0 years old (or approximately grades 2-4), and be a monolingual English speaker. Participants were excluded if they achieved scores greater than 2 ½ standard deviations below the mean for same-aged peers, or if they were reading more than 2 grade levels below the stimulus stories, as they would be unable to complete the experimental task. The participants also had to have hearing within the normal range in order to hear the stimulus and questions presented via auditory modality. Hearing status was indicated by the participant's parent via a consent form, which contained demographic questions pertaining to education, occupation, reading levels,

and other areas related to the research task (e.g. sensory needs, allergies). In addition, parents completed the parent/caregiver questionnaire from the Childhood Autism Rating Scale-2nd Edition (CARS-2; Schopler, Van Bourgondien, Wellman, & Love, 2010) to assist in verifying the ASD diagnosis and assessing current skill levels. Researchers also completed the standard version of the CARS-2 following the session to measure severity of current symptoms of ASD under the direct supervision of the principal investigator. The following T-scores were obtained for the current sample of participants using the following severity scale: 15-29 as minimal-to-no symptoms of ASD (n=2), 30-36.5 for mild-to-moderate symptoms of ASD (n=1), and 37+ as severe symptoms of ASD (n=2). This yielded an average T-score of 34.6 (SD=16.06, Range = <20-54). Researchers did not collect exact IQ scores, but each participant's parent reported that the participant's IQ was measured at school and was within the average range; therefore there was no concern for intellectual disability interfering with the statistical analysis. In addition to filling out the report, the participant's caregiver signed a consent form and the participant gave verbal assent before researchers began conducting the experiment. At completion of the session, the participants were compensated $20 for the session. The study was conducted in full compliance with the Institutional Review Board at Kansas State University.

Experimental task materials

The task itself was comprised of story contexts/vignettes involving a gender-neutral child, Pat, that were utilized in previous research regarding irony comprehension with permission from the principle investigator (Burnett, 2015). Keeping the story-lines consistent across experiments aided with data consistency when comparing with comprehension studies of language-impaired and typically developing children. The study conducted by Burnett in 2015 discussed the versions of the experimenter-generated stories and how they were adapted to meet criteria. For

instance, each study was rated according to the degree of negativity it possessed in efforts to confirm that negative and positive stories were had sufficiently different outcomes (Burnett, 2015). Although previous studies utilized remarks that were strongly negative, neutral, and positive, the current study included only strongly positive and strongly negative remarks. This remark was the final sentence in the vignette, and each vignette was approximately four sentences in length. The Flesch-Kincaid (measuring grade level) reading level averaged 2.77 for the Auditory set and 1.35 for the Written set. Overall, the average Flesch-Kincaid for the 22 stories was a 2.09, with the minimum level being 0.30 and the maximum difficulty being 4.0. Of the 11 total vignettes in each group, six were used for data collection (J.P. Kincaid, R.P. Fishburne, R.L. Rogers, & B.S. Chissom (1975). Three filler stories (i.e. positive contexts with ironic remarks, negative contexts with literal criticisms, and neutral contexts and remarks) and two practice stories were also included throughout the experimental task (Burnett, 2015).

Every vignette was accompanied by two black and white illustrations that visually represented the main events of the vignette. These illustrations were the same that were utilized in Burnett's prior research with permission from the researcher (Burnett, 2015). These illustrations were simple; containing only relevant aspects of the events mentioned in the vignette and with no visible facial expressions. The current study did not want more complex illustrations nor facial expressions to aide in terms of comprehension strategies, as this was not the focus of the current study. The two illustrations were presented to the child prior to the vignette being presented, and were present throughout the entire task.

Open-ended questions were given for each vignette to assess comprehension of speaker meaning, intent, and attitude. The current study utilized the same questions presented in prior research of this nature with permission from the researcher (Burnett, 2015). Previous research

indicated that children tend to repeat the answer from speaker meaning to the speaker intent question. Thus, a forced-choice intent probe was also utilized with a yes/no possible response in efforts to clearly determine comprehension of speaker intent (Burnett, 2015). These questions, along with the vignettes, were audio-recorded for presentation to the participants by a female speaker using Praat software. The same recordings that were utilized in Burnett's 2015 study were given to participants in the current study to stay consistent (Burnett, 2015).

Procedure

The experimental task consisted of listening to or reading short vignettes. Participants were seated at a table across from the investigator, and were presented with two illustrations accompanying each vignette prior to the presentation of the vignette itself. These illustrations remained visible until the child had completed answering the entire set of questions for the specific vignette. There were 22 total vignettes that were divided into two modality groups: auditory and written. Each modality set had 11 short vignettes. Two practice vignettes were initially presented at the beginning of each set to familiarize the child with the procedure and to ensure that participants were comprehending the format. Each set also contained three filler stories alongside the experimental stories. There were no story context repetitions within participants (e.g. each participant heard only one version of the "Spill" story) in efforts to stay consistent with prior research (Burnett, 2015). The order of presentation for the two sets and vignettes was randomized for every participant, with an exception of the practice vignettes which were presented initially for each context. After the vignette, the investigator played audio-recorded follow-up questions regarding what took place in the story, speaker meaning, affect, and intent in efforts to assess participant's comprehension. These questions were presented in a fixed order of presentation in efforts to stay consistent with previous research (Burnett, 2015).

Each session was administered individually, during one session that was approximately one-hour long. The sessions were all conducted at the K-State Speech and Hearing Center in Campus Creek Complex. Sessions were recorded via a wall-mounted camera and a recording system as well as a video camera positioned on a tripod. Sessions activities were carried out by an undergraduate student under direct supervision of the principal investigator. Parents could observe the testing session through a one-way mirror from an observation room. Children were given breaks as needed to minimize fatigue and to maintain attention to the tasks at hand.

Scoring and reliability

Participant's responses to the experimental questions were scored based off the video/audio-recorded sessions as well as the researcher's data collected during the session. For each comprehension question, the participants received a score of 1 (correct response) or a score of 0 (incorrect response) as follows:

For the fact question, the participant received a score of one if they could provide the main aspects of the vignette. They received a score of 0 if they did not comprehend the story, or if there were main points missing that would impede the answers to the rest of the questions. For the speaker meaning question, the participant received a score of 1 if they took the utterance that was said by the speaker and portrayed what was meant by the utterance (thus, for ironic remarks they provided the opposite meaning). For example, if the speaker said, "nice job" ironically, the researchers expect that the meaning statement would include something about how it was not actually a nice job. The participant received a score of 0 if they did not refer to the comment, but rather how they were wanting someone to feel or the intent. It is noted that for scoring purposes, the participants occasionally discussed the meaning during another question. Researchers deemed that the participant should receive a score of 1 for the speaker's meaning question if the

client discussed the meaning correctly at any point throughout questioning. For instance, if the child doesn't mention it when talking about meaning but in the intent question refers to "nice job" as meaning "not a nice job" then he received a score of 1. For the speaker attitude question, the participant received a score of 1 if they conveyed how Pat was feeling accurately for the situation. They received a score of 0 if the feeling they stated did not match an appropriate emotion for how Pat should feel based upon the intent of the story. For the speaker intent question, the participant received a score of 1 if they stated why the main character would say the remark to the other character in the story and/or how they wanted to make the other person feel. They also received a score of 1 if they mentioned that the main character was being sarcastic. The participant received a score of 0 if they stated what was meant by the utterance rather than about how the main character was intending to make the other person feel by saying the utterance. For the forced choice question, the participant received a score of 1 if they answered the yes/no question correctly. They received a score of 0 if they did not answer the question correctly. They received a score of N/A (which was recorded as a 0 for data analysis purposes) if the child answered "maybe," "kind of," or said "I don't know."

The undergraduate investigator who conducted the session assigned the initial score, however each undergraduate investigator scored responses for every participant and discussed if there were any discrepancies in scoring methods. 20% of participant's responses were also scored from video-recorded sessions by a third scorer who was blind to the research questions but trained in standardized testing administration.

Results

Possible scores for each story type (e.g., written modality and novel/situation-specific remark, auditory modality and conventional remark) ranged from 0 to 3. Possible scores for the modality condition (written vs. auditory) ranged from 0-6 (see Table 1 for descriptive statistics). No significant differences (α=.05) were found using two-tailed paired *t*-tests based on the conventionality of the statement or modality of presentation, as it relates to the participant's ability to interpret speaker meaning, attitude or intent.

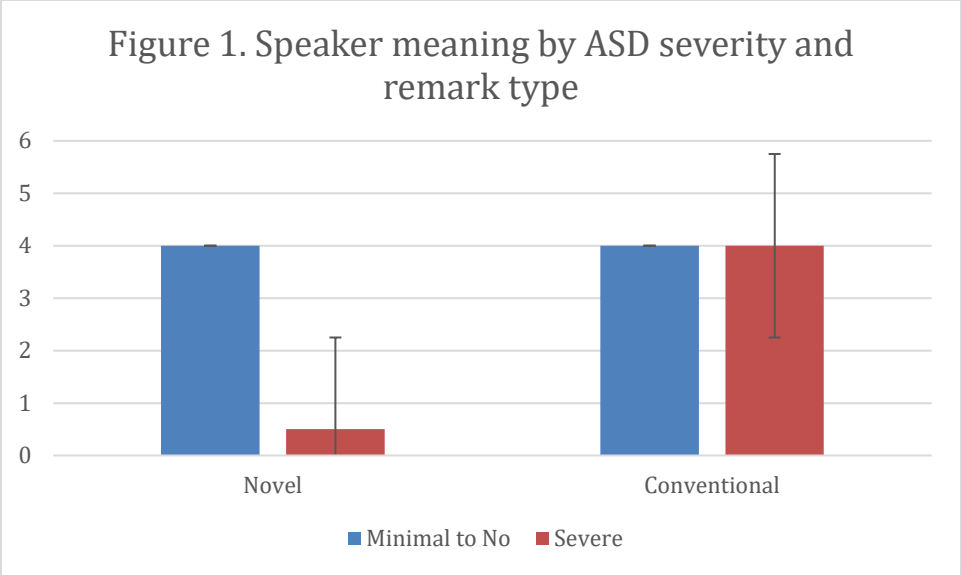
Table 1. Average scores by modality and remark type (possible scores for each range from 0-6).

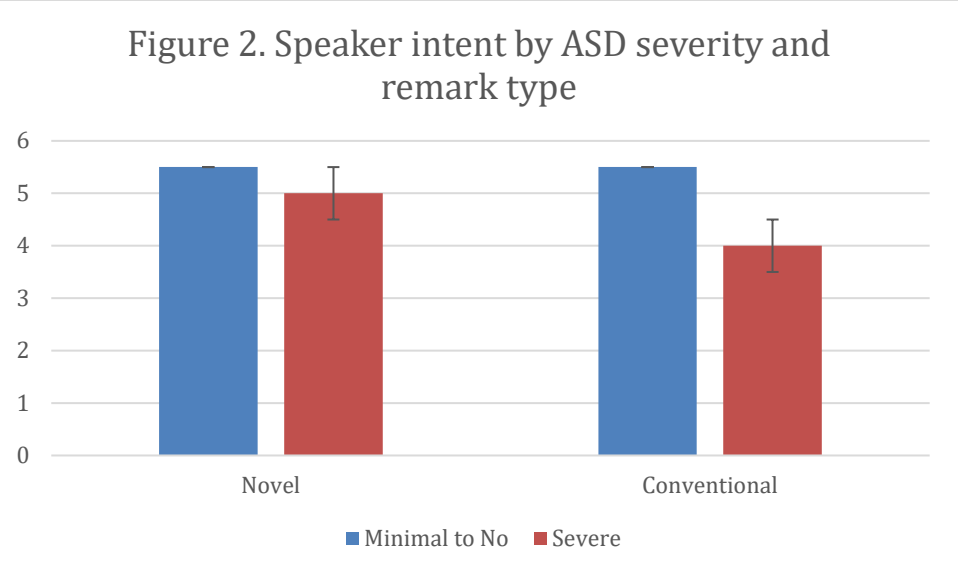
Question Type		Story Modality		Remark Type	
		Written	Auditory	Novel	Conventional
Meaning	*M*	3.0	3.0	2.2	3.6
	SD	2.0	1.73	2.28	1.67
Affect	*M*	5.8	5.8	5.8	5.8
	SD	0.45	0.45	0.045	0.045
Intent	*M*	4.2	4.8	4.6	4.2
	SD	1.48	1.64	1.67	1.48

For speaker meaning, results for each participant indicated that some performed equally across remark type whereas others performed better for the conventional remarks. For speaker affect, there was a ceiling effect in that most children were able to determine if the speaker was feeling a positive or negative emotion. For speaker intent, there were two possible ways that the child could answer the question. An open-ended question followed by a forced choice question

allowed the child to indicate if the speaker, i.e. Pat, wanted to make the listener feel good or bad about what had occurred. Children were able to perform equally for both remark types for speaker intent. Further exploration of the role of question type and the performance of the participants, particularly with the ability of children with ASD to answer various types of questions, would be helpful in future research.

Further exploration of the role of severity of ASD diagnosis and the performance of the participants was conducted. While there were no statistically significant differences between the performance of the two children with minimal to no symptoms (according to CARS-2) as compared to the two children with severe symptoms, there does appear to be some potential for further study in the area of speaker meaning. For inferring meaning, children with severe symptoms of ASD may demonstrate much more difficulty than children with minimal symptoms when encountering novel/situation-specific remarks. When encountering conventionally used phrases, there may be little distinction related to severity if ASD (see Figure 1). In the area of speaker intent, there may be a role of symptom severity, but current findings did not indicate that it would be an important factor (see Figure 2).





Discussion

This study examined if the conventionality of a statement had an impact on comprehension ability for children with ASD. Keeping in mind previous findings regarding the graded salience hypothesis, it was hypothesized that participants in this study would be able to comprehend conventional remarks with more ease than novel/situation-specific remarks (Giora, Fein, & Schwartz, 1998). After analyzing results, researchers failed to reject the null hypothesis.

Current findings were compared to prior research by Burnett (2015) for children with typical language skills, ages 7 and 8 years, and (2014) for children with specific language impairment. Comparisons are made by examining the number of correct responses since the number of total possible responses varied slightly between studies. The children with specific language impairment averaged 2.5 correct responses for speaker meaning given conventional remarks and 2 correct responses given novel remarks. Children with typical language achieved 4.5 and 4.16, respectively. The children with ASD scored between those two groups for conventional remarks with an average of 3.6 correct responses. However, for novel/situation-specific remarks, they appeared more like the children with language impairment with an average of 2.2 correct responses. These comparisons signal an important role for remark type when exploring the ability of children with ASD and children with language impairment. They also support further evidence that the graded salience hypothesis is a potential framework for understanding how children develop these skills and for potential therapy strategies.

Although results for the current study were not statically significant, it should be noted that there are several limitations to the study may account for the lack of significance. For instance, this study had a small sample size (n=5) making results not as easily generalizable. The participants were also variable in terms of the severity of their ASD diagnosis, as determined by their scores on the CARS-2. In addition, their comprehension abilities differed greatly; with one participant being consistently correct throughout a majority of the experimental tasks and other participants being quite limited in their performance abilities. Due to this variability along with the small sample size, it was difficult to determine a pattern among results. Participant's behavior may have affected the reliability of results for several children as well, as some children struggled to maintain attention throughout all experimental tasks. Determining if the participants

had received prior training in ironic comprehension strategies was also not controlled by researchers, which may have contributed to some participant's relative success in comparison to those who had received no prior training. Based on these findings, it is inconclusive as to if conventionality plays a role in irony comprehension for children with ASD. Further research must be conducted with more selective inclusionary criteria and a larger sample size in order for the role of these factors in ironic comprehension to be determined effectively. Future studies also may want to consider the impact that behaviors have on the overall score, as client's with more severe ASD symptoms had difficulty with maintaining attention to the experimental task. The scores reflect this, as the participant with the least severe symptoms also had the most success with the experimental task. Although no direct correlation can be derived from this, researchers do believe this phenomenon should be noted as future studies could explore this further. Overall, future research is needed before determining any clinical implications. If results were found to be statistically significant, it could impact the course of therapy and training techniques for teaching children with ASD how to correctly interpret ironic utterances. For instance, if the null hypothesis would have been rejected, suggesting that literal remarks were more easily comprehended, clinicians would need to be mindful as to how they present ironic statements to individuals with ASD (i.e. presenting less salient remarks during initial stages of therapy for more effective treatment). Researchers do believe that further research would be beneficial so that clinicians are able to use the most effective treatment options and aide in the comprehension of non-literal language techniques for children with ASD.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th edition). Washington, DC.

Burnett, D. (2014). Exploring the role of conventionality for irony comprehension in children with and without language impairment. Poster presented at the Symposium for Research in Child Language Disorders, Madison, WI.

Burnett, D. (2015). Exploring the role of conventionality in children's interpretation of ironic remarks. *Journal of Child Language,* 43, pp 1267-1288.

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.

Giora R., 1997. Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics*, 7 , 183-206.

Giora R., Fein, O. & Schwartz, T. (1998). Irony: graded salience and indirect negation. *Metaphor and Symbol*, 13, 83-101.

Giora R. & Fein, O. (1999). Irony: context and salience. *Metaphor and Symbol*, 14, 241-257

Glenwright, M., & Pexman, P. M. (2010). Development of children's ability to distinguish sarcasm and verbal irony*. *Journal of Child Language, 37*(2), 429-51.

Hollingshead A. B. (1975). Four factor index of social status. Unpublished manuscript, Department of Sociology, Yale University, New Haven, CT.

J.P. Kincaid, R.P. Fishburne, R.L. Rogers, & B.S. Chissom (1975). Derivation of new readability formulas (Automated Readability Index, Fog Count, and Flesch Reading Ease Formula) for navy enlisted personnel, Research Branch Report No. 8-75. Naval Air Station, Memphis, TN

National Institute of Mental Health."Autism Spectrum Disorder." *U.S National Library of Medicine*. U.S. National Library of Medicine, Oct. 2016. Web. 23 Nov. 2016.

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. dx.doi.org.er.lib.k-state.edu/10.1007/s10803-010-1131-7

Appendix A

General Comprehension Question Format (Burnett, 2015)

Fact question: *Can you tell me what happened in the story?*

Speaker meaning: *What did Pat mean by [target remark]?*

Speaker attitude: *How did Pat feel when [even outcome]?*

Speaker intent open-ended: *Why did Pat say, [target remark]?*

Speaker intent forced-choice: *Did Pat want to make Pat's friend feel [bad/good]?*

Example story context, target remarks, comprehension questions, and illustrations for the 'Spill' story (Burnett, 2015)

Strongly negative context + Conventional ironic remark

Pat asks a friend to carry a vase of water to the table.

Pat's friend is not being very careful and slips.

The vase breaks and water spills everywhere.

Pat says, "Smart move."

Comprehension questions

Fact question: Can you tell me what happened in the story?

Speaker meaning: What did Pat mean by "Smart move"?

Speaker attitude: How did Pat feel when Pat's friend spilled the water?

Speaker intent open-ended: Why did Pat say, "Smart move"?

Speaker intent forced-choice: Did Pat want to make Pat's friend feel bad?

Illustration Accompanying Spill Story