The development of children's comprehension and appreciation of riddles

Harry R. M. Purser[1], Jo Van Herwegen[2], & Michael S. C. Thomas[3]

[1]Division of Psychology, Nottingham Trent University

[2]Institute of Education, University College London, University of London

[3]Developmental Neurocognition Laboratory, Birkbeck College, University of London

Running head: Development of riddle comprehension

Word count: 5299

E-mail: harry.purser@ntu.ac.uk

The authors have no conflict of interest to declare

Acknowledgements

We would like to thank all of the children that took part in our research as well as Amy Rolls for her help with coding. This research was funded by The British Academy and the Leverhulme Foundation, grant number SG161069.

Abstract

Humor appreciation and understanding is important for children's social relationships. The current study examined the associations between riddle comprehension, riddle appreciation, and smiling/laughter in children from a wide age-range aged 4 to 11 years old, as well as how cognitive processing style relates to riddle comprehension. Style was distinguished between local and global language processing at the sentence level. The results showed that only children above the age of 8 years old showed a reliable relationship between humor comprehension and smiling/laughter. These findings show that laughter should not be taken as an automatic indicator of explicit understanding. In addition, higher vocabulary ability was independently associated with better humor comprehension. This demonstrates a separable role language proficiency in humor comprehension and suggests avenues for future research in atypical populations known to have difficulties in this area.

Humor provides a window on the interface of cognition, language, and social functioning and is an under-explored avenue of investigation in developmental psychology: children's social, cognitive, and linguistic development have all been linked to humor development (Ely & McCabe, 1994; Semrud-Clikeman & Glass, 2008; Southam, 2005). Context appears to play an important role, in the sense that jokes only occur when the context is safe or non-threatening (see discussion in Degabriele & Walsh, 2010). In addition, one also needs understanding of the context *per se* (Mitchell, Graesser & Louwerse, 2010), because the incongruity is often presented at the end of joke and thus, in order to understand the joke, information given at beginning of joke must be integrated with information presented at the end. Weak Central Coherence (WCC) describes the tendency (which need not reflect an impairment; Happé & Frith, 2006) not to integrate piecemeal information into a coherent whole and thus WCC would be expected to impair joke understanding in typically-developing children, just as it has been found to impair language understanding in autism spectrum disorders (Happé, 1997; Jolliffe & Baron-Cohen, 2000).

Comprehension is only part of the psychological response to humor, however: appreciation and enjoyment have also been formally linked to humor. Some older literature partially addressed questions of how comprehension, appreciation, and enjoyment are related developmentally, leaving some basic questions open. Some 50 years ago, Zigler, Levine, and Gould (1966) administered a 'mirth response test' to children in grades two to five (mean ages 7.5 to 10.5). Children were shown cartoons taken from magazines and three measures were taken: whether the child found the cartoon funny, a facial mirth score, and a metric of comprehension. No developmental trend was apparent for finding the cartoons funny, whereas comprehension increased with age. The pattern for enjoyment was more complex, with a marked drop in facial mirth between grades 4 and 5 (age 9-11), which the authors attributed to a lack of satisfaction in understanding the cartoons.

Subsequent experimental research involving measures of both comprehension and appreciation or enjoyment has tended to focus on verbal riddles. Whitt and Prentice (1977) investigated the riddle appreciation and comprehension of first, third, and fifth graders (average ages 7.4, 9.6, and 11.2, respectively), noting that comprehension of homonymic riddles (those based on words/phrases with different meanings sounding the same, e.g., "Why are fish smart? They go around in schools") rose with age while self-reported appreciation fell. In addition, Bruno, Johnson, and Simon (1987) administered a joke completion task in which participants chose between humorous-incongruous and factual endings to children in grades 1-3 (7-10 years), 4-5 (10-13 years), and 8 (13-16 years), asking them also to explain the jokes. Older children tended to select the humorous endings more often than younger children and mostly gave good explanations for those jokes for which they selected the 'correct' humorous ending. These studies show that humor appreciation and understanding change with age.

Although there is a paucity of experimental research into children's riddle comprehension, results from the related area of irony comprehension tell a somewhat different story. In a large study, Dews and colleagues (Dews et al., 1996) found a sharp rise in comprehension of irony between the ages of five and six years, with greater subjective ratings of the humor of ironic criticism associated with older children (aged up to 9), in contrast to the age-related decline in appreciation seen for riddles in the literature discussed above. The authors noted that irony may present a particular difficulty for comprehension because the listener must understand both the speaker's beliefs *and* the speaker's beliefs about the beliefs of the listener (Winner, 1988; Winner & Gardner, 1993). One way of reconciling age-related increases in appreciation of irony with decreases in appreciation of riddles is Masten's (1989) observation that children prefer humor that challenges them and is not obvious: irony can be harder to understand than simple homophonic riddles.

The results of the experimental studies discussed above should also be considered in the light of observational studies, which converge on the notion that children can successfully produce riddles, amongst other forms of figurative language, from as young as 5- or 6-years-old. Observation of dyads of preschool children (Esposito, 1980; Garvey, 1977) have found that a small but significant proportion of speech by very young children (aged three to five) involved play with sounds. This linguistic play develops further with age: Ely and McCabe (1994) showed that almost a quarter of kindergarten children's utterances involved language play of some kind, including riddles and other forms of humor. In the only spontaneous riddle session observed in the Ely and McCabe (1994) study (from a corpus of almost 7000 child utterances), a child produced five distinct true riddles and also contributed responses to riddles from other children. Although riddles themselves were rare, numerous other forms of humor were described, including wordplay, sarcasm, and obscene/toilet references. These observational findings of children's varied linguistic humor use further raise the question of whether children understand the humor that they use or not.

Considered together, the literature above indicates that older children are better at understanding simple humorous material than younger children, but that enjoyment and appreciation appear to reduce over time, at least of simple verbal jokes. The basic questions that remain only partly answered or unanswered are as follows: do children smile and laugh more at jokes that they understand? Do they rate jokes as funnier if they understand them? If so, at what age do these relationships emerge?

The primary aim of this study was to address these questions, revisiting Zigler et al's (1966) attempt to characterize humor development in terms of (1) understanding, (2) self-reported appreciation, and (3) facial reaction, together in a single developmental study, 50 years onwards. Instead of cartoons, riddles were used, partly to afford comparison with more recent literature (which has tended to use such stimuli) but also because such stimuli reflect

what people tend to mean by ‘jokes’ and primary school children show a fondness for remembering riddles (Cunningham, 2005).

The secondary aim was to examine the possible role of central coherence (Happé & Frith, 2006) in understanding of, self-reported appreciation of, and facial reaction to jokes in children, given this factor’s identified role in language understanding in autism spectrum disorders (Happé, 1997; Jolliffe & Baron-Cohen, 2000). For the current study, we will examine this ‘coherence’ at the level of sentence comprehension, as local rather than global processing; it is beyond the scope of this introduction to discuss whether this might reflect a more domain-general construct. The questions here were: do children with greater local processing bias tend to understand jokes less well, and specifically to interpret them literally? Do children with greater local processing bias find riddles less funny, and do they smile and laugh at them less?

It was predicted that comprehension of riddles would improve with age and that appreciation of riddles would decrease as children get older. However, we anticipated that appreciation of riddles would depend on whether or not children understood the jokes. In addition, we predicted that children with greater local processing bias would find riddles less funny, and also smile and laugh at them less.

Method

Participants

Seventy-four children aged 4 to 11 years old took part in the experiment, recruited for the [blinded for submission].

Participant characteristics are provided in Table 1. Socio-economic status data were available for 72 of the participants, based on the Office of National Statistic Indices of

Multiple Deprivation (2010), a ZIP-code-based database. The sample characteristics (*M* = 13.32, *SD* = 11.85, range = 0-46.2) indicated that participants were represented from all five quintiles, with the mean in the second quintile (quintile group 1 has the least deprived ZIP codes, 5 the most deprived, and 3 the 'averagely' deprived ZIP codes). The frequency distribution across the quintiles was as follows: Q1 = 38, Q2 = 18, Q3 = 5, Q4 = 7, Q5 = 3, no data = 3.

Table 1 about here

Materials

Riddles task

Initially, 20 semantic riddles were chosen from a wide variety of websites on the basis of having simple vocabulary and not appearing obscure in any way. These were screened by a pool of five developmental psychologists, who selected the 12 riddles that seemed the most amusing and likely to engage children. A full list of the riddles used is provided in Appendix 1, along with 12 control stimuli, each of which began with the same question as its corresponding riddle but ended with a non-humorous answer. Although there would inevitably be more experience with, and more complete knowledge of, the vocabulary used in the test items associated with the older children relative to the younger ones, this was deemed necessary on the basis of piloting in order to maintain the interest of older children and allow the examination of development.

Twenty-four video stimuli were created by filming silhouettes of adults telling riddles and riddle-like control stimuli, using a digital camera. The films were of silhouettes in order to eliminate any nonverbal cues from facial expressions or other bodily gestures, while

providing a social attentional focus for participants (see Figure 1) and keeping participants on task. The stimuli were spoken without obvious joking emphasis or prosody, such that the riddles and control stimuli sounded similar.

Figure 1 about here

Participants sat in front of a 15" laptop computer on a table top and wore headphones. Participants were told: "We are going to play a game where we watch some videos of people telling jokes from behind a curtain. Sometimes they're real jokes but sometimes the people just say things that sound like jokes. All you have to do is watch the videos and listen carefully to what the people say in the video. Then you have to let me know how funny you thought the video was by pointing to one of these pictures." At this point, participants were shown a visual scale including Very Funny, Funny, OK, Not Funny for rating their appreciation of the video (see Figure 2).

Figure 2 about here

The stimuli (Supplementary Materials 1) were presented using Microsoft PowerPoint software. There was a single practice trial including the following riddle: "What kind of room has no walls? A mushroom". When the video had finished playing, the PowerPoint presentation was moved on to the screen shown in Figure 2 and the participant was asked, "How funny was that video?". The experimenter noted the response on a score sheet. If the participant did not point to one of the four response options, the experimenter reminded the participant of each possible option by pointing to each and stating what each meant. If the participant still did not respond, they were asked to give their 'best guess'. At this point, if a child did not respond or did not appear to understand the requirements of the task, they were thanked and the experiment was ended. None of the 74 children included in the study had any apparent difficulty in understanding the task.

Next, participants were presented 12 experimental trials. Half of these were riddles and half were control stimuli. Riddles and control stimuli were semi-randomized into two orders, with order B being the reverse list order of order A and also reversing whether each stimulus ending was a joke or control ending. Half of the participants in each age group were presented with Order A, with the other half presented with Order B.

Participants were filmed by the laptop's built-in camera during the experiment and recorded with Quicktime, so that their facial reactions to the stimuli could be subsequently coded (see 'Scoring' section below). The SoundFlower driver was used so that Quicktime captured the video of the participants but the audio played via PowerPoint, rather than that captured by the laptop's microphone. This afforded accurate editing of the video, time-locked to the start of each trial.

After the experimental trials had been completed, participants were told, "Now we're going to watch some of the same videos again. I'll remind you of what you thought of the video and ask you to tell me why you thought it was funny or not funny, or just OK." Starting with the practice trial, participants were then shown the riddles, but not the control stimuli, in the same order that they had encountered them previously. Instead of rating each video, they were told the rating that they had given (e.g., "You thought this was just OK") and were asked, "Can you explain why? Why was it very funny/funny/just OK/not funny?" The experimenter transcribed these responses, which were categorized after the session.

Sentence Completion Task

In order to examine local processing bias, a slightly modified version of the Sentence Completion Task (Booth & Happé, 2010) was administered. Participants were told, "We're going to do a quick game that is totally different. I'm going to read out some sentences and you're going to finish them off for me. Let's try it." Fourteen sentence stems were then read

out, 10 of which were designed to elicit a possible local completion (e.g., You can go hunting with a knife and….", intended to invite the local response "fork", whereas a context-appropriate global response would be "gun"), with 4 filler items ("I was given a pen and…"). Responses were digitally recorded and then scored after the session, as local, global or other. The dependent variable was the total number of local completions (the maximum was 10).

Background measures

Background measures were collected to characterize levels of receptive vocabulary skill and social aptitude, two of the identified determinants of individual differences in humor comprehension (e.g., Southam, 2005). In order to verify that participants had age appropriate vocabulary abilities, vocabulary comprehension scores were obtained from the *British Picture Vocabulary Scales-3* (BPVS; Dunn, Dunn, Styles, & Sewell, 2009) and converted to standardized scores. In this task participants are shown 4 pictures and need to point to the picture that corresponds to the word spoken by the examiner. There was a maximum of 14 sets with 12 items per set. Children started with the set recommended for their age and finished with the set in which they made eight or more errors.

Parents completed the *Social Aptitudes Scale* (SAS), which is part of the Development and Well-being Assessment (Goodman, Ford, Richards, Gatward, & Meltzer, 2000) and assesses socially skilled behaviour in children. This questionnaire includes 10 questions with (unstandardized) scores ranging from 0 to 40. It has a good internal consistency with a Cronbach's alpha of .88 (Goodman et al., 2000).

Procedure

Ethical approval was obtained from the School of Psychology, University of Nottingham. Parental consent and child assent were obtained from all children who took part in the study. As a reward for participating, children were given a token that could be used to

play a variety of traditional fairground-style games at the Summer Scientist Week public engagement event.

Scoring of riddles task

The video footage from the riddles task of each participant was edited using Apple iMovie. Clips of the experimental trials were edited to begin at the onset of each audio stimulus and ended with the initiation of the participant's verbal response. These clips were then coded for the presence or absence of smiling and/or laughter (0 for no smiling or laughter, 1 for any smiling and/or laughter) by a coder blind to whether each stimulus was a riddle or a control as no sound was played during the coding of the videos. The time taken for the initiation of any smiling or laughing was also recorded, measured from the end of the audio stimulus on each trial.

Explanations were scored according to the following scheme:

- *Explicit understanding*: child explained the joke, pointing out the lexical ambiguity
- *Explicit lack of understanding*: child identified that they did not understand joke or know why they rated the joke that way.
- *Literal*: child interpreted the meaning of the joke literally, e.g., 'The students were bright, so they were shining light from their bodies.'
- *Echolalic*: child just repeated fragments of what was said in the joke, e.g., 'There were students' or 'The students were bright.'
- *Other*: responses that did not fit into the above categories, e.g., 'Because I liked it.'

Results

Descriptive statistics are provided in Table 2.

Table 2 about here

Although exploratory data analyses did not indicate any serious deviation from normality for continuous measures, nonetheless Spearman correlations and logistic regressions were used throughout to obviate standard concerns about meeting parametric assumptions for smaller subsamples.

The simple relationship of age to humor measures

Spearman correlations revealed that age was not significantly associated with variability in appreciation ratings of the riddles, $r_s(72) = -.18$, $p = .127$, but that younger children tended to rate the control stimuli as funnier than older children, $r_s(72) = -.51$, $p < .001$. Older children tended to smile or laugh at jokes more than younger children, $r_s(72) = .24$, $p = .039$, whereas there was no reliable age related variance in smiling/laughing at the control stimuli, $r_s(72) = .08$, $p = .523$. Furthermore, older children were reliably better able to comprehend the riddles than younger children, $r_s(72) = .68$, $p < .001$, but were not significantly less likely to interpret the riddles literally, $r_s(72) = -.17$, $p = .141$.

The unique relationship of age to humor measures

To complement the previous analyses and explore the *unique* relationships between measures, a series of multilevel multiple logistic regressions were run, with participant and trial (i.e., the particular joke or control stimulus) as random factors and appreciation (self-rated), comprehension, and smiling/laughing as dependent variables, and age, BPVS (receptive vocabulary), SES (socio-economic status), and SAS (social aptitude) as fixed predictor variables (see Table 3).

Table 3 about here

These analyses revealed that age, having controlled for the other predictors, was not significantly associated with appreciation (funniness ratings) of the riddles. However, older

children were reliably better able to explain the riddles than younger children, although they were not significantly more or less likely to interpret the riddles literally. Turning to the control stimuli, there were no reliable predictors of smiling/laughing, but appreciation (rating non-jokes as funny in this case) significantly declined with age.

The unique relationship of participant characteristics to experimental measures

Independently from age, several other participant characteristics reliably predicted experimental outcomes. Lower appreciation was significantly predicted by higher SES, with lower appreciation of control stimuli significantly associated with receptive vocabulary knowledge. Comprehension was also significantly predicted by receptive vocabulary knowledge (BPVS). Smiling/laughing was not reliably associated with any other participant characteristics. Finally, literal interpretations of riddles were significantly predicted by poorer vocabulary knowledge.

Local processing

Local processing was not significantly correlated with age, r_s (72) = -.227, p = .052, appreciation of riddles, r_s r(72) = .100, p = .398, appreciation of control stimuli, r_s (72) = .022, p = .852, comprehension, r_s (72) = -.114, p = .335, or smiling/laughing at riddles, r_s (72) < .001, p = .998. Multilevel logistic regressions with participant and trial as random variables showed no significant relationship between local processing and appreciation, Z = 0.129, p = .898, local processing and comprehension, Z = -1.10, p = .273, local processing and literal interpretation, Z = 1.37, p = .170, nor local processing and smiling/laughing, Z = -.090, p = .928.

In order to examine whether results were evident or not for particular ages, affording comparison to literature discussed in the introduction section, the following analyses

exploring associations between experimental variables were split by age group. All of the following analyses were multilevel logistic regressions with participant and trial as random variables.

Appreciation and comprehension

Greater appreciation reliably predicted successful comprehension of riddle stimuli, $Z = 4.44$, $p < .001$. This may appear to be incongruent with the pattern of results in Table 3, given that age is negatively (though not reliably) related to appreciation, but *positively* related to comprehension. One way of explaining this would be to state that there are age-related changes in outcome measures indicated in Table 3, but those are logically independent of the relationships between variables that change within the testing session, because appreciation and comprehension are causally related to the changing stimuli presented, whereas age is not.

A further analysis split by age-group showed that this relationship was not significant for 4- to 5-year-olds, $Z = .014$, $p = .988$, at least in part because children in that age group could explain the riddles on only 3.5% of occasions, and was marginally non-significant for 6- to 7-year-olds, $Z = 1.79$, $p = .074$, but was significant for both 8- to 9-year-olds, $Z = 3.37$, $p < .001$, and 10- to 11-year-olds, $Z = 3.66$, $p < .001$.

Smiling/laughing

Smiling/laughing was significantly associated with greater appreciation of riddles, $Z = 6.26$, $p < .001$, and with successful comprehension of the riddle, $Z = 2.43$, $p = .015$. However, an analysis split by age-group showed that this latter effect was only evident for 6- to 7-year-olds, $Z = 2.392$, $p = .017$, and 8- to 9-year-olds, $Z = 2.13$, $p = .033$, but not 4- to 5-year-olds, $Z = .513$, $p = .608$, or 10- to 11-year-olds, $Z = -.543$, $p = .587$. As with the ‘appreciation and

comprehension' analysis above, the lack of effect for 4- to 5-year-olds is likely due to the scarcity of successful comprehension events in that group.

Discussion

The current study examined the development of humor comprehension and appreciation in typically developing children using a cross-sectional design. In addition, the impact of understanding riddles on appreciation and smiling/laughter was investigated within each age group. Finally, the relationship between local processing bias and the comprehension and appreciation of riddles was explored.

There was no evidence of age-related change in self-rated appreciation of the riddles. However, older children did tend to smile and laugh more at the jokes and demonstrated greater understanding, in that they could explain the riddles better than younger children. These findings are in line with those of previous studies that have shown that comprehension of riddles increases with age (Whitt & Prentice, 1977; Zigler et al., 1966), but not with the finding of a concomitant decrease in the appreciation of riddles (Whitt & Prentice, 1977). Although Whitt and Prentice (1977) suggested that children's liking for homophonic riddles might decline over time, this would presumably depend on the particular riddles used as stimuli: those in the current study were chosen on the basis of piloted suitability across the whole age range. It should be noted that appreciation of non-joke control materials declined with age. A parsimonious explanation is that younger children tend to appreciate materials that have the structure of jokes regardless of whether they are actually jokes or not. We will return to this point below.

The current study expanded previous research by exploring the various relationships between comprehension, appreciation, and laughter/smiling when hearing riddles. Higher funniness ratings predicted successful comprehension only in 8- to 11-year-olds, with 4- to 5-

year-olds almost entirely unable to successfully explain any of the jokes. Putting this together with the results outlined above, there appears to be an age-dependent emergent relationship between appreciation and comprehension.

As might be expected, smiling/laughing was also associated with higher funniness ratings of riddles. For 6- to 9-year-olds only, smiling/laughing was also related to successful comprehension of the riddle. A tentative explanation for this might be that younger children either do not understand or are unable to successfully express an explanation for riddles that they nevertheless find amusing, whereas the oldest children – who successfully explained most of the riddles – might only smile or laugh at jokes that are not too 'obvious' and make them 'think twice' (Masten, 1989). Given these possibilities, children's appreciation of humor should not be taken as firm evidence of understanding. This latter notion fits with a discussion by Whitt and Prentice (1977), who noted that children's understanding of riddles developed between the ages of six and eleven, despite enjoyment being evident throughout that age range. Social Aptitude Scale scores uniquely predicted smiling/laughing in our sample, raising the possibility that elicited positive affect might – at least in part – reflect something beyond mere comprehension: namely, a willingness or ability to behave appropriately in a context where smiling and laughing might be expected.

The results above showing minimal comprehension in 4- to 5-year-olds and only a reliable relationship between appreciation and comprehension in 8- to 11-year-olds might appear to conflict with previous findings that children younger than eight years often produce verbal humor (Ely & McCabe, 1994), including riddles, and even preschool children produce humor-like verbal play (Esposito, 1980). A face-value interpretation would be simply that younger children often use humor that they do not understand: instead, they might be amused by the incongruence alone (see McGhee, 1979) or by learned associations of amusement to the relevant humor. An alternative account would be that comprehension is neither important

nor necessary for much of younger children's humor, such as toilet humor, rhymes, and nonsense stories, which appear superficial relative to semantic wordplay and riddles. A further possibility is that the inherent metacognitive demands of explicit comprehension tasks impose performance limits on younger children's demonstration of underlying competence in understanding humor. The use of a broader range of methods for assessing children's comprehension of humorous materials is likely to be a fruitful avenue of future investigation.

Whatever the ultimate explanation may prove to be for the age-related differences in comprehension, comprehension itself was predicted not just by age, but also by receptive vocabulary. It is unsurprising that superior vocabulary knowledge – which might be considered a proxy for language knowledge more generally – would enhance comprehension, simply because full understanding of a joke directly depends on adequate knowledge of the relevant word meanings (and language structures) and would also facilitate the working memory required to integrate the riddle's meaning overall (e.g., Hulme, Maughan, & Brown, 1991).

The only other notable relationship between a background measure and an experimental measure was that SES predicted appreciation of the riddles. Although other explanations of this relationship are possible, it is well established that higher SES is associated with greater intelligence (e.g., Cuff, 1933; Von Stumm & Plomin, 2015), presumed to be mediated by differential access to support/ resources (Hayes, 1962) or parental language input (Hoff, 2003). The suggestion would therefore be that that higher SES supports general cognitive ability, in turn supporting necessary components of appreciation, such as language processing and working memory.

Addressing the final set of aims, tendency to local processing was not associated with any of the experimental measures. These findings are at odds with previous studies that have

suggested that a local processing bias affects context understanding in general language (Happé & Frith, 2006) and that this can interfere with the comprehension of jokes (Mitchell et al., 2010). One possibility is, of course, that the comprehension of the type of homophone-based riddles used in the current study is simply not particularly disrupted by any tendency to local processing: the integration requirements of such riddles might be too low to cause problems for local processors. Another is that participants were allowed as long as they liked to attempt to explain the riddles and were therefore afforded as much time as they might need to overcome any difficulties imposed by any local processing tendency. It may therefore be necessary to assess comprehension in a speeded fashion in any future research.

Given that humor may be very personal, future studies may want to further explore individual differences in comprehension, in both typically and atypically developing children. In addition, the current study used a cross-sectional design as it highlights individual differences. Yet, these individual differences may mask developmental trends. Therefore, the findings in the current study should be followed-up by longitudinal studies. Riddles are also only one particular type of humor and thus, future studies should also examine a wider range of humor types to examine developmental trends in humor understanding and appreciation. The interpretation of the results of the current study is constrained by the fact that the same riddles were experienced by children of all ages, so that any age-related differences in appreciation, smiling/laughing, or comprehension are differences with respect to quite simple homophonic riddles and should not be generalised to some more global notion of humor. More generally, laboratory-based experiments necessarily operationalize and formalize stimuli and responses in a way that abstracts them from day-to-day enculturated spontaneous humorous exchanges. Ultimately, such experimental studies are valuable principally as complements to more in-depth observational research.

In sum, although what people find funny or not might be very personal, there are clear developmental trends in riddle understanding, with comprehension increasing with age. In addition, although younger children might smile to show that they somehow recognize someone is telling them a joke, only in older children (older than 4/5) is laughter or smiling a reliable predictor of appreciation. The relationship of smiling/laughter to comprehension may be more complex than this, at least with riddles, with appropriate smiling or laughter appearing to be a good indicator of humor comprehension only for certain age groups. Given that humor is a central part of our social communication and impacts on inter-personal relationships and intra-personal emotional functioning (Martin, 2007), understanding what children find funny and how this relates to cognitive processing styles (including WCC, memory, language abilities, etc) will allow further examination of humor processing in populations that experience difficulties with humor processing (e.g., children with autism spectrum disorders or those learning an additional language) and will allow for the development of interventions to fully integrate these groups. Children's laughter in response to the use of humor, while evidence of change in their affective state, is not necessarily a marker of understanding: some laughter may just be elicited by the pragmatic context. A closely related notion is that riddles (or some other set of humorous materials) might only be 'funny' to a particular age group, such that funniness is highly age-related. Given that development is rarely taken into account when studying humor, especially in atypical populations, fuller understanding of these phenomena seems a worthwhile goal for future research.

Open Practices Statement

The experiment reported in this article was not formally preregistered. The data have been made available on a permanent third-party archive; requests for the data or materials can be sent via email to the lead author at [email].

References

Booth, R., & Happe, F. (2010). "Hunting with a knife and ... fork": Examining central coherence in autism, attention deficit/hyperactivity disorder, and typical development with a linguistic task. *Journal of Experimental Child Psychology*, *107*(4), 377 - 393. doi: 10.1016/j.jecp.2010.06.003

Bruno, R. M., Johnson, J. M., & Simon, J. (1987). Perception of humor by regular class students and students with learning disabilities or mild mental retardation. *Journal of Learning Disabilities*, *20*(9), 568-570. doi: 10.1177/002221948702000910

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, *96*(1), 31-42. doi: 10.1037/0022-0663.96.1.31

Cuff, N. B. (1933). Relationship of socio-economic status to intelligence and achievement. *Peabody Journal of Education*, *11*(3), 106-110. doi: 10.1080/01619563309535185

Cunningham, J. (2005). Children's humor chapter. In W. G. Scarlett, S. Naudeau, D. Salonius Pasternak, & I. Ponte (Eds.), *Children's Play*. Thousand Oaks, CA: SAGE Publications, pp.93–109.

Degabriele, J., & Walsh, I. P. (2010). Humor appreciation and comprehension in children with intellectual disability. *Journal of Intellectual Disability Research*, *54*(6), 525-537. doi: 10.1111/j.1365-2788.2010.01277.x

Dews, S., Winner, E., Kaplan, J., Rosenblatt, E., Hunt, M., Lim, K., ... & Smarsh, B. (1996). Children's understanding of the meaning and functions of verbal irony. *Child development, 67(6),* 3071-3085. doi: 10.1111/j.1467-8624.1996.tb01903.x

Dunn, L. M., Dunn, D. M., Styles, B., & Sewell, J. (2009). *The British Picture Vocabulary Scale III* – 3rd Edition. London: GL Assessment.

Ely, R., & McCabe, A. (1994). The language play of kindergarten children. *First Language, 14(40),* 19-35.

Esposito, A. (1980). Children's play with language. *Child Study Journal, 10,* 207-216.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41,* 1149-1160. doi: 10.3758/BRM.41.4.1149

Garvey, C. (1977). Play with language and speech. In C. Mitchell-Keman & S. Ervin (Eds), *Child Discourse*. New York: Academic Press.

Goodman, R., Ford, T., Richards, H., Gatward, R., & Meltzer, H. (2000). The Development and Well-Being Assessment: description and initial validation of an integrated assessment of child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry* 41(5), 645–655. doi: 10.1111/j.1469-7610.2000.tb02345.x

Happé, F. G. (1997). Central coherence and theory of mind in autism: Reading homographs in context. *British Journal of Developmental Psychology*, *15*(1), 1-12. doi: 10.1111/j.2044-835X.1997.tb00721.x

Happé, F., & Frith, U. (2006). The weak coherence account: detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *36*(1), 5-25. doi: 10.1007/s10803-005-0039-0

Hayes, K. J. (1962). Genes, drives, and intellect. *Psychological Reports, 10(2)*, 299-342. doi: 10.2466/PR0.10.2.299-342

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development, 74*(5), 1368-1378. doi: 10.1111/1467-8624.00612

Hulme, C., Maughan, S., & Brown, G. D. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language, 30*(6), 685-701. doi: 10.1016/0749-596X(91)90032-F

Jolliffe, T., & Baron-Cohen, S. (2000). Linguistic processing in high-functioning adults with autism or Asperger's syndrome. Is global coherence impaired? *Psychological Medicine, 30*(5), 1169-1187. doi: 10.1017/S003329179900241X

Martin, R. A. (2007). *The psychology of humor: An integrative approach*. Burlington, MA: Elsevier Academic Press.

Masten, A. S. (1989). Humor appreciation in children: Individual differences and response sets. *Humor: International Journal of Humor Research, 2*(4), 365-384.

McGhee, P. (1979). *Humor: Its origin and development*. San Francisco: Freeman.

Mitchell, H.H., Graesser, A.C. & Louwerse, M.M. (2010). Effect of context on humor: A constraint-based model of comprehending verbal jokes. *Discourse Processes, 47*(2), 104–129. doi: 10.1080/01638530902959893

Semrud-Clikeman, M., & Glass, K. (2008). Comprehension of humor in children with nonverbal learning disabilities, reading disabilities, and without learning disabilities. *Annals of Dyslexia*, *58*(2), 163-180. doi: 10.1007/s11881-008-0016-3.

Southam, M. (2005). Humor development: An important cognitive and social skill in the growing child. *Physical & Occupational Therapy in Pediatrics*, *25*(1-2), 105-117. doi: 10.1080/J006v25n01_07

Von Stumm, S., & Plomin, R. (2015). Socioeconomic status and the growth of intelligence from infancy through adolescence. *Intelligence*, *48*, 30-36. doi: 10.1016/j.intell.2014.10.002

Whitt, J. K., & Prentice, N. M. (1977). Cognitive processes in the development of children's enjoyment and comprehension of joking riddles. *Developmental Psychology*, *13*(2), 129-136. doi: 10.1037/0012-1649.13.2.129

Winner, E. (1997). *The point of words: Children's understanding of metaphor and irony*. Cambridge, MA: Harvard University Press.

Winner, E., & Gardner, H. (1993). Metaphor and irony: Two levels of understanding. ding. In A. Ortony (Ed.), *Metaphor and thought* (2nd ed., pp. 425-446). Cambridge, UK: Cambridge University Press.

Zigler, E., Levine, J., & Gould, L. (1966). Cognitive processes in the development of children's appreciation of humor. *Child Development*, *37*(3), 507-518. doi: 10.2307/1126675

Table 1.

Participant characteristics by age group, including mean chronological age, Vocabulary knowledge from British Picture Vocabulary Scales (standardized scores; BPVS) and Social Aptitudes Scale (SAS).

	Age		BPVS		SAS	
Age Group	Mean	*SD*	Mean	*SD*	Mean	*SD*
4-5 (N=20)	5.25	0.54	104.42	26.59	22.47	5.24
6-7 (N=17)	6.89	0.62	100.82	27.79	22.00	7.87
8-9 (N=19)	8.82	0.59	103.88	29.36	24.17	6.93
10-11 (N=18)	10.94	0.69	102.67	10.24	25.50	7.82

Table 2.
Descriptive statistics by age group for the main experimental measures. Ratings are the average per trial (max=4); Smiling (smiling/laughing), Comprehension and Local Processing are proportions of all applicable trials.

	Measure											
	Joke Appreciation		Control Appreciation		Comprehension		Joke Smiling		Control Smiling		Local Processing	
Age-Group	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
4-5	1.48	0.61	1.21	0.74	0.04	0.09	0.33	0.34	0.30	0.31	0.25	0.15
6-7	1.39	0.59	0.74	0.64	0.18	0.22	0.43	0.34	0.24	0.23	0.21	0.18
8-9	1.37	0.56	0.36	0.37	0.38	0.24	0.43	0.29	0.29	0.28	0.14	0.09
10-11	1.21	0.63	0.33	0.43	0.51	0.29	0.58	0.36	0.36	0.30	0.18	0.18

Table 3.

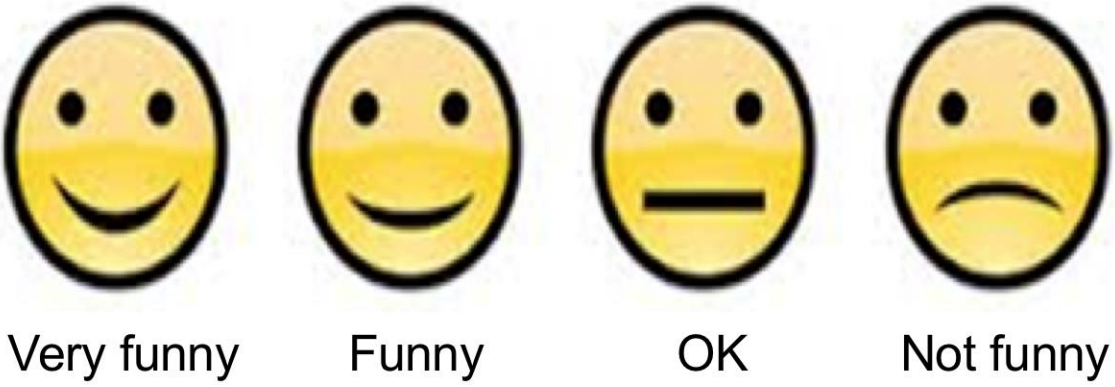
*Multilevel logistic regressions of appreciation, comprehension, smiling/laughing, literal interpretations, control appreciation (ratings of non-jokes), and control smiling/laughing as outcome variables. Participant ID was modelled as a random effect, with fixed effects of age, BPVS (receptive vocabulary), SES (socio-economic status), and SAS (social aptitude). Appreciation has been recoded to a binary variable where 1 represents ratings of 'OK' to 'very funny') and represents 0 ratings of 'terrible' and 'not funny'). *p < .05, **p < .001*

	Standardized effect size *Z*			
Outcome	Age	SES	BPVS	SAS
Appreciation	-1.70	**-2.37***	.474	.896
Comprehension	**2.84****	-1.90	**2.44***	.582
Smiling/laughing	1.74	.793	-.120	1.32
Control appreciation	**-3.85****	-.591	-1.48	-.767
Control smiling/laughing	.857	.994	-.564	.853

Figure 1.

Still image of example 'silhouette' video shown to participants.

Figure 2: Funniness scale



Appendix 1. Overview of the stimuli (joke/control endings).

What is the best day to go to the beach? Sun-day / Whenever the weather is good

Why do bears have fur coats? Because they'd look silly wearing jackets / To keep them warm when it's cold

Why do humming birds hum? Because they forgot the words / It's the sound their wings make

What kind of tree can fit into your hand? A palm tree / A baby tree

Why did the robber take a bath? Because he wanted to make a clean getaway / Because he was dirty

What is the hardest part about skydiving? The ground / Daring to jump out of the plane

What do you call a deer with no eyes? No idea ['No-eye-deer'] / A blind deer

Why do bees have sticky hair? Because they have honey combs / Because honey gets stuck to them

Why is there a gate around cemeteries? Because people are dying to get in/ To keep people out when they're closed

How do fleas travel from place to place? By 'itch-hiking / By jumping

Why did the teacher wear sunglasses? Because his students were bright / Because it was a sunny day?

Why do jelly beans go to school? Because they want to become smarties / Because they want to learn maths