Running head: CONSISTENCY BETWEEN AFFECTIVE CUES

Consistency between verbal and non-verbal affective cues: A clue to speaker credibility

Randall L. Gillis[a] & Elizabeth S. Nilsen[a]

[a]University of Waterloo

200 University Avenue West, Waterloo, Ontario, CANADA

Send correspondence to:

Elizabeth Nilsen
Psychology Department
University of Waterloo
200 University Avenue West
Waterloo, Ontario, Canada N2L 3G1
enilsen@uwaterloo.ca
Tel: 518-888-4567 x33991
Fax: 519-746-8631

Acknowledgements

Abstract

Listeners are exposed to inconsistencies in communication; for example, when speakers' words (i.e., verbal) are discrepant with their demonstrated emotions (i.e., non-verbal). Such inconsistencies introduce ambiguity, which may render a speaker to be a less credible source of information. Two experiments examined whether children make credibility discriminations based on the consistency of speakers' affect cues. In Experiment 1, school-age children (7- to 8-year-olds) preferred to solicit information from consistent speakers (e.g., those who provided a negative statement with negative affect), over novel speakers, to a greater extent than they preferred to solicit information from inconsistent speakers (e.g., those who provided a negative statement with positive affect) over novel speakers. Preschoolers (4- to 5-year-olds) did not demonstrate this preference. Experiment 2 showed that school-age children's ratings of speakers were influenced by speakers' affect consistency when the attribute being judged was related to information acquisition (speakers' believability, 'weird' speech), but not general characteristics (speakers' friendliness, likeability). Together, findings suggest that school-age children are sensitive to, and use, the congruency of affect cues to determine whether individuals are credible sources of information.

(179 words)

Keywords: children; selective trust; speaker credibility; non-verbal cues; emotion recognition; affective cues; communication;

Consistency between verbal and non-verbal affective cues: A clue to speaker credibility

Speakers convey their feelings and intentions through both the content of their statements and the manner with which they make their utterances. While speakers' non-verbal cues often support the literal meaning of their statements, there are instances when the emotional valence of verbal (i.e., what they say) and non-verbal information (i.e., how they say it) are inconsistent [e.g., the Grinch displaying a giant grin while stating, "They're finding out now that no Christmas is coming... the Whos down in Whoville will all cry 'Boo Hoo,'" (Seuss, 1966)]. Such inconsistencies can be exploited to produce nuanced communicative behaviour (e.g., verbal irony) or may occur inadvertently when a speaker attempts to mask their true emotions. Moreover, inconsistencies in communicative cues occur more often when individuals are lying (DePaulo, Stone, & Lassiter, 1985; Feldman & White, 1980). Given that inconsistencies occur in the context of everyday communication and may be indicative of unreliable information, we sought to examine whether children detect these inconsistencies and what perceptions they hold of such speakers. Specifically, we examined whether preschool and school-age children attend to the consistency of affective cues when deciding on the credibility of speakers.

Adopting a discriminating stance towards speakers is adaptive given that children are exposed to vast amounts of new information from others and are unable to absorb everything (e.g., Harris, 2002). Further, at times it could be detrimental to trust an individual's verbal testimony (e.g., individuals lie or unknowingly deliver incorrect information). A large body of research has demonstrated that children are selective when deciding from whom to solicit information (see Mills, 2013 for a review) and show a preference for learning from individuals who possess various characteristics, in particular, individuals who are accurate and knowledgeable (e.g., Corriveau, Meints, & Harris, 2009; Koenig, Clement & Harris, 2004;

Koenig & Harris, 2005; Scofield & Behrend, 2008). When inferring a speakers' credibility, children also rely on the non-verbal cues speakers provide as they deliver information (defined here as both facial expression and tone of voice). For example, Birch, Akmal and Frampton (2010) found that 2-year-olds preferred to learn from individuals who displayed confident, as opposed to uncertain, non-verbal cues (also see Brosseau-Liard, Cassels, & Birch, 2014). Despite this research, it is unclear how children apply their sensitivity to non-verbal cues in complex communicative situations, for example, when a speaker's words conflict with her non-verbal affect.

Children's ability to interpret non-verbal cues precedes their ability to interpret the words uttered by speakers (O'Grady, 2005). Within their first year of life, infants show sensitivity to vocal tones and facial expressions indicative of different affective states (Barrera & Maurer, 1981; Fernald, 1993; Kuchuk, Vibbert & Bornstein, 1986). At 4 years old, children can use a speaker's vocal affect to interpret ambiguous messages (Berman, Chambers, & Graham, 2010) and can label emotions from facial expressions (Ridgeway, Waters, & Kuczaj, 1985). By 5 years of age, children recognize that individuals are capable of expressing emotions that differ from those that they are experiencing (Wellman & Liu, 2004). Research on children's interpretation of inconsistent messages (i.e., those where the verbal information is incongruent with the non-verbal information) highlight developmental differences in children's reliance on various affective cues. During the preschool years, children demonstrate a verbal primacy, in that their interpretations of the emotional valence of inconsistent utterances are based primarily on the verbal content (as opposed to non-verbal cues; e.g., Lawrence & Fernald, 1993 [as cited in Friend, 2001]; Morton & Trehub, 2001; Friend & Bryant, 2000; Friend, 2000; 2003; Waxer & Morton, 2011). For example, if they hear, "My mommy gave me a treat," said in a sad voice,

they are likely to interpret the speaker's emotion as happy (Morton & Trehub, 2001). The

reliance on non-verbal information increases through childhood until adulthood when a non-

verbal primacy is demonstrated (e.g., Solomon & Ali, 1972; Morton & Trehub, 2001; Mehrabian

& Ferris, 1967; Argyle, Alkema, & Gilmour, 1971; Reilly & Muzekari, 1986). Interestingly,

when faced with speakers whose words differ from the affect cues displayed, children tend not to

notice the inconsistencies until about 6-7 years of age (Morton & Trehub, 2001; Rotenberg,

Simourd & Moore, 1989). However, implicit sensitivity to inconsistency is shown earlier: when

judging speakers' emotions, children as young as 4 years old take longer to respond to

inconsistent, compared to consistent, utterances (Morton & Trehub (2001).

       While young children are discerning learners and show implicit sensitivity to

inconsistencies in communicative cues, it is not clear whether these two areas of development

converge. Addressing this, the present work examined whether preschool and school-age

children form different perceptions of a speaker's credibility based on (in)consistencies between

her verbal and non-verbal cues—specifically inconsistencies between affective information

conveyed in verbal form (i.e., the words) and affective information conveyed non-verbally

(facial expression/tone). As both children and adults demonstrate greater inconsistency in their

verbal and non-verbal communications when they are lying versus telling the truth (DePaulo et

al. 1985; Feldman & White, 1980), inconsistency would be a useful heuristic for determining

whether information from a speaker is accurate or not. There are several other reasons to believe

that (in)consistency would be an important cue for children. Inconsistent affective cues introduce

ambiguity into communication because the speaker is conveying different information through

the different communicative channels. Therefore, to avoid misunderstanding the message, it

would be beneficial for children to rely on speakers who provide clear information (i.e.,

delivered with consistent communicative cues). Indeed, past work has found that children prefer to learn from speakers who use clear, as opposed to ambiguous references (Gillis & Nilsen, 2013). In addition, children prefer speakers who are consistent in other ways; for example, they trust individuals whose non-verbal cues (e.g., saying "wow!" and smiling) are consistent with the context (e.g., looking in a box with a toy) as opposed to inconsistent (e.g.., looking in a box without a toy; Chow, Poulin-Dubois, & Lewis, 2008; Chiarella & Poulin-Dubois, 2013).

We expected that children would be more likely to solicit information from speakers who showed consistency, as opposed to inconsistency, between the verbal content of their statements and their facial/tonal expressions. Two studies tested this premise in varying contexts: in Experiment 1 children made choices about information sources in a context where the goal was to acquire the 'accurate' facts from a story (epistemic goals); whereas in Experiment 2, children were asked whether they would like to use the personal preferences of a speaker to help them choose an unknown prize (non-epistemic goals). To understand what inferences might be driving children's choices, we examined children's attributions of the speakers in terms of attributes relevant to information delivery (believability, 'weirdness' of speech) as well as those unrelated (likeability, friendliness).

Experiment 1

To examine whether children use verbal/non-verbal (in)consistency as a cue to speaker credibility, preschool-age (4- to 5-year-old) and school-age (7- to 8-year-old) children completed a speaker affect task in which they were exposed to a speaker who either delivered information that was consistent (e.g., a positive statement said in a positive tone of voice) or inconsistent (e.g., a positive statement said in a negative tone of voice). Children then chose whether they wanted to receive a new piece of information about a story from that speaker or from a speaker

of which they had no prior knowledge. That is, they could choose to solicit information from the speaker that they had watched making a video-recorded statement, or someone they had not seen. This procedure was different from a common speaker reliability procedure where children are presented with two speakers and then are required to choose which speaker they would like to 'learn' from. The rationale for this change in methodology was to create a scenario where children were exposed to one source of information and could choose to rely on this person or not. This methodology also had the advantage of reducing the working memory demands (i.e., children did not have to hold in mind the type of information from more than one speaker).

Method

      Participants. Twenty 4- and 5-year-olds (12 males, $M = 62.15$ months, $SD = 5.58$) and 22 7- and 8-year-olds (11 males, $M = 96.09$ months, $SD = 8.01$) were tested individually within a research laboratory. The ethnic background was reflective of the mid-sized North American community the sample was drawn from, with the majority Caucasian. Six children were tested but not included in the analyses due to difficulties completing the task ($n = 2$) or comprehending English ($n = 4$). As assessed by a standardized measure of receptive vocabulary, (i.e., the Test of Language Development Primary Third Edition [TOLD-P:3; Newcomer & Hammill, 1997]), all children possessed language skills sufficient to understand the speaker statements.

      Materials and procedure.

      Speaker affect task. The children's task was to watch video-recorded speakers, one at a time, and decide whether to solicit information from the shown speaker or from another individual (whom children had no information about; Figure 1). Speakers differed in the consistency with which they delivered affective information. Specifically, speakers provided positive or negative verbal information, and positive or negative non-verbal cues. This allowed

for four speaker types: consistent positive (Pos-Verbal/Pos-NV: positive statement said with positive non-verbal cues; e.g., "I found my favourite book today," said with a smile), consistent negative (Neg-Verbal/Neg-NV: negative statement said with negative non-verbal cues; e.g., "I lost my favourite book today," said with a frown) and two inconsistent speakers (Pos-Verbal/Neg-NV: positive statement said with negative non-verbal cues; e.g., "I found my favourite book today," said with a frown, and Neg-Verbal/Pos-NV: negative statement said with positive non-verbal cues; e.g., "I lost my favourite book today," said with a smile). For positive non-verbal cues, speakers were instructed to sound happy, smile, and use speech that was faster, higher pitched, had more pitch variability and more intensity. For negative non-verbal cues, speakers were instructed to sound sad, bring their eyebrows down into a sad frown, and use speech that was slower, lower pitched, with less pitch variability and less intensity. The audio files of speaker's statements were analyzed with the program PRAAT (Boersma, 2001). Analyses for all paralinguistic variables (duration of utterance, pitch mean, pitch standard deviation, intensity) revealed that, as designed, there were differences across non-verbal valence (ps < .001), but not verbal valence (ps > .17). All 12 speakers were Caucasian women with brown hair pulled back from their face. Each speaker wore a different colour t-shirt. The type of information provided by each speaker (i.e., inconsistent versus consistent) and the content of the statement, was counterbalanced across children. The order in which children encountered each of the 12 speakers and the order of the type of information delivered was randomized.

Children were told that their task was to figure out a story by soliciting details from different speakers. They completed 12 trials (three per speaker type) while seated at a table in front of a computer and storybook. Each trial began with the children watching a video-recorded speaker making a statement. Next, children decided whether they wanted to solicit a missing

detail of the story from that speaker or from a different individual that they had no information about (i.e., *speaker choice*; "Do you want this girl to help you figure out part of the story or another girl?"). Choices were scored '1' if the speaker was chosen or '0' if the other girl was chosen (i.e., a possible total of 3 for each of the 4 speaker types). After their choice, children rated how believable the speaker was with the aid of a visual scale (i.e., *speaker rating*; "How much do you believe this girl? Not at all, not much, mostly or very much"). Ratings were scored from a 1 (not at all) to a 4 (very much). Of note, to ensure that children understood the word "believe," 30% of the sample were assessed. All of these children, including the 4-year-olds, demonstrated appropriate knowledge of the word, "believe".

Each page of the storybook depicted a question about the story (e.g., "What did Johnny eat for breakfast?") as well as two contradicting responses with pictures: one from the speaker and one from the other girl (i.e., each girl was pictured with their response in a speech bubble). Importantly, children did not see the page depicting the girls' responses until *after* making their decision. This ensured that participants did not base their responses on their own preferences. To highlight that there were 'correct' answers and to increase motivation for obtaining accurate information, children were told that at the end of the task, the real story would be consulted and they would receive a stamp for every correct detail.

Feelings rating task. A task was administered to determine whether there were age differences in affect recognition abilities. Children watched 12 new videos depicting different speakers than those in the Speaker Affect task, but who said the same statements (i.e., resulting in the same four speaker types). After watching each speaker, children rated (with the aid of a visual scale) how the speaker was feeling, from 1(mostly sad) to 3(mostly happy).

Results

All dependent variables were subject to a 2(Age: preschool- versus school-age) X 4(Speaker Type) mixed model ANOVA.

Speaker choice. Results for children's speaker choices revealed a significant 2-way interaction, $F(3, 120) = 31.06$, $p < .001$, $\eta_p^2 = .44$, which was explored by conducting repeated measures ANOVA for each age group (Table 1). For both age groups, there was a main effect of speaker type (preschool-age: $F(3, 57) = 7.55$, $p < .001$, $\eta_p^2 = .28$; school-age: $F(3, 63) = 34.48$, $p < .001$, $\eta_p^2 = .62$). Paired t-tests (with Bonferroni correction, .05/6 comparisons) were conducted to interpret the data.

Preschool-age. Preschool-age children chose to solicit information from the consistently positive speakers (i.e., Pos-Verbal/Pos-NV), over unknown speakers, to a greater extent than both of the inconsistent speakers (Neg-Verbal/Pos-NV: $t(19) = 3.24$, $p = .004$, $d = .70$; Pos-Verbal/Neg-NV: $t(19) = 3.86$, $p = .001$, $d = 1.03$, as well as the consistently negative speakers, $t(19) = 2.93$, $p = .001$, $d = .90$. There were no other differences between speaker types, ps > .44. One-sample t-tests revealed that preschoolers chose consistently negative speakers (over unknown speakers), as well as both types of inconsistent speakers, less than expected by chance (Neg-Verbal/Neg-NV: $t(19) = 3.56$, $p = .002$, $d = 1.03$; Neg-Verbal/Pos-NV: $t(19) = 2.30$, $p = .03$; Pos-Verbal/Neg-NV: $t(19) = 4.61$, $p < .001$. They chose consistently positive speakers, over unknown speakers, at chance-levels, $p = .44$.

School-age. School-age children chose to solicit information from consistently positive speakers, over unknown speakers, to a greater extent than both inconsistent speakers (Neg-Verbal/Pos-NV: $t(21) = 8.44$, $p < .001$, $d = 2.61$; Pos-Verbal/Neg-NV: $t(21) = 7.09$, $p < .001$, $d = 2.25$), but not more than consistently negative speakers ($p = .05$). Importantly, they also preferred to solicit information from consistently negative speakers, over unknown speakers,

more than both inconsistent speakers (Neg-Verbal/Pos-NV: $t(21) = 5.85$, $p < .001$, $d = 1.79$; Pos-Verbal/Neg-NV: $t(21) = 4.83$, $p < .001$, $d = 1.52$). There was no difference between children's preference for the two inconsistent speakers, $p = .33$. Therefore, school-age children showed a preference for both types of consistent speakers, relative to an unknown speaker, over both types of inconsistent speakers. Further analyses revealed that consistently positive speakers were chosen more often than chance: $t(21) = 3.50$, $p = .002$, while both inconsistent speakers were chosen less often than chance (Neg-Verbal/Pos-NV: $t(21) = 9.76$, $p < .001$; Pos-Verbal/Neg-NV: $t(21) = 7.31$, $p < .001$), suggesting that these speakers were actively avoided. Consistently negative speakers were chosen at chance-levels, $p = .35$.

Comparisons between age groups revealed that, relative to school-age children, preschool-age children were less likely to choose the consistently negative speakers relative to an unknown speaker, $t(40) = 3.18$, $p = .003$, $d = 1.00$), with no other significant differences, ps > .04 (Please see supplemental material for additional Generalized Linear Mixed Model analyses that produced a similar pattern of results).

Speaker belief ratings. Analyses on children's speaker ratings revealed a significant 2-way interaction between age and speaker type, $F(3, 120) = 43.44$, $p < .001$, $\eta_p^2 = .52$ (Table 1). To further explore this interaction, two repeated measures ANOVAs were conducted (one for each age group). For both age groups, there was a significant effect of speaker type (preschool-age: $F(2, 57) = 9.98$, $p < .001$, $\eta_p^2 = .34$; school-age: $F(3, 63) = 61.76$, $p < .001$, $\eta_p^2 = .75$. To interpret these effects, follow up t-tests were conducted with Bonferroni correction.

Preschool-age. Preschool-age children rated the consistently positive speakers as more believable than both of the inconsistent speakers (Neg-Verbal/Pos-NV: $t(19) = 4.62$, $p < .001$, d

= .1.57; Pos-Verbal/Neg-NV: $t(19) = 4.56$, $p < .001$, $d = 1.54$), as well as the consistently

negative speakers, $t(19) = 3.72$, $p = .001$, $d = 1.31$). There were no other differences, ps > .65.

Therefore, preschool-age children rated consistently positive speakers as more believable than

the other three speaker types.

School-age. School-age children rated the consistently positive speakers as more

believable than both of the inconsistent speakers (Neg-Verbal/Pos-NV: $t(21) = 13.28$, $p < .001$, d

= .3.97; Pos-Verbal/Neg-NV: $t(21) = 7.67$, $p < .001$, $d = 2.82$), and the consistently negative

speakers, $t(21) = 3.78$, $p = .001$, $d = .71$. Importantly, they rated consistently negative speakers as

more believable than both the inconsistent speakers (Neg-Verbal/Pos-NV: $t(21) = 9.48$, $p < .001$,

$d = 2.77$; Pos-Verbal/Neg-NV: $t(21) = 6.19$, $p < .001$, $d = 2.02$). There was no difference

between children's ratings of the two inconsistent speakers, $p = .36$. Therefore, school-age

children rated both types of consistent speakers as more believable than both types of

inconsistent speakers.

Of note, compared to school-age children, preschool-age children rated both types of

inconsistent speakers as more believable (Neg-Verbal/Pos-NV speakers: $t(40) = 4.74$, $p < .001$, d

= 1.45; Pos-Verbal/Neg-NV speakers: $t(40) = 3.09$, $p = .004$, $d = .96$, with a trend for

consistently negative speakers to be rated as less believable, $t(40) = -2.53$, $p = .02$, $d = .78$).

Speaker feeling ratings. Importantly, as per children's responses on the Feelings Ratings

Task (Table 1), we did not find a significant effect of age (or an interaction between age and

speaker type) on judgements of speaker affect (ps > .15). This suggests that preschool-age and

school-age children were similar in their ability to determine affect. There was a main effect of

speaker type, $F(3, 120) = 78.85$, $p < .001$, $\eta_p^2 = .66$. Follow-up comparisons (with Bonferroni

correction) revealed that there were significant differences between the affect ratings of all the

speaker types, such that consistently negative speakers were rated as the most sad, next was the the Pos-Verbal/Neg-NV, then Neg-Verbal/Pos-NV, to the consistently positive speakers who were rated as being the most happy, all $ps < .006$.

Discussion

Results showed that school-age children preferred to solicit new information from speakers who demonstrated consistency between the words uttered and the non-verbal cues provided. Specifically, in a context where the goal was epistemic (i.e., children were seeking accurate information), they solicited information from consistently positive speakers as well as consistently negative speakers (over unknown speakers) to a greater extent than inconsistent speakers. Further, consistent speakers were rated as being more believable, suggesting that school-age children are sensitive to the fact that inconsistency renders information less credible.

A developmental progression was observed in the application of this verbal/non-verbal consistency principle: in contrast to school-age children, preschool-age children did not show a general preference for consistency, instead, they tended not to solicit information from consistently negative speakers at an equivalent rate to the inconsistent speakers. Preschoolers' responses suggest that they avoid soliciting information from speakers who provide negative verbal or non-verbal information. It is unlikely that these age-differences can be attributed to affect recognition abilities as there were no age differences in children's ratings of the speakers' feelings. That is, across both ages, children distinguished the affect between the various speaker types. Interestingly, children used the non-verbal affect of the speakers more than the verbal statement when forming these affect judgments. This findings is contrary to the lexical primacy demonstrated in the past (e.g., Morton & Trehub, 2001). It is likely the exaggerated nonverbal cues in the present study made this information more salient (see Eskritt & Lee, 2003).

What remains unclear, however, is *why* school-age children might prefer to solicit information from consistent speakers (and why preschoolers do not). It may be that children generally attribute consistent speakers with globally favourable characteristics without specifically recognizing them as better informational sources. It is also unclear whether the pattern of results demonstrated is specific to the acquisition of accurate information, or whether children would show a preference for a consistent speaker in non-epistemic contexts.

In addition, it is not clear that the school-age children explicitly detected the inconsistencies in messages – and, related, that the preschool performance was due to difficulty with detecting the inconsistency. Certainly, previous work has found that the explicit recognition of inconsistency between verbal/non-verbal information begins around 6 – 7 years of age (e.g., though this work only involved audio information, Morton & Trehub, 2001). In order to clarify these issues, a second experiment was conducted.

Experiment 2

One purpose of Experiment 2 was to examine whether children use speakers' cue consistency in a context where the goal is not about acquiring accurate information. As such, children were presented with the same speaker videos as Experiment 1, but the choices and ratings made by the children were different: children were told that they could choose to keep a sticker (hidden in a cup) that was either the favorite sticker of the speaker or the favorite sticker of another girl they had not been exposed to. Thus, the decision was about whether to use the speakers' personal preference or not. This change allowed us to examine whether children's decisions change when the context is such that there is no 'right' answer. A second goal of Experiment 2 was to assess whether children view consistent speakers in globally favourable ways; as such, children were asked to rate the speakers on attributes that were unrelated to

information-quality, i.e., their friendliness and likeability. Finally, a third goal of Experiment 2 was to examine whether children were sensitive to the differences in speakers' communicative cues; therefore, children were asked to rate speakers on the 'weirdness/trickiness' of their speech.

Method

Participants. Twenty three children aged 4- and 5-years-old (12 males, $M = 61.35$ months, $SD = 6.10$) and 21 children aged 7- and 8-years-old (12 males, $M = 95.29$ months, $SD = 7.86$) were recruited. Five additional children were tested, but their data were not included in the analyses as they did not complete the task. All children possessed language skills sufficient to understand the statements, as assessed by the TOLD-P:3 (Newcomer & Hammill, 1997), a standardized receptive vocabulary task.

Materials and procedure.

Speaker personal preference task. Children were exposed to the same speakers as in Experiment 1 (i.e., who varied in the consistency with which they delivered affective information). Children were told that each speaker's favourite sticker was in a cup. The cups were opaque so the children were not able to see the stickers. The children's task was to listen to each speaker and then decide whether to keep that speaker's favourite sticker or to keep the favourite sticker of another girl they had no exposure to. Thus, we were able to determine whether consistency impacted children's decision to use or not use a speakers' personal preference.

Each of the 12 trials began with the children watching a video-recorded speaker making a statement (i.e., consistent or inconsistent; three trials for each of the four speaker types). Next, children decided whether they wanted to keep the sticker the speaker preferred or the sticker that

another girl preferred (i.e., *speaker choice*; "This girl likes the sticker in this cup best, the other girl likes the sticker in this cup best, which one do you want?"). While hearing these instructions, children saw pictures of the speaker from the video and an unknown girl, with cups beside them, and speech bubbles from each girl saying, "This is my favourite sticker." Importantly, children did not open the cups or see the stickers until the end of the task. In this way, we ensured that participants were not basing their responses on the stickers themselves. Once children made their choice, the cup with the sticker they wanted to keep was placed in a pile that they were told was theirs to keep. Children's choices for stickers were scored as '1' if they chose the sticker from the speaker or '0' if they chose the sticker from the other individual. The trials were counterbalanced/randomized in the same way as they were in Experiment 1.

After their choice, children were asked questions about the speaker: "How much do you like this girl? Not at all, not much, mostly, very much" (*speaker liking*), "How friendly is this girl? Not at all, not much, mostly, very much" (*speaker friendliness*). Ratings ranged from a 1 (not at all) to a 4 (very much) and the verbal instructions of the rating were accompanied by visual scales depicting the options in differently sized bars. Children were also asked to rate the communicative quality of the speakers: "Was there anything weird or tricky about what this girl said? Yes or no" (*speaker weirdness*). They received a score of '1' if children indicated that the speaker's responses were 'weird' and '0' if not.

Results

To examine whether age and speaker type influenced children's speaker choices and ratings 2 (Age: preschool- versus school-age) X 4 (Speaker type) mixed model ANOVAs were conducted for all variables. See Table 2.

Speaker choice. Children's choice of speaker data revealed an interaction between age and speaker type, $F(3, 126) = 4.75$, $p = .004$, $\eta_p{}^2 = .10$). To explore the interaction, two repeated measures ANOVAs were conducted (one for each age group). For both age groups, the effect of speaker type was significant (preschool-age: $F(3, 66) = 3.30$, $p = .03$, $\eta_p{}^2 = .13$; school-age: $F(3, 60) = 4.81$, $p = .005$, $\eta_p{}^2 = .19$).

Preschool-age. Once the Bonferroni correction was applied, none of the preschool-age children's speaker choices were significantly different from one another ($ps > .02$), suggesting that preschool-age children did not interpret any speaker type as having more desirable personal preferences. However, preschool-age children chose the stickers that the Neg-Verbal / Neg-NV speakers liked at less than chance, $t(22) = -2.65$, $p = .01$, $d = .57$, and the rest of the speakers at chance $ps > .30$, suggesting that preschool-age children avoided choosing stickers that consistently negative speakers liked.

School-age. School-age children chose the stickers preferred by consistently positive speakers (over unknown speakers), to a greater extent than the inconsistent Pos-Verbal / Neg-NV speakers: $t(20) = 3.28$, $p = .004$, $d = 1.11$. No other comparisons were significant, $ps > .03$. School-age children chose the stickers that the consistently positive speakers liked more often than chance, $t(20) = 3.24$, $p = .004$, $d = .70$; other speakers were chosen at chance-levels, $ps > .006$.

Comparisons between the age groups revealed that, relative to school-age children, preschool-age children were less likely to choose the consistently positive speakers, $t(42) = 2.42$, $p = .02$, $d = .71$). There were no significant differences between the age groups in their choices of inconsistent speakers or consistently negative speakers, $ps > .06$ (Please see supplemental

material for additional Generalized Linear Mixed Model analyses that produced a similar pattern

of results).

Speaker likeability ratings. Results revealed a main effect of speaker type, $F(3, 126) =$

20.861, $p < .001$, $\eta_p{}^2 = .33$, but no significant interaction between speaker type and age, $p = .25..$

Follow up t-tests were conducted: children gave consistently positive speakers higher likeability

ratings than all other speakers types (Neg-Verbal / Neg-NV: $t(43) = 6.92$, $p < .001$, $d = 1.03$;

Pos-Verbal / Neg-NV: $t(43) = 6.28$, $p < .001$, $d = .89$; Neg-Verbal / Pos-NV: $t(43) = 5.87$, $p <$

.001, $d = .97$). No other speaker types were significantly different from one another, $ps > .10$.

Thus, children rated consistently positive speakers as more likeable and tended to rate speakers

who demonstrated any kind of negativity as less likeable.

Speaker friendliness ratings. Regarding children's ratings of *speaker friendliness*, results

revealed a significant interaction between age and speaker type, $F(3, 126) = 3.86$, $p < .001$, $\eta_p{}^2 =$

.08). To further explore the interaction, repeated measures ANOVAs were conducted for each

age group.

Preschool-age. For the preschool-age children, there was a significant effect of speaker

type, $F(3, 66) = 6.22$, $p = .001$, $\eta_p{}^2 = .22$. Follow-up comparisons revealed that preschoolers rated

the consistently positive speakers as friendlier than all other speakers except the Neg-Verbal/Pos-

NV speakers (Neg-Verbal / Neg-NV: $t(22) = 2.92$, $d = .53$, $p = .008$; Pos-Verbal / Neg-NV: $t(22)$

$= 2.98$, $d = .51$, $p = .007$). Once Bonferoni was applied, none of the other comparisons were

significant, $ps > .012$.

School-age. For the school-age children, there was a significant effect of speaker type, $F(3,$

$60) = 15.09$, $p < .001$, $\eta_p{}^2 = .43$). Paired t-tests (with Bonferroni correction) revealed that school-

age children rated the consistently positive speakers as friendlier than all other speakers (Neg-

Verbal / Neg-NV: $t(20) = 4.91$, $p < .001$, $d = 1.42$; Pos-Verbal / Neg-NV: $t(20) = 7.29$, $p < .001$,

$d = 1.96$; Neg-Verbal / Pos-NV: $t(20) = 5.29$, $p < .001$, $d = 1.24$). No other comparisons were

significantly different from one another, ps $> .15$. Therefore, school-age children were more

likely to judge speakers as friendly if they demonstrated consistently positive cues. In addition,

children perceived speakers showing any negativity (verbal or non-verbal) as less friendly.

Comparisons between the age groups revealed that, relative to preschool-age children,

school-age children rated the consistently positive speakers as more friendly, $t(42) = 2.59$, $p =$

$.01$, $d = .80$), with no other significant differences, ps $> .23$.

Speaker weirdness-of-speech ratings. When analysing children's responses about the

weirdness or trickiness of the speakers' speech, results revealed a significant interaction between

age and speaker type, $F(3, 126) = 12.74$, $p < .001$, $\eta_p^2 = .23$). To further explore interaction, two

repeated measures ANOVAs were conducted (one for each age group). For both age groups,

there was a significant effect of speaker type (preschool-age: $F(3, 66) = 4.96$, $p = .004$, $\eta_p^2 = .18$;

school-age: $F(3, 60) = 37.02$, $p < .001$, $\eta_p^2 = .65$; Please see supplemental material for additional

Generalized Linear Mixed Model analyses that produced a similar pattern of results).

Preschool-age. Once the Bonferroni correction was applied to the comparisons, none of

the preschool-age children's speaker choices were significantly different from one another, ps $>$

$.01$. However, the trend was for the preschoolers to rate the inconsistent speakers as speaking in

a more weird or tricky manner.

School-age. School-age children rated both types of inconsistent speakers as speaking in a

manner that was more weird or tricky than the consistently positive speakers (Pos-Verbal / Neg-

NV: $t(20) = 6.52$, $p < .001$, $d = 2.14$; Neg-Verbal / Pos-NV: $t(20) = 5.68$, $p < .001$, $d = 1.83$), as

well as the consistently negative speakers: (Pos-Verbal / Neg-NV: $t(20) = 7.15$, $p < .001$, $d =$

2.14; Neg-Verbal / Pos-NV: $t(20) = 6.18$, $p < .001$, $d = 1.83$); no other comparisons were significant, ps >.21. This indicates that school-age children detected the inconsistency in speakers' cues.

Comparisons between the age groups revealed that, relative to preschool-age children, school-age children were more likely to indicate that the inconsistent speakers said something weird or tricky (Pos-Verbal / Neg-NV: $t(42) = 4.16$, $p < .001$, $d = 1.26$; Neg-Verbal / Pos-NV: $t(42) = 2.92$, $p = .006$, $d = .87$). There were no significant age group differences in the ratings of the consistent speakers, ps > .28.

Discussion

It was found that within this non-epistemic context children's tendency to use a speaker's preference (relative to an unknown speaker's preference) did not vary based on consistency. There was no difference in the degree to which the consistently negative speaker's preference was used relative to the inconsistent speaker's preference. Further, unlike in Experiment 1, inconsistent speakers were not actively avoided by school-age children (i.e., their preferences were not used fewer times than expected by chance). Rather, the only choice that differed from chance was the consistently positive speakers.

Children did not seem to consider consistency in their attribution ratings; rather, the affect valence (positive versus negative) seemed more important. Both age groups rated the consistently positive speakers higher on likeability than the other speaker types, and there were no differences between their ratings of consistently negative speakers and inconsistent speakers. With regards to ratings of friendliness, school-age children demonstrated the same pattern of results, while preschool-age children's ratings were generally higher for speakers who delivered positive non-verbal affect.

Finally, we found that only the school-age group reliably detected speakers'

inconsistency. While preschoolers' ratings showed a trend towards rating inconsistent speakers

as speaking in a "weird or tricky" manner, it was not significantly so. Moreover, their detection

of the inconsistency in affect cues was significantly less accurate than the school-age children

who showed a clear distinction between ratings of the consistent speakers (i.e., positive and

negative) relative to the inconsistent speakers.

General Discussion

We examined children's use of consistency in affective cues when, first, deciding from

whom to acquire accurate information (Experiment 1) and, second, deciding whose personal

preferences to use (Experiment 2). We also examined children's ratings of consistent versus

inconsistent speakers' believability, likeability, friendliness, and 'weirdness' of speech. Together

results provide insight into the affective cues used by children as cues to credibility within

different contexts, as well as the kinds of attributions children make based on the speakers'

affective cues.

Consistency between the affective cues appears to be an important cue for school-age

children when there is the epistemic goal of gaining accurate information. In Experiment 1, when

deciding whether to solicit information from a speaker they had seen in a video-recording (or an

unknown speaker) school-age children chose to acquire information from consistent speakers to

a greater extent than inconsistent speakers. This was also the case for the consistently negative

speakers, an important finding given that in the past children have been shown to avoid soliciting

information from speakers who are described or demonstrated to display negative behaviour /

affect (Landrum, Mills, & Johnston, 2013; Mascaro & Sperber, 2009). The chance-comparisons

reveal a pattern whereby school-age children actively rejected the opportunity to gain

information from inconsistent speakers. The findings expand upon previous work demonstrating young children's preference for informational sources who demonstrate affect that is consistent with the context (Chow et al., 2008). Interestingly, past work shows that 6 year old children are not attuned to consistency between a speaker's confidence and accuracy when making credibility judgements (unlike adults; Tenney, Small, Kondrad, Jaswal, & Spellman 2011). It may be the case that children are sensitive to affective calibration (as demonstrated here) at an earlier developmental stage than they are to calibration in terms of confidence/accuracy. In the present work, children rated inconsistent speakers as using a more weird/tricky manner of speaking and as being less believable than both consistent speakers, suggesting that the inconsistency in affect cues may be interpreted as a marker of deception. Certainly, past work has demonstrated that by the age of 9-years-old children view individuals who show a mismatch between the valence of the words and their facial affect as being less truthful and predict that the facial expressions of an individual who was lying would differ from his/her statements (Rotenberg et al., 1989). It may also be the case that children are showing sensitivity to the fact that inconsistent affect cues renders information more ambiguous and therefore less reliable. This interpretation would suggest children show sensitivity to violations of Gricean maxims (Grice, 1975), in particular the maxim of Manner in which speakers are expected to avoid ambiguity. Certainly, past work has shown that by 7 years old, children avoid speakers who make ambiguous referential statements (Gillis & Nilsen, 2013). Preschoolers have also shown to be sensitive to violations of other maxims (e.g., Relation, Quality, Quantity; Eskritt, Whalen, & Lee, 2008).

While there was no significant difference between the two consistent speaker types, the consistently positive speakers were chosen at greater than chance levels in an information-gathering context, while the consistently negative speakers were chosen at chance-levels.

Further, the consistently positive speakers were rated as more believable than the consistently negative speakers. Thus, it appears that while inconsistency is an important cue to speaker credibility, the school-age children still showed some evidence of preferring speakers who show positivity (in addition to consistency). This is in keeping with the work of Boseovski (2012), which demonstrates that, in general, children prefer speakers who provide positive information.

School-age children showed flexibility in how they applied the consistency principle both in terms of context and relevant attributes. First, the cues that school-age children used depended on the goals of the task (i.e., epistemic or non-epistemic). More specifically, in Experiment 1 where the goal was to acquire accurate information, children used consistency as a cue to guide their choices. In contrast, in a context that was not about knowledge acquisition (i.e., when children had the opportunity to use the personal preferences of a speaker when choosing a sticker, Experiment 2), consistency did not play a role in children's speaker choices. In this context there was no difference between the rate of choosing consistently negative speakers and inconsistent speakers and they did not clearly reject the inconsistent speakers. Second, when rating speaker attributes, school-age children used consistency cues for those attributes that are important to information acquisition but did not for those attributes unrelated to information acquisition. That is, both consistently positive and consistently negative speakers were rated as having greater clarity of speech and more believability, relative to inconsistent speakers. In contrast, the consistently negative speakers were not rated as being more likeable or friendly. Rather, in these more interpersonal domains, demonstrating positive affect seemed more important and any degree of negative affect (verbal or nonverbal) rendered a speaker less favourable.

Preschoolers did not show the same use of the consistency principle to infer speaker credibility as school-age children did. Even in a context involving an epistemic goal (Experiment 1), all speakers, except the consistently positive speakers, were chosen at less than chance levels and did not differ from each other. The preschool-age children were at the age where they would be starting to recognize that individuals could hold a feeling that was different from what was being displayed (Wellman & Liu, 2004). However, it did not appear that they recognized the inconsistency in cues. Namely, their ratings, albeit showing a trend of detecting inconsistency, did not demonstrate a robust appreciation there was anything 'tricky or weird' about the way the inconsistent speakers spoke (with their detection significantly less pronounced than the school-age children). This finding is comparable to previous work demonstrating that explicit recognition of inconsistency between verbal/non-verbal affective information emerge around 7 years of age (Morton & Trehub, 2001; Rotenberg et al. 1989). Difficulty with detecting inconsistency likely accounts for the preschooler's speaker choices. Instead of relying on the consistency principle, preschoolers showed particular sensitivity to negative information and tended to avoid any type of negative information, non-verbal or verbal (in keeping with a negativity bias; Vaish, Grossmann & Woodward, 2008).

Being sceptical of individuals who display inconsistent cues may be adaptive for children's knowledge acquisition (i.e., to avoid deception or potential miscommunication). Still, there are a number of ways in which inconsistency is intentionally exploited in our language system to achieve specific communicative goals. For example, figurative language relies heavily on a mismatch between the words uttered and the affect with which the words are delivered. Moreover, there are times when a situation may explain why the valence of a speaker's words differ from her non-verbal affect (e.g., saying, "I am going to play my favourite game, soccer" in

a sad voice makes sense if the speaker is looking out the window at stormy weather). As such, it would be useful for future work to examine whether children treat all inconsistent speakers similarly or whether children use the speakers' knowledge of the situational context to potentially accept inconsistent cues as appropriate.

In sum, the findings here highlight that consistency in affect cues is an important factor in determining speaker credibility. By the age of 7 to 8 years old, children are more likely to rely on speakers who show consistency between *what* is said and *how* it is said, and rate these speakers as more believable than inconsistent speakers. These findings contribute to a large body of literature revealing the broad range of cues that children use to decide whether someone is a credible source of information. Moreover, findings highlight the fact that over development children become increasingly sensitive to nuanced cues that are relevant to source credibility, as well as the contexts in which such cues are relevant (or not).

References

Argyle, M., Alkema, F., & Gilmour, R. (1971). The communication of friendly and hostile attitudes by verbal and non-verbal signals. *European Journal of Social Psychology*, *1*, 385-402.

Barrera, M. E., & Maurer, D. (1981). The perception of facial expressions by the three-month-old. *Child development*, 203-206.

Berman, J. M. J., Chambers, C. G., & Graham, S. A. (2010). Preschoolers' appreciation of speaker vocal affect as a cue to referential intent. *Journal of Experimental Child Psychology, 107,* 87 – 99.

Birch, S. A. J., Akmal, N., & Frampton, K. L. (2010). Two-year-olds are vigilant of others non-verbal cues to credibility. *Developmental Science, 13*, 363-369.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International 5,* 341-345

Boseovski, J. J. (2012). Trust in testimony about strangers: Young children prefer reliable informants who make positive attributions. *Journal of Experimental Child Psychology, 111,* 543 – 551.

Brosseau-Liard P, Cassels T., Birch, S. (2014). You seem certain but you were wrong before: Developmental change in preschoolers' relative trust in accurate versus confident speakers. *PLoS ONE 9(9):* e108308. doi:10.1371/journal.pone.0108308

Chiarella, S. S., & Poulin-Dubois, D. (2013). Cry babies and Pollyannas: Infants can detect unjustified emotional reactions. *Infancy*, *18*, 81-96.

Chow, V., Poulin-Dubois, D., & Lewis, J. (2008). To see or not to see: Infants prefer to follow the gaze of a reliable looker. *Developmental Science*, *11*, 761-770.

Corriveau, K. H., Meints, K., & Harris, P. L. (2009). Early tracking of informant accuracy and

inaccuracy. *British Journal of Developmental Psychology, 27*, 331–342.

DePaulo, B. M., Stone, J. I., & Lassiter, G D. (1985). Deceiving and detecting deceit. In B. R.

Schlenker (Ed.), *The self and social life* (pp. 323-370). New York: McGraw-Hill.

Eskritt, M., & Lee, K. (2003). Do actions speak louder than words? Preschool children's use of

the verbal-nonverbal consistency principle during inconsistent communications. *Journal

of Nonverbal Behavior*, *27*(1), 25-41.

Eskritt, M., Whalen, J. & Lee, K. (2008). Preschoolers can recognize violations of the Gricean

maxims. *British Journal of Developmental Psychology, 26,* 435 – 443.

Feldman, R. S., & White, J. B. (1980). Detecting deception in children. *Journal of

Communication, 30,* 121-128

Fernald, A. (1993). Approval and disapproval: Infant responsiveness to vocal affect in familiar

and unfamiliar languages. *Child Development, 64,* 657 – 674.

Friend, M. (2000). Developmental changes in sensitivity to vocal paralanguage. *Developmental

Science*, *3*, 148-162.

Friend, M. (2001). The transition from affective to linguistic meaning. *First Language, 21,* 219

– 243.

Friend, M. (2003). What should I do? Behavior regulation by language and paralanguage in early

childhood. *Journal of Cognition and Development*, *4*, 161-183.

Friend, M., & Bryant, J. B. (2000). A developmental verbal bias in the interpretation of

discrepant messages. *Merrill-Palmer Quarterly, 46,* 342-369.

Gillis, R. L., & Nilsen, E. S. (2013). Children's use of information quality to establish speaker

preferences. *Developmental Psychology*, *49*, 480-490.

Grice, H. P. (1975). Logic and conversation, In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics* (Vol. 3, pp. 41 – 85). New York, NY: Academic Press.

Harris, P. L. (2002). What do children learn from testimony? In P. Carruthers, S. Stich, and M. Siegal (Eds.), *The Cognitive Basis of Science* (pp. 316-334). Cambridge University Press.

Koenig, M. A., Clément, F., & Harris, P. (2004). Trust in testimony: Children's use of true and false statements. *Psychological Science, 15*, 694-698.

Koenig, M. A., & Harris, P. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child Development, 76,* 1261-1277.

Kuchuk, A., Vibbert, M., & Bornstein, M. H. (1986). The perception of smiling and its experiential correlates in 3-month-olds. *Child Development*, *57*, 1054-1061.

Landrum, A. R., Mills, C. M., & Johnston, A. M. (2013). When do children trust the expert? Benevolence information influences children's trust more than expertise. *Developmental Science, 16*, 622-638.

Mascaro, O., & Sperber, D. (2009). The moral, epistemic, and mind reading components of children's vigilance towards deception. *Cognition, 112,* 367–380.

Mehrabian, A., & Ferris, S. R. (1967). Inference of attitudes from nonverbal communication in two channels. *Journal of Consulting Psychology*, *31*, 248-252.

Mills, C. M. (2013). Knowing when to doubt: Developing a critical stance when learning from others. *Developmental Psychology*, *49*, 404-418.

Morton, J. B., & Trehub, S. E. (2001). Children's understanding of emotion in speech. *Child Development*, *72*, 834-843.

Newcomer, P. L. & Hammill, D. D. (1997). *Test of language development-primary 3 (TOLDP:3).* Austin, TX: ProEd.

O'Grady, W. (2005). *How children learn language*. Cambridge University Press.

R Development Core Team (2011). R: A language and environment for statistical computing. R

Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL

http://www.R-project.org/.

Reilly, S. S., & Muzekari, L. H. (1986). Effects of emotional illness and age upon the resolution

of discrepant messages. *Perceptual and Motor Skills, 62,* 823–829.

Ridgeway, D., Waters, E., & Kuczaj, S. A. (1985). Acquisition of emotion-descriptive language:

Receptive and productive vocabulary norms for ages 18 months to 6

years. *Developmental Psychology*, *21*(5), 901.

Rotenberg, K. J., Simourd, L., & Moore, D. (1989). Children's use of a verbal-nonverbal

consistency principle to infer truth and lying. *Child Development, 60,* 309-322.

Scofield, J., & Behrend, D. A. (2008). Learning words from reliable and unreliable speakers.

Cognitive Development, 23, 278-290.

Seuss, Dr. (1966) How the Grinch Stole Christmas! TV Short, Retrieved 20 February 2014 from

http://www.imdb.com/title/tt0060345/quotes.

Solomon, D., & Ali, F. A. (1972). Age trends in the perception of verbal reinforcers.

Developmental Psychology, *7*, 238-243.

Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: the

negativity bias in social-emotional development. *Psychological Bulletin*, *134*, 383-403.

Waxer, M., & Morton, J. B. (2011). Children's judgments of emotion from conflicting cues in

speech: Why 6-year-olds are so inflexible? *Child Development*, *82*, 1648-1660.

Wellman, H. M., & Liu, D. (2004). Scaling of Theory-of-Mind Tasks. *Child development*, *75*,

523-541.

Table 1

Children's speaker choices and speaker ratings in the Speaker Affect Task and Feelings Rating Task Experiment 1

	Speaker Choice		Speaker Belief Ratings		Speaker Feeling Ratings	
	Proportion of speaker choices (SD)		Mean speaker ratings 1 – 4 (SD)		Mean speaker ratings 1 – 3 (SD)	
	Preschool-age	School-age	Preschool-age	School-age	Preschool-age	School-age
Consistent Speakers						
Positive Verbal/ Positive Non-verbal	.57 (.38)	.69 (.25)	3.49 (.57)	3.44 (.53)	3.00 (.00)	2.97 (.10)
Negative Verbal/ Negative Non-verbal	.27 (.28)	.56 (.30)	2.28 (1.18)	3.02 (.65)	1.15 (.20)	1.06 (.22)
Inconsistent Speakers						
Negative Verbal/ Positive Non-verbal	.30 (.39)	.11 (.19)	2.40 (.80)	1.45 (.47)	2.45 (.64)	2.27 (.75)
Positive Verbal/ Negative Non-verbal	.24 (.25)	.15 (.22)	2.37 (.86)	1.59 (.76)	1.74 (.73)	1.77 (.79)

Table 2

Children's speaker choices and speaker ratings, Experiment 2

	Speaker Choice		Speaker Likeability Ratings		Speaker Friendliness Ratings		Speaker Weirdness-of-speech Ratings	
	Proportion of speaker choices (SD)		Mean speaker ratings 1 – 4 (SD)		Mean speaker ratings 1 – 4 (SD)		Proportion of 'weird' rating (SD)	
	Preschool	School	Preschool	School	Preschool	School	Preschool	School
Consistent Speakers								
Positive Verbal/ Positive Non-verbal	.51 (.26)	.71 (.30)	3.23 (.80)	3.48 (.49)	3.23 (.81)	3.73 (.36)	.06 (.13)	.03 (.10)
Negative Verbal/ Negative Non-verbal	.33 (.30)	.46 (.31)	2.57 (1.00)	2.63 (.46)	2.72 (1.08)	2.94 (.70)	.07 (.14)	.03 (.10)
Inconsistent Speakers								
Negative Verbal/ Positive Non-verbal	.57 (.29)	.43 (.41)	2.75 (.88)	2.63 (.68)	3.17 (.92)	2.97 (.79)	.26 (.35)	.60 (.43)
Positive Verbal/ Negative Non-verbal	.55 (.33)	.37 (.31)	2.60 (1.01)	2.40 (.67)	2.75 (1.04)	2.70 (.65)	.20 (.33)	.67 (.41)

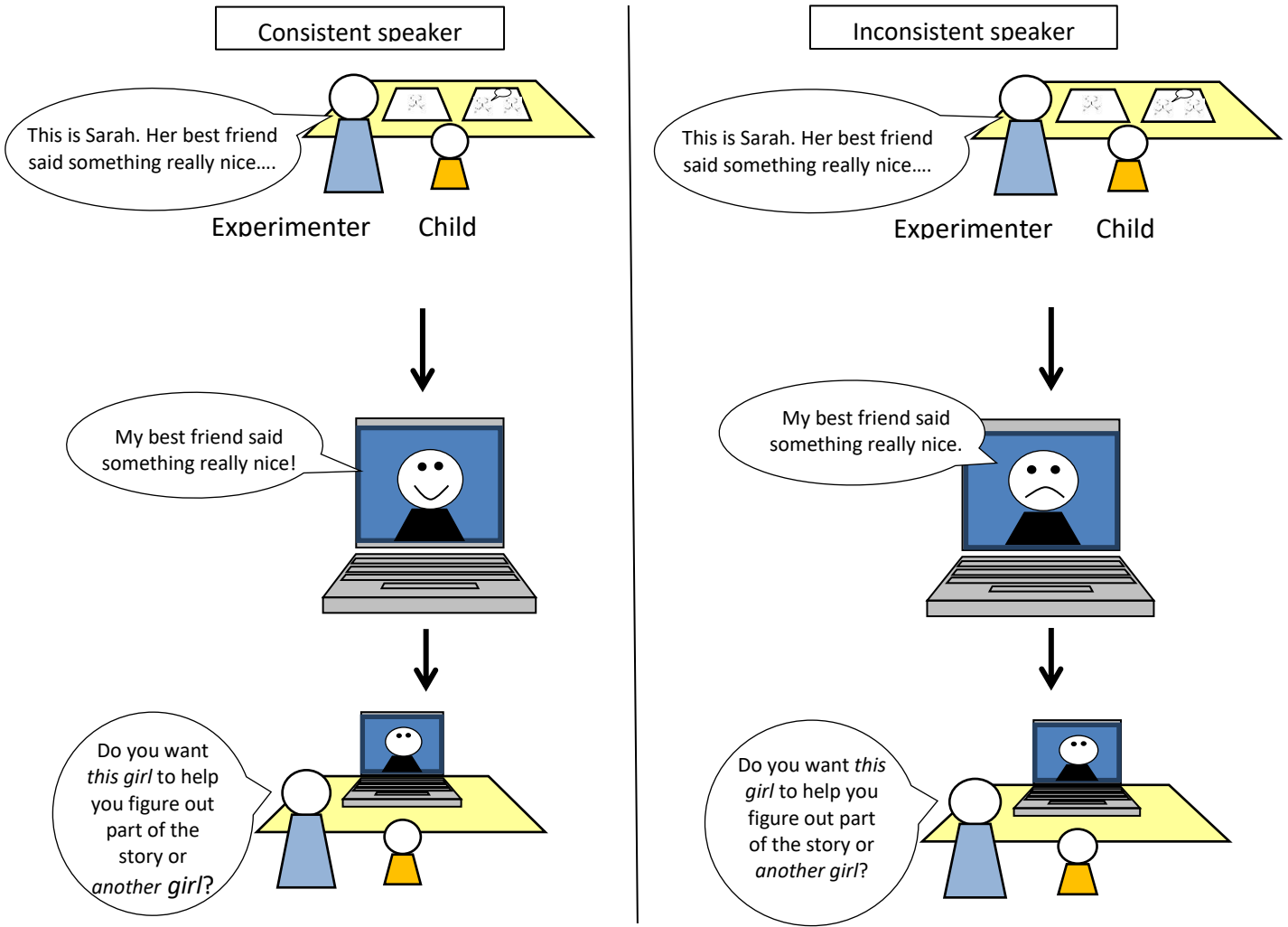*Figure 1*. Example of inconsistent (e.g., Pos-Verbal/Neg-NV) and consistent (e.g., Pos-Verbal/Pos-NV)

speaker trials, Experiment 1