Research Article

The Development of the Liking Gap: Children Older Than 5 Years Think That Partners Evaluate Them Less Positively Than They Evaluate Their Partners

Wouter Wolf[1], **Amanda Nafe**[1], **and Michael Tomasello**[1,2]

[1]Department of Psychology and Neuroscience, Duke University, and [2]Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

Abstract

After two strangers have briefly interacted with one another, both believe that they like their partner more than their partner likes them. A plausible explanation for this *liking gap* is that people are constantly worrying about how others are evaluating them. If so, one would expect the liking gap to emerge in young children as they become more concerned with their reputations and the impression they make on other people. The current study (N = 241 U.S. children; age range = 4–11 years) supported this hypothesis, showing a liking gap beginning when children were 5 years old, the age at which they first become concerned with other people's evaluations of them. Moreover, the liking gap became more pronounced as children got older. These findings provide the first developmental description of the liking gap and support the hypothesis that this phenomenon is related to individuals' concerns for how others evaluate them.

People generally tend to hold unrealistically favorable views about themselves and their individual abilities (e.g., Alicke, 1985; Kruger & Dunning, 1999), but this optimism does not extend to their perception of their social lives. We tend to think that others have richer and more active social lives than we do (Deri et al., 2017), are more connected to others than we are (Whillans et al., 2017), and have less desire to connect with us than we do with them (Epley & Schroeder, 2014). This social skepticism not only affects our holistic perception of our social world but also influences our perception of individual social interactions in our daily lives.

Boothby et al. (2018) provided an especially striking demonstration of this skepticism. After two strangers interacted with one another briefly, both reported liking the other person more than they thought the other person liked them, a social illusion referred to as the *liking gap*. This tendency to underestimate how positively one is evaluated by another individual even persists beyond the initial interaction and can be found in established social relationships a year after first meeting the other person (Boothby et al., 2018). Boothby et al. proposed that the liking gap is the consequence of individuals during interactions constantly monitoring others' evaluating them. In doing so, people worry about being evaluated negatively, perhaps particularly

Corresponding Author:
Wouter Wolf, Duke University, Department of Psychology and Neuroscience
E-mail: wouter.wolf@gmail.com

when interacting with a stranger. This anxiety might make negative thoughts more salient or relevant, causing individuals to become more skeptical about the impression they make on others.

If the liking gap is indeed driven by people's concern for how they are evaluated by others, then, in childhood, the liking gap should emerge in parallel with the emergence of children's concern for their reputation and the impression they make on other people. Children start to manage their reputations, implicitly at first, around 5 years of age (Engelmann & Rapp, 2018; Silver & Shaw, 2018). When observed by a same-age peer, 5-year-olds are less likely to steal stickers (Engelmann et al., 2012), and they steal less when they are told that they are being watched by an invisible adult (Piazza et al., 2011). Additionally, 5-year-olds refrain from cheating more often if they have a positive reputation to maintain, whereas 4-year-olds do not (Fu, Heyman, Qian, et al., 2016). In addition to shaping their reputation by avoiding behavior they think will be negatively perceived, 5-year-old children also engage in positive behaviors to shape their reputation. When 5-year-olds were asked how many stickers to give to another individual (with no cost to themselves), they were more likely to choose the generous option when the recipient could see them than when the recipient could not see them (Buhrmester et al., 1992; Leimgruber et al., 2012). Moreover, similar increases in 5-year-olds' generosity has been seen in response to being observed by a neutral, unaffected peer (Fujii et al., 2015). Interestingly, 5-year-olds become particularly prosocial when the recipient either can reciprocate later or is part of their (minimal) in-group (Engelmann et al., 2012).

There is one study in which, under very specific circumstances, 4-year-olds were found to engage in some form of reputation management as well (Rapp et al., 2019). In this study, the children were in the room with three peers, each of whom was asked to donate stickers to another group of children at the same time. The children were separated by dividers and could not see each other, but at the end of the donation period, the name of the highest or lowest donor would be displayed on a screen visible to them all. Compared with children in a control condition, both 4- and 5-year-old children donated more when they believed that the name of the child donating the most stickers would be displayed. However, the experimental setup in this study allowed children to directly compare themselves with their peers, perhaps introducing a competitive or gamelike component to the game. Overall, then, the general consensus in the literature is that behavior motivated by reputational concerns reliably emerges in most circumstances when children are around 5 years old.

Statement of Relevance

We humans spend much of our time worrying about what others think of us. What is striking is that, at least by adulthood, we experience a liking gap: We inaccurately believe that others like us less than they actually do. In this study, we investigated the developmental origins of the liking gap. We found that children do not seem to experience the gap until they are around 5 years old. Independent evidence indicates that age 5 is when children start monitoring how other people are evaluating them. We also found that the liking gap increased in magnitude from age 5 through middle childhood (age 11). It is probable that the liking gap is linked to processes of self-presentation, impression management, and conservation of the self, all of which increase over childhood into adolescence. These concerns shape our social reality by influencing the way we think others evaluate us and, ultimately, the way we evaluate ourselves.

After reputation management starts to emerge at 5 years of age, these skills become more sophisticated over the following years. During this time, children gain a better understanding of self-representational motives (Banerjee, 2002; Banerjee et al., 2012). Between 6 and 11 years, they become more concerned with appearing fair (Shaw et al., 2014) and modest (Watling & Banerjee, 2007). They also gain a better grasp of how to strategically engage in self-representational behavior, doing so especially reliably around age 8 (Aloise-Young, 1993; Banerjee & Yuill, 1999)—although the way in which they do so seems to vary by culture (Fu, Heyman, Cameron, & Lee, 2016). This is also the point during development when children start to engage in competitive altruism when other people are watching (Herrmann et al., 2019).

If the liking gap is indeed driven by reputational and impression-management concerns, then one would expect this phenomenon to have a similar ontogenetic trajectory: emerging at 5 years of age, followed by continuous development through middle childhood. To test this hypothesis, we conducted a study in which we examined the liking gap in children between 4 and 11 years old who were initially unfamiliar with one another. Participants engaged in a brief interaction (i.e., building a tower together), after which they were asked in several ways to indicate how positively they felt about their partner as well as how positively they thought their partner felt about them.

Table 1. Distribution of Participants per Age Group and Dyad and the Average Age Difference per Age Group

Age (years)	Dyad gender composition				Average age difference (years)[a]
	Girl–girl	Boy–boy	Mixed	Total	
4	10	9	16	35	−0.29
5	14	11	9	34	0.18
6	19	11	15	45	0.19
7	9	12	12	33	−0.03
8	13	11	15	39	0.15
9	6	9	12	27	−0.21
10	5	10	5	20	0.33
11	1	5	2	8	0.52

[a]A negative number indicates that the children in this age group were, on average, younger than their partners when their partners were from other age groups.

Method

Participants and design

In total, 260 children (130 pairs; 4–12 years old) participated in this study, which was approved by the institutional review board of Duke University. Participants were paired on the basis of age (birth dates within 18 months of each other). Participants were tested at a local museum in Durham, North Carolina, as well as at various events throughout this same area. We tested 132 boys and 128 girls (58.0% White, 15.3% mixed race, 9.6% Hispanic, 6.9% Asian, 3.4% Black, and 1.5% Middle Eastern; 5.0% did not report race). Approximately 60% of families reported an annual income of $60,000 and above; roughly another 22% of our sample opted not to disclose this information.

Of the original sample, five dyads (i.e., 10 participants) were excluded from analyses because they did not complete the task (1 dyad) or because parents reported that their child was diagnosed with autism or a developmental delay (4 dyads). In addition, some participants were excluded because they did not seem to understand the questions in the dependent measure ($n = 5$). During data collection, it appeared that it was particularly difficult to recruit 12-year-olds, resulting in only three 12-year-olds in the sample. Because of the limited sample size for this age group, we decided not to include the 12-year-olds in our analyses. Finally, one participant showed an extreme dislike for their partner, resulting in a liking-gap score 2.5 times more extreme than the second-most extreme score on that side of the distribution. Because this made us concerned that something unusual might have happened during the procedure, we analyzed the data with and without this participant, yielding similar results. We report our analyses without this participant because including their data caused a high level of kurtosis in the distribution of liking-gap scores. Consequently, the final sample consisted of 241 children.

Table 1 shows the distribution of dyad gender composition within each group for the final sample.[1]

Procedure

When parents with a child passed by a testing location where two experimenters were present, the parents were asked whether they were willing to have their child participate in the study. If so, we informed the parent (and the child) that we were looking for two children around the same age who did not know each other to participate in a study in which they would build a tower together for 5 min. They were told that after the task, an experimenter would ask the child a couple questions about their experience, after which the child would be able to pick out a prize in exchange for their help. Finally, the parents were told that instructions and the interaction (but not the answering of the questions) would be video recorded to make sure that no experimenter errors occurred and that no outside factors in the testing area interfered with the procedure. If the parent and child agreed to participate, one experimenter started to go through the informed-consent and demographic forms with the parent, while the other would look for a second child in the appropriate age group to complete the dyad.

After consent was obtained from both parents, Experimenter 1 started the cooperative tower-building task. In the original Boothby et al. (2018) paradigm, participants engaged in conversation for 5 min. Holding a conversation for 5 min without having anything else to do, however, can be challenging for younger children. We therefore presented them with a tower-building task specifically designed to serve as a vehicle for social interaction, while making sure that they did not perceive the task as competitive or that they were being judged on performance. To do so, we presented them with the following instructions:
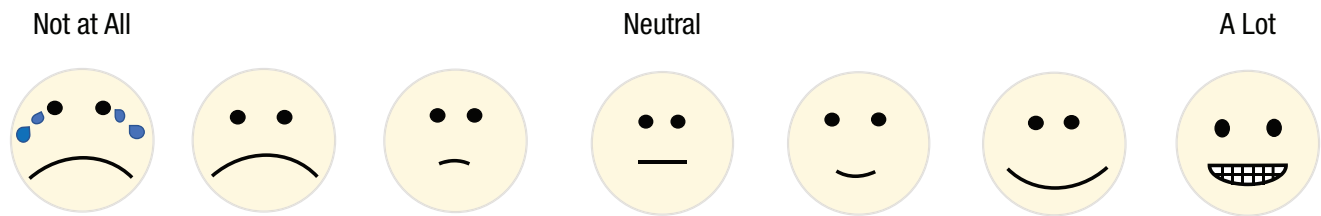
Not at All   Neutral   A Lot

Fig. 1. Smiley-face Likert-type scale that children used to answer the six questions about their partner in the cooperative task.

The two of you are going to build a tower together. You can build any kind of tower that you want, but the important part is that you do it together. Okay? You will have 5 minutes to build the tower and at the end of these 5 minutes we will ask you a couple questions about how it went and then get you your prize. Okay you are all set to get started building your tower together.

At 4 min into the play session, the children were told that they had 1 min remaining. If participants said they had completed the task before the 5-min mark, Experimenter 1 would let them know they still had time left and would suggest using that time to make the tower better. Experimenter 1 and Experimenter 2 typically watched some of the play session but also tried to look somewhat preoccupied in order to create an environment for the children in which they felt safe as well as independent with regard to their task. After 5 min, Experimenter 1 stopped the camera and directed one child to come with them and the other child to go with Experimenter 2. The experimenters were both as far as possible from each other and other community members in order to a create a private space for the child to answer questions. Next, children answered questions about their experience, after which they were thanked and received their prize.

Measures

We asked each child six questions about their partner in the cooperative task. It was explained to the participants that all questions would be read out loud by the experimenter for all children at all ages (even though some older participants indicated that they could read the questions themselves). Answers were given on a tablet using a 7-point smiley-face Likert-type scale (see Fig. 1). Furthermore, although the questions were communicated verbally, participants were asked to select the answer on the tablet themselves, even if they also gave a verbal answer.

To assess the degree to which children comprehended the nature of the Likert scale, we used a non-task-related practice question first. That is, before

engaging the child with the liking-gap questions, the experimenters first asked the child whether they liked ice cream and, if so, to indicate on the Likert scale how much. If the child's verbal answer was not similar to what they had selected on the screen, the experimenter reviewed the answer to make sure the child had indeed selected what they had intended. For example, if a child said they loved ice cream but selected the face with the smallest smile, the experimenter would say, "So that smiley face says you like it, but you only like it a little, see how the face is just a little bit happy? Is that how you feel about ice cream?" If the child did not respond to this explanation and kept pressing a button that was incongruent with their verbal answer, the experimenter made a note in the data file so that this participant's data would be excluded from analysis.

Once the child's comprehension of the Likert scale was assessed, the child was asked two sets of three questions, which were based on the initial questions from the liking-gap study for adults (Boothby et al., 2018) but adapted for children. The order in which the sets of questions were presented to each participant was counterbalanced within the dyad. One set of questions was aimed at measuring the participants' feelings toward their partner after interacting with them. These questions included, "How much do you like the other boy/girl?" "How much would you like to play with the other boy/girl again?" and "How much would you like the other boy/girl to be your friend?" (Cronbach's α = .73). The other set of questions assessed how the participants thought their partner felt about them. These questions included, "How much did the other boy/girl like you?" "How much would the other boy/girl like to play with you again?" and "How much would the other boy/girl like to be friends with you?" (Cronbach's α = .74). The sets of questions were then compiled into a partner-evaluation scale and a scale indicating how the participant thought their partner had evaluated them. Next, a liking-gap score for each participant was created by subtracting the participant's score indicating how they thought their partner evaluated them from that participant's score indicating how they evaluated their partner.

Table 2. Comparison of Models Including Age, Age Difference, and Dyad Gender Composition

Model	P(M)	P(M|data)	BF$_M$	BF$_{10}$	Error %
Age	0.077	0.554	14.889	1.000	
Age + age difference	0.077	0.236	3.705	0.426	0.005
Age + age difference + (Age × Age Difference)	0.077	0.101	1.354	0.183	0.004
Null	0.077	0.039	0.487	0.070	0.003
Dyad gender composition + age	0.077	0.028	0.351	0.051	1.364
Age difference	0.077	0.018	0.225	0.033	0.005
Age + dyad gender composition + age difference	0.077	0.012	0.148	0.022	1.214
Age + dyad gender composition + age difference + (Age × Age Difference)	0.077	0.005	0.062	0.009	0.930
Dyad gender composition	0.077	0.002	0.022	0.003	0.025
Age + dyad gender composition + (Dyad Gender Composition × Age)	0.077	0.002	0.020	0.003	0.923
Age + dyad gender composition + age difference + (Dyad Gender Composition × Age)	0.077	0.001	0.014	0.002	17.682
Dyad gender composition + age difference	0.077	8.396×10^{-4}	0.010	0.002	1.565
Age + dyad gender composition + age difference + (Dyad Gender Composition × Age) + (Age × Age Difference)	0.077	4.769×10^{-4}	0.006	8.613×10^{-4}	1.244

Note: P(M) is the likelihood of the prior. P(M|data) is the likelihood of the posterior given the data. BF$_M$ is the Bayes factor (BF) for the model in each row compared with the average BFs of the other models. BF$_{10}$ is the BF in favor of the alternative over the null hypothesis. Error % shows the margin of error in the Markov chain Monte Carlo approximation of the BF.

Results

The standardized residuals of the liking-gap scores, controlled for age (skewness = .55; kurtosis = .775), indicated that the data were appropriate for using standardized Bayesian statistical models with normal likelihoods and uninformed priors in JASP (Version 0.12.2; JASP Team, 2020) and SPSS Version 26. Because participants engaged in the task in dyads, it was necessary to test whether the data within dyads were consequentially nonindependent. Because the intraclass correlation (ICC) of the liking-gap scores within the dyads was lower than .45 (ICC = .104), we concluded that there was no consequential nonindependence in the data (Kenny et al., 2006). We therefore analyzed the data at the individual level.

Given that, across different age groups, there was an unequal distribution of the different types of dyads as well as age differences relative to the participants' partners (see Table 1), we wanted to make sure that these factors did not confound or moderate the effect of age on the liking-gap scores. We therefore conducted a Bayesian generalized linear model with age, age difference, and dyad gender composition as independent variables and the liking-gap score as the dependent variable. We then compared a basic model with only age as a predictor variable with models that also included age difference, dyad gender composition, and their interactions as predictors. The Bayes factors (BFs) in Table 2 (i.e., BF$_M$ and BF$_{10}$) show that the model with only age as a predictor fitted the data better than models that also included age difference, dyad gender composition, and any of their interactions. This suggests that there was no effect of age difference and dyad gender composition on the liking-gap scores and that these variables did not confound or moderate the effect of age on liking-gap scores. Subsequent analyses therefore did not include age difference and dyad gender composition as predictor variables.

A Bayesian linear regression with age predicting liking-gap scores showed a positive effect of age (Bayesian estimate of standardized regression coefficient: mode = 0.062, 95% highest-density interval [HDI] = [0.023, 0.101]),[2] meaning that the tendency of children to report that they liked their partner more than they thought their partner liked them (i.e., the liking gap) was more extreme for older children than for younger children. To get a better idea about whether this was a gradual trend or whether this tendency emerged at a specific age, we placed participants into different 1-year age groups (e.g., 4-year-olds, 5-year-olds) and conducted a Bayesian analysis of variance (ANOVA) with age group as a factor and liking-gap score as a dependent variable. As Table 3 and Figure 2 show, there was no difference between how positively 4-year-olds reported feeling about their partner and how positively they thought their partner evaluated them (i.e., the Bayesian 95% HDI of the liking-gap score contains zero). However, in all other age groups, children on average evaluated their partner more positively than they thought their partner evaluated them (although the data from the 11-year-olds need to be interpreted with caution because of the limited sample size in that age group). This suggests that the liking gap in children emerges when they are between 4 and 5 years old.

Table 3. Unstandardized Bayesian Parameter Estimates for Liking-Gap Scores in Different Age Groups

Age (years)	Posterior			95% highest-density interval
	Mode	*M*	Variance	
4	0.076	0.076	0.011	[−0.126, 0.279]
5	0.324	0.324	0.011	[0.118, 0.529]
6	0.281	0.281	0.008	[0.103, 0.460]
7	0.384	0.384	0.011	[0.175, 0.592]
8	0.316	0.316	0.010	[0.124, 0.508]
9	0.481	0.481	0.014	[0.251, 0.712]
10	0.583	0.583	0.019	[0.315, 0.851]
11	0.500	0.500	0.047	[0.076, 0.924]

We also examined whether the overall effect of age on the liking gap was exclusively driven by its emergence at age 5 or whether there was still an effect of age on liking-gap scores after age 5. A Bayesian linear regression on the data for the 5- to 11-year-olds with age predicting liking-gap scores still showed a positive effect of age (Bayesian estimate of standardized regression coefficient: mode = 0.049, 95% HDI = [0.002, 0.097]). This suggests that even after the liking gap emerges at age 5, it becomes more extreme as children grow older.

Finally, we looked at whether changes in the liking-gap scores at different ages were caused by changes in how the participants evaluated their interaction partner or changes in how they thought their interaction partner evaluated them. A Bayesian linear regression showed no effect of age on how participants evaluated their interaction partner (Bayesian estimate of standardized regression coefficient: mode = −0.011, 95% HDI = [−0.063, 0.041]) and a credible negative relationship between age and how participants thought their interaction partner evaluated them (Bayesian estimate of standardized
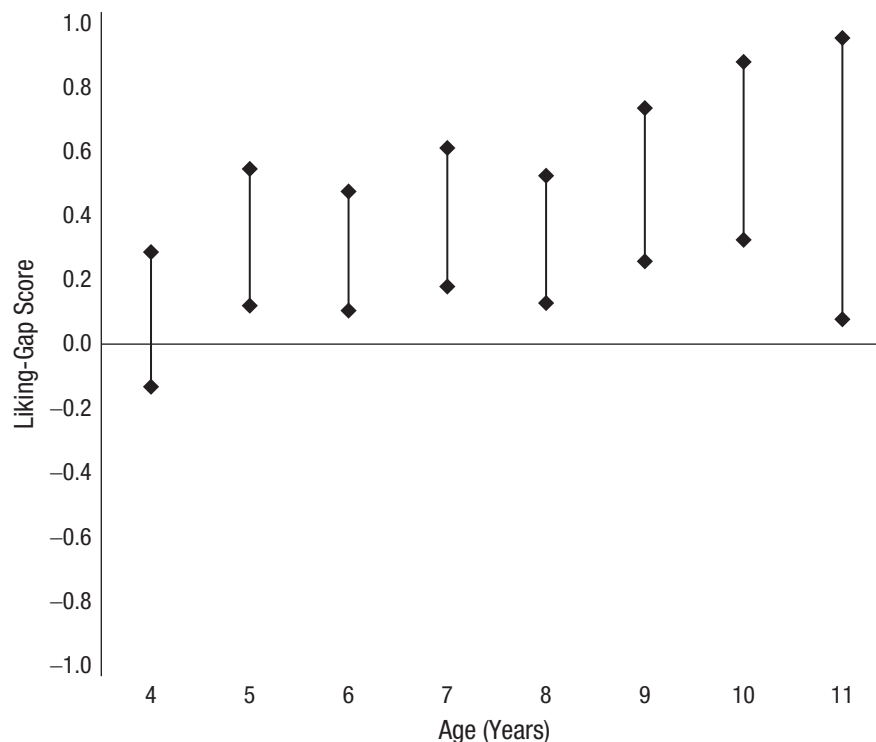

Fig. 2. Liking-gap score as a function of age group. The graph shows Bayesian 95% highest-density intervals, which indicate the 95% most likely values of the average liking-gap score for each age group.

regression coefficient: mode = −0.073, 95% HDI = [−0.127, −0.020]). However, when we looked specifically at the change between the 4-year-olds and the 5-year-olds, where the liking gap first appeared, a Bayesian between-subjects *t* test showed that children's evaluation of their partner became more positive (Bayesian estimate of between-age difference: mode = 0.398, 95% HDI = [0.078, 0.717]), whereas their perception of how their partner evaluated them did not change (Bayesian estimate of between-age difference: mode = 0.150, 95% HDI = [−0.225, 0.526]). In contrast (and in line with the overall trend), for the 5- to 11-year-olds, Bayesian regressions showed no effect of age on how participants evaluated their interaction partner (Bayesian estimate of standardized regression coefficient: mode = −0.045, 95% HDI = [−0.111, 0.021]) and a credible negative effect of age on how participants thought their interaction partner evaluated them (Bayesian estimate of standardized regression coefficient: mode = −0.095, 95% HDI = [−0.162, −0.028]). This suggests that the emergence of the liking gap when children are between 4 and 5 years old and its widening (i.e., becoming more extreme) when they are between 5 and 11 years old are driven by different psychological mechanisms.

Discussion

The current results show that the liking gap emerges when children are around 5 years old. At age 4, children did not evaluate their partner more positively than they thought their partner evaluated them, whereas children who were 5 years old or older did. This effect was not influenced by the gender composition of the dyad or the age difference between participants. Additionally, we found that, after its emergence, the liking gap became more pronounced between ages 5 and 11.

Additional analyses showed that the initial emergence of the liking gap was driven by 5-year-olds evaluating their partner more positively than 4-year-olds did, while their perceptions of how their partner evaluated them did not change between 4 and 5. In contrast, the widening of the liking gap (i.e., the gap becoming more extreme) between 5 and 11 years was driven by children perceiving their partner's evaluation of them as less positive (while their evaluation of their partner remained the same). These findings suggest that there are two important developmental processes underlying the emergence and development of the liking gap during childhood.

First, between 4 and 5 years old, children's evaluations of strangers after a single interaction become more positive. This is the age at which children are coming to understand others' beliefs (theory of mind), but this social-cognitive development is unlikely to be the key factor in the initial emergence of the liking gap, as it seems to be driven by a change in children's attitudes toward others, not their perception of others' attitudes toward them. A more plausible explanation is that between ages 4 and 5, many children are rapidly exposed to interactions with novel individuals, in particular same-age peers, for example, because of mandatory kindergarten enrollment at age 5 in their local public school system. This increased exposure to interactions with peers might contribute to a decrease in stranger anxiety between these ages and a consequent increase in the social value of interacting with novel peers and the social relationships that could emerge from them. Yet although 5-year-olds seem to evaluate novel individuals more positively than do 4-year-olds, their perceptions of how other people evaluate them do not match this trend, suggesting that an increase in the social value of interacting with novel individuals does not automatically bring with it a heightened concern for how others are evaluating oneself.

Second, after the initial liking gap has emerged at age 5, children become increasingly concerned with how other people are evaluating them (while their evaluation of those others remains the same). This suggests that after the emergence of the liking gap between ages 4 and 5, its subsequent development is primarily driven by increased social concern with other people's evaluations of the self. A likely explanation here is more cognitive—namely, that children's understanding of their interaction partners' self-presentational motives increases after they are 5 years old (Aloise-Young, 1993; Banerjee, 2002; Banerjee et al., 2012; Herrmann et al., 2019; Watling & Banerjee, 2007), and this shapes their interpretation of their partners' behavior toward them. Before understanding that other people have self-reputational concerns of their own, children might interpret a partner's positive social behavior during an interaction as a manifestation of a genuinely positive attitude toward them. However, as children get older, they start to understand that their interaction partners' friendly behavior might be motivated by their desire to make a good impression (e.g., appearing friendly and likeable) and that this behavior is therefore not necessarily indicative of that partner evaluating them positively. Consequently, children's increased understanding of other people's self-reputational concerns and impression-management strategies might leave them with more uncertainty when trying to infer from their partner's behavior how positively they are being evaluated.

If this is the case, the question remains how children's social-cognitive development underlies the emergence of phenomena related to impression management and reputational concerns during middle childhood (i.e., after age 5). One potentially important social-cognitive

skill that develops between the ages of 5 and 11 is the capacity to use second-order theory of mind, which allows children to process more complex recursively structured social information. For example, at age 6, children have been shown to understand and express not only that other people hold beliefs but also that other people hold beliefs about their (i.e., the children's) beliefs (Grueneisen et al., 2015; Perner & Wimmer, 1985). In a similar fashion, after age 6, children might start to understand not only that their interaction partner holds an attitude toward them (e.g., evaluates them positively or negatively) but also that their partner holds an attitude toward them having an attitude toward their partner. In other words, second-order theory of mind might be a prerequisite for understanding that other people worry about what you think of them. Consequently, the capacity to engage in higher-order theory of mind, emerging at around age 6, might be the social-cognitive starting point for understanding other people's self-representational motives at that age and may have contributed to the increased social skepticism and widening of the liking gap between ages 5 and 11 in the current study.

Overall, it seems that our findings on the development of the liking gap are consistent with previous research on the development of children's concerns for their reputation and the impression they make on other people. Crucially, this suggests that, as Boothby et al. (2018) proposed, the liking gap is indeed the result of individuals' concern for the impression they are making on others during interactions, perhaps especially with strangers. It appears that around the age at which children begin to worry about their reputation and start to manage the impression they make on other people, their perception of how others are evaluating them becomes more negative relative to how positively they are evaluating their partner. Thus, children from 5 years old and adults might both experience the social anxiety inherent to interacting and thus being evaluated by others, potentially causing negative thoughts to become more salient and relevant, which, in turn, causes individuals to become more skeptical about what others think of them.

The question remains to what extent these findings can be generalized to different populations. Although recruitment at a museum allowed us to include a broader audience than is typically the case in laboratory samples, the majority of parents described their socioeconomic status as at least middle class. More broadly speaking, the research was conducted in a midsized U.S. city, meaning that one should be cautious when generalizing the existence and development of this phenomenon to non-Western populations (Henrich et al., 2010; Nielsen et al., 2017). In addition, the question

remains to what extent individual differences moderate the existence of the liking gap. It is plausible that individual differences impact the way children engage in or appreciate interactions. In this regard, it would be particularly important to look at the relationship between individual differences in children's liking gaps and socially anxious dispositions and attachment styles. It is not implausible that, in some cases, shyness, social anxiety, or insecure attachment could be a manifestation or a consequence of a relatively high discrepancy between how much children like other people and how much they think other people generally like them back.

Furthermore, because the task and, probably, the dependent measure of the current study were not designed to be suitable for individuals older than 11 or 12, how the liking gap develops through late childhood into early adolescence remains an open question. Research suggests that children's and adolescents' concern for reputation and status within their peer groups peaks at early adolescence (LaFontana & Cillessen, 2010). If future research on the development of the liking gap specifically in these age ranges found a similar peak in early adolescence, this would further consolidate the relationship between the liking gap and individuals' concerns for their reputation and the impression they make on others.

Overall, the current results demonstrate the development of the liking gap through early and middle childhood and illustrate the emergence of people's concern for the impression they make on others during social interactions. It appears that such concerns profoundly shape our social reality, even from an early age, by influencing the way we think others evaluate us and thus, ultimately, shaping the way in which we evaluate ourselves.

Transparency

ORCID iD

Wouter Wolf https://orcid.org/0000-0003-4379-1343

Notes

1. We initially preregistered a sample size of 36 children per age group (12 per dyad gender composition), in line with the sample used by Boothby et al. (2018). Because of limits in the number of participants in the same age group who did not know each other, we decided to allow participants to be one age group apart (and we accounted for this age difference in the analysis). In the later stages of data collection, however, we ended up with dyads in which one participant belonged to a group for which sufficient data had already been collected, whereas the other participant belonged to a group for which data collection was still ongoing. Because we had no reason to believe that collecting more data than specified in the preregistration would meaningfully change the results, we decided to include these children in the data set rather than send them away or exclude them. Consequently, our new sampling goal was to recruit at least 12 participants per cell (instead of exactly 12 participants). In the final stage of collecting data, we were interrupted by the institutional shutdown of all data collection because of the COVID-19 pandemic. We therefore decided to cease data collection indefinitely (before analyzing the data). The preregistration link is https://osf.io/uvyfn.

2. Bayesian 95% HDIs contain the estimated 95% most likely values of the parameter of interest, in this case the standardized regression coefficient. If the 95% HDI does not contain zero, then the effect is deemed statistically credible.

References

Alicke, M. D. (1985). Global self-evaluation as determined by the desirability and controllability of trait adjectives. *Journal of Personality and Social Psychology*, *49*(6), 1621–1630. https://doi.org/10.1037/0022-3514.49.6.1621

Aloise-Young, P. A. (1993). The development of self-presentation: Self-promotion in 6- to 10-year-old children. *Social Cognition*, *11*(2), 201–222. https://doi.org/10.1521/soco.1993.11.2.201

Banerjee, R. (2002). Children's understanding of self-presentational behavior: Links with mental-state reasoning and the attribution of embarrassment. *Merrill-Palmer Quarterly*, *48*(4), 378–404. https://doi.org/10.1353/mpq.2002.0015

Banerjee, R., Bennet, M., & Luke, N. (2012). Children's reasoning about self-presentation following rule violations: The role of self-focused attention. *Child Development*, *83*(5), 1805–1821. https://doi.org/10.1111/j.1467-8624.2012.01813.x

Banerjee, R., & Yuill, N. (1999). Children's explanations for self-presentational behaviour. *European Journal of Social Psychology*, *29*(1), 105–111.

Boothby, E. J., Cooney, G., Sandstrom, G. M., & Clark, M. S. (2018). The liking gap in conversations: Do people like us more than we think? *Psychological Science*, *29*(11), 1742–1756. https://doi.org/10.1177/0956797618783714

Buhrmester, D., Goldfard, J., & Cantrell, C. (1992). Self-presentation when sharing with friends and nonfriends. *Journal of Early Adolescence*, *12*(1), 61–79. https://doi.org/10.1177/0272431692012001004

Deri, S., Davidai, S., & Gilovich, T. (2017). Home alone: Why people believe others' social lives are richer than their own. *Journal of Personality and Social Psychology*, *113*(6), 858–877. https://doi.org/10.1037/pspa0000105

Engelmann, J. M., Herrmann, E., & Tomasello, M. (2012). Five-year olds, but not chimpanzees, attempt to manage their reputations. *PLOS ONE*, *7*(10), Article e48433. https://doi.org/10.1371/journal.pone.0048433

Engelmann, J. M., & Rapp, D. J. (2018). The influence of reputational concerns on children's prosociality. *Current Opinion in Psychology*, *20*, 92–95. https://doi.org/10.1016/j.copsyc.2017.08.024

Epley, N., & Schroeder, J. (2014). Mistakenly seeking solitude. *Journal of Experimental Psychology: General*, *143*(5), 1980–1999. https://doi.org/10.1037/a0037323

Fu, G., Heyman, G. D., Cameron, C. A., & Lee, K. (2016). Learning to be unsung heroes: Development of reputation management in two cultures. *Child Development*, *87*(3), 689–699. https://doi.org/10.1111/cdev.12494

Fu, G., Heyman, G. D., Qian, M., Guo, T., & Lee, K. (2016). Young children with a positive reputation to maintain are less likely to cheat. *Developmental Science*, *19*(2), 275–283. https://doi.org/10.1111/desc.12304

Fujii, T., Takagishi, H., Koizumi, M., & Okada, H. (2015). The effect of direct and indirect monitoring on generosity among preschoolers. *Scientific Reports*, *5*(1), Article 9025. https://doi.org/10.1038/srep09025

Grueneisen, S., Wyman, E., & Tomasello, M. (2015). "I know you don't know I know. . ." Children use second-order false-belief reasoning for peer coordination. *Child Development*, *86*(1), 287–293. https://doi.org/10.1111/cdev.12264

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Herrmann, E., Engelmann, J. M., & Tomasello, M. (2019). Children engage in competitive altruism. *Journal of Experimental Child Psychology*, *179*, 176–189. https://doi.org/10.1016/j.jecp.2018.11.008

JASP Team. (2020). *JASP* (Version 0.12.2) [Computer software]. https://jasp-stats.org/previous-versions/

Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. Guilford Press.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77*(6), 1121–1134. https://doi.org/10.1037//0022-3514.77.6.1121

LaFontana, K. M., & Cillessen, A. H. N. (2010). Developmental changes in the priority of perceived status in childhood and adolescence. *Social Development*, *19*(1), 130–147. https://doi.org/10.1111/j.1467-9507.2008.00522.x

Leimgruber, K. L., Shaw, A., Santos, L. R., & Olson, K. R. (2012). Young children are more generous when others are aware of their actions. *PLOS ONE*, *7*(10), Article e48292. https://doi.org/10.1371/journal.pone.0048292

Nielsen, M., Haun, D., Kärtner, J., & Legare, C. H. (2017). The persistent sampling bias in developmental psychology: A call to action. *Journal of Experimental Child Psychology*, *162*, 31–38. https://doi.org/10.1016/j.jecp.2017.04.017

Perner, J., & Wimmer, H. (1985). "John *thinks* that Mary *thinks* that. . ." attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology*, *39*(3), 437–471. https://doi.org/10.1016/0022-0965(85)90051-7

Piazza, J., Bering, J. M., & Ingram, G. (2011). "Princess Alice is watching you": Children's belief in an invisible person inhibits cheating. *Journal of Experimental Child Psychology*, *109*(3), 311–320. https://doi.org/10.1016/j.jecp.2011.02.003

Rapp, D. J., Engelmann, J. M., Herrmann, E., & Tomasello, M. (2019). Young children's reputational strategies in a peer group context. *Developmental Psychology*, *55*(2), 329–336. https://doi.org/10.1037/dev0000639

Shaw, A., Montinari, N., Piovesan, M., Olson, K. R., Gino, F., & Norton, M. I. (2014). Children develop a veil of fairness. *Journal of Experimental Psychology: General*, *143*(1), 363–375. https://doi.org/10.1037/a0031247

Silver, I. M., & Shaw, A. (2018). Pint-sized public relations: The development of reputation management. *Trends in Cognitive Sciences*, *22*(4), 277–279. https://doi.org/10.1016/j.tics.2018.01.006

Watling, D., & Banerjee, R. (2007). Children's understanding of modesty in front of peer and adult audiences. *Infant and Child Development*, *16*(3), 227–236. https://doi.org/10.1002/icd.450

Whillans, A. V., Christie, C. D., Cheung, S., Jordan, A. H., & Chen, F. S. (2017). From misperception to social connection: Correlates and consequences of overestimating others' social connectedness. *Personality and Social Psychology Bulletin*, *43*(12), 1696–1711. https://doi.org/10.1177/0146167217727496