



City Research Online

City St George's, University of London

Citation: Sijtsma, H., Lemmers-Jansen, I. L. J., Van Buuren, M. & Fett, A-K. (2026). Age- and Sex-Related Differences in Trust During Adolescence and Adulthood. *International Journal of Developmental Science*, doi: 10.1177/2192001x261499917

This is the published version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/38568/>

Link to published version: <https://doi.org/10.1177/2192001x261499917>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

Age- and Sex-Related Differences in Trust During Adolescence and Adulthood

International Journal of
Developmental Science

1–15

© The Author(s) 2026



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/2192001X261499917

journals.sagepub.com/home/dev**H Sijsma¹** , **ILJ Lemmers-Jansen¹** , **M Van Buuren¹**  and **A-K Fett^{2,3}** 

Abstract

The propensity to trust is thought to change across the lifespan and has been hypothesized to differ between sexes. However, the existing literature is inconclusive regarding age- and sex-related associations with trust toward others. We sought to specifically investigate age and sex differences in initial trust toward unknown others and in trust in repeated interactions with human and computer partners who either behaved trustworthy or untrustworthy. In the current study, cross-sectional data of six projects were pooled and included 304 participants aged 13–58 years (51% female). Initial trust toward a human partner showed a quadratic association with age. No age-related effect was observed for computer partners. During repeated interactions, older age was associated with more pronounced reductions in trust toward untrustworthy partners. No age effect was found for trust toward a trustworthy partner. Sex was unrelated to levels of trust in all analyses. Limitations of the study are the cross-sectional nature of the data, highlighting the need for longitudinal research to better understand developmental changes in trust.

Keywords

initial trust, adaptation of trust, age-related associations, sex-related differences, life-span

Received: February 18, 2026; accepted: October 5, 2026

At a macro level, trust can be seen as a vital ingredient for the cohesion and functioning of societies (Newton, 2001; Uslaner, 2002), as it plays a key role in economic transactions, politics, and democracy (Leković, 2012; Misztal, 1996; Stolle, 2002). On a micro level (i.e., for an individual), trusting others can be risky, as it requires the willingness to be vulnerable while there is a risk of getting betrayed. However, being able to show this kind of vulnerability is key to initiating and sustaining successful and mutually beneficial social interactions.

Trust at the micro level can be portrayed as a social dilemma where one person invests resources in another person, not knowing if their trust will be reciprocated (also see Lewicki et al., 2006). These interpersonal trust decisions play a role in different types of relationships, from encounters with strangers to relationships in the workplace, with family, friends, and in romantic relationships (Simpson & Vieth, 2021). Trust in strangers is important during encounters on the street or the train, but also when functioning in society, trust in strangers is essential for online transactions, rental and shared property services, or public transport and traffic, where one trusts others to behave according to rules and the common good.

The integrative neuropsychoeconomic (NPE) model of trust explains how insights and knowledge from diverse disciplines (economy, psychology, and neuroscience) interact in shaping trust behavior (Krueger & Meyer-Lindenberg, 2019). According to this model, trust arises from the interplay of three psychological systems which are motivation, affect, and cognition. The social dilemma of trust decisions balances the anticipation of reward (motivational system) and the risk of getting betrayed (affective system), creating

¹Department of Clinical, Neuro- and Developmental Psychology, Faculty of Behavioral and Movement Sciences, Vrije Universiteit Amsterdam, Research Institute LEARN!, Institute Brain and Behavior Amsterdam, Amsterdam, The Netherlands

²Department of Psychology and Neuroscience, School of Health and Medical Sciences, City St George's University of London, London, UK

³Department of Psychological medicine, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

Corresponding author:

Hester Sijsma, Vrije Universiteit Amsterdam, Research Institute LEARN!, Institute Brain and Behavior Amsterdam, Van der Boechorststraat 7 1081 BT Netherlands.

Email: h.sijsma@vu.nl

uncertainty whether to trust or not (Krueger & Meyer-Lindenberg, 2019). When making these decisions, one balances self-regarding versus other-regarding interests (cognition) (Krueger & Meyer-Lindenberg, 2019). In turn, these psychological systems are related to large brain networks involved in (the anticipation of) reward and reinforcement learning, the regulation of social behavior, cognitive control processes, and social cognition (Krueger & Meyer-Lindenberg, 2019). As suggested by Murtin et al. (2018), individual characteristics, such as age and sex, are key determinants of trust that are not well understood. Erikson's well-known Theory of Psychosocial Development (Erikson, 1963) explains that the first years of life are crucial and formative for learning to trust others; however, also after childhood when one grows more independent, the development and change of trust remains important.

Indeed, some research has suggested that trust changes over the lifespan (Fett, Gromann, et al., 2014; Sutter & Kocher, 2007) and these changes may be associated with age-related changes in the psychological and brain systems that are central in the NPE model. With increasing age, the psychological systems such as reinforcement and feedback learning (Nussenbaum & Hartley, 2019) and social cognition (Hall et al., 2021; Valle et al., 2015; Van der Graaff et al., 2014) develop and change into more advanced forms (for example into abstract forms of social-cognitive thinking such as second-level recursive thinking: 'I think that you think that he/she thinks'). Processes related to the motivational system such as expectations of whether others can be trusted (Ma et al., 2020) or preferences regarding how stakes should be divided (Westhoff et al., 2020) change with age as well. Furthermore, research shows structural and functional maturational changes in brain networks important for trust (e.g., involved in reward and learning, cognitive control processes, and social cognition) (Andrews et al., 2021; Crone & Fuligni, 2020; Krabbendam et al., 2024; Larsen & Luna, 2018; Sydnor et al., 2021). Moreover, not only these social-cognitive, learning, and brain processes, but also one's expectations and preferences change as people grow older, which may underlie age-related changes in trust decisions.

However, the results of experimental studies examining age-related changes in trust are still relatively scarce and inconclusive. Some studies that investigated the initial trust one has when interacting with someone unknown showed that this type of trust increases across adolescence (Sijtsma, Lee, et al., 2023: [ages 11–16]; Van den Bos et al., 2012: [ages 11–21]) and adulthood (Sutter & Kocher, 2007: [ages 8–68]), but other findings did not point in this direction (Fett, Shergill, et al., 2014: [ages 13–18]; Holm & Nystedt, 2005: [20 versus 70 years]; Lemmers-Jansen, Fett, Shergill, et al., 2019: [ages 13–19]; Van de Groep et al., 2020: [ages 12–18]). When social interactions continue beyond initial trust, individuals

can change and adapt their trust in response to the demonstrated trustworthiness of the interaction partner in a tit-for-tat manner (Duncan et al., 2025). Reward and learning processes come into play, as there is a chance to get betrayed by an untrustworthy person. It may also happen that the other player is more trustworthy than expected. Both situations may result in the need to adapt one's own trust as the repeated social interaction unfolds. Some studies that used the trust game showed age-related changes in the adaptation of trust toward the partner's trustworthiness (Lemmers-Jansen, Fett, Shergill, et al., 2019: [ages 13–19]; Lee et al., 2016: [young versus mid-, and late adolescents]; Sijtsma, Lee, et al., 2023: [ages 11–16]; Van den Bos et al., 2012: [ages 10–23]; Westhoff et al., 2020: [ages 8–23]), whereas others did not find evidence for this relationship (Fett, Shergill, et al., 2014: [ages 13–18]; Lemmers-Jansen et al., 2017: [ages 16–27]; Telga & Lupiáñez, 2021: [21 versus 65 years]).

It is difficult to speculate how these differences in findings arise, but some of these discrepancies may be explained by restricted age ranges, small sample sizes, or differences in the sample composition (e.g., in terms of sex). Moreover, all of the above studies used similar experimental paradigms based on the trust game (Berg et al., 1995) to study trust behavior. This game is a well-known behavioral economic task simulating a social interaction between two players (a trustor and a partner). However, many variants of the game exist. Thus, another reason for the different findings with regards to age-related changes in initial trust may be differences in paradigms. In sum, there is some evidence for age-related changes in initial and adaptive trust during the unfolding of a social interaction; however, not all studies point in this direction.

The age-related development of trust may not occur in a similar fashion for everyone. To better understand what supports or constrains trust, it is important to understand how trust is shaped by different factors. Besides age, another important demographic characteristic that has been suggested to impact levels of trust is sex (Murtin et al., 2018). A meta-analysis investigating sex differences showed that adult males have higher levels of initial trust than females, although the effect was small (Van den Akker et al., 2020). Similar results were found in adolescents and young adults (Lemmers-Jansen et al., 2017: [ages 16–27]), however this difference was not replicated in a younger sample (Lemmers-Jansen, Fett, Shergill, et al., 2019: [ages 13–19]). Evolution-based explanations of sex differences in trust hold that males may be typically more risk-taking than females to acquire goods and achieve a high status, leading to higher levels of initial trust (Van den Akker et al., 2020). Lower levels of initial trust in females could also be related to females expecting lower levels of reciprocal behavior from others compared to males (Buchan et al., 2008; Dorough & Glöckner, 2019) or to greater risk aversion (Borghans et al., 2009; Dohmen et al., 2011).

Sex differences in relation to adaptive trust have hardly been examined. Results from previous studies did not show sex differences during interactions with trustworthy partners (Lemmers-Jansen et al., 2017; Lemmers-Jansen, Fett, Shergill, et al., 2019). However, during interactions with untrustworthy partners, sex differences in relation to age were found, as one relatively small study in adolescents showed that younger males decreased their trust more strongly than older males (Lemmers-Jansen, Fett, Shergill, et al., 2019: [ages 13–19]), whereas in a slightly older sample, stronger decreases in trust were found in older males (Lemmers-Jansen et al., 2017: [ages 16–27]). This pattern suggests a quadratic change for males, but not females. This sex difference might also suggest that males are more punishing toward untrustworthy others or that females prioritize maintaining friendly relationships with others, e.g., according to social expectations placed on them (Eagly & Wood, 2011; Haselhuhn et al., 2015; Lemmers-Jansen et al., 2017).

In sum, existing findings on age- and sex-related differences in trust showed inconsistent patterns. These inconsistent patterns may be related to characteristics of the studies (e.g., small sample sizes) or differences in trust game designs (e.g., instructions that participants receive, the specific timing and order of events within the task, or the algorithm to model the trustworthiness level of the partner). Differences in the trust game design make it hard to pool data of smaller studies for re-analysis to examine effects of demographic characteristics such as age- and sex-related patterns in a bigger sample. Our research group has conducted six projects with highly similar trust game procedures and designs, which in the current paper were pooled to investigate age- and sex related effects in trust.

Notably, one important difference between studies is the type of real or hypothetical partner that is used. For example, some studies used a presumed human partner where participants were informed that they would play with a human partner (but they actually played with a computer partner), while in other studies, participants played with a computer partner and were informed about this as such. Perhaps participants respond differently knowing that the interaction partner is a computer versus believing that the partner is human. In light of the increasing development of artificial intelligence (AI) in everyday social interactions in which computers are increasingly viewed as social actors instead of technical tools (Wang & Goel, 2022), it is important to understand how trust in human-AI interactions may differ from human-human interactions (Li et al., 2024; Reinecke et al., 2025; Yang & Rau, 2024). In recent years, people started seeing the usefulness and ease of use of technology in all sorts of fields (see the Technology Acceptance Model by Davis (1989); also see Marangunić and Granić (2015)) and, due to ongoing innovations, computers are getting equipped with human-like Theory of Mind capabilities (Wang & Goel, 2022). People may recognize and appreciate

Table 1. Age and Sex Distribution Per Project.

Project	Age (M (SD))	Sex Distribution (n Male/n Female)
TAP	16 (1.6)	50/50
Trust	32.3 (8.7)	21/19
DECOP	34.2 (9.5)	14/35
COGPID	41.3 (12.6)	20/15
Beslist	20.6 (2.7)	25/22
GROUP	33.4 (10.2)	19/14

Note: TAP = Trust in Adolescence Project; DECOP = Decision Making and Context Processing in Psychosis; COGPID = Cognitive Processes in Decision making; GROUP = Genetic Risk and Outcome in Psychosis project.

these “social skills”, impacting the way they respond to computer partners. Therefore, we additionally examined whether trust differed during trust game interactions with presumed human versus computerized counterparts.

Our key research questions were (1) whether patterns of trust were associated with age and (2) whether such associations were sex-dependent, and/or (3) dependent on the nature of the counterpart (trustworthy versus untrustworthy, human versus computer). First, we hypothesized an age-related increase of initial trust from adolescence to adulthood, and the level of initial trust to be higher in males compared to females. Second, we hypothesized an increase in adaptive trust from adolescence to adulthood, expressed through a stronger age-related increase in investments throughout repeated trust game rounds when interacting with a trustworthy partner and a stronger age-related decrease in investments when interacting with an untrustworthy partner. We did not have hypotheses about sex differences regarding the adaptation of trust over repeated game rounds and trust differences regarding the nature of the counterpart, as findings of previous studies are inconclusive. Likewise, analyses regarding trust toward human versus computer partners were exploratory.

Method

Participants

The current analyses included data of six different projects conducted by the investigator team and colleagues at three different research sites in the Netherlands (NL) and the United Kingdom (UK). In total, 304 participants from the general population completed the trust game (mean age = 26.6, standard deviation = 11.9, age range = 13–58, 51% female). Detailed information on the participants per project can be found in Table 1. The UK projects were conducted at King’s College London, Institute of Psychiatry, Psychology and Neuroscience (projects: Trust in Adolescence Project (TAP), Trust project (Trust), and Decision Making and Context Processing in Psychosis (DECOP)). The Dutch projects were conducted at the

University of Maastricht and the Vrije Universiteit Amsterdam (projects: Cognitive Processes in Decision making (COGPID), Beslist, and an add-on project from the Genetic Risk and Outcome in Psychosis project (GROUP)). The TAP, Trust, DECOP, COGPID, and the GROUP add-on project were funded by the Dutch Research Council ('De Nederlandse Organisatie voor Wetenschappelijk Onderzoek'). Beslist was funded by the Hersenstichting.

The projects mostly used convenience sampling. TAP participants were recruited via collaborations with two local schools and were drawn from multiple classes. Participants in the Trust and DECOP projects were recruited via colleagues, local adverts, and through a community volunteer database (Mindsearch), online via websites (e.g., Gumtree, Callforparticipants), and recruitment circular emails at Institute of Psychiatry, Psychology and Neuroscience at King's College London. COGPID participants were recruited through adverts in the local area. Participants of the Beslist project were recruited at educational institutions, through snowball sampling, and personal networks. Participants in the GROUP add-on project were recruited from the original GROUP participant pool (Korver et al., 2012). Inclusion criteria for all participants were having a good command of their local native language, not having a history of neurological disorders, current psychiatric diagnosis, or taking psychotropic medication. The studies did differ in terms of their inclusion criteria for age: TAP age 13–19, Trust and DECOP age 18–65, COGPID age 18–60, Beslist age 16–31, Group age 18–60.

The projects were approved by the local research ethics committees (TAP by London-Surrey Borders; Trust by Barking and Havering; DECOP by London-Harrow; COGPID by the Medical Research Ethics Committee of the University Medical Centre Maastricht; Beslist by the Medical Research Ethics Committee of the Vrije Universiteit Medical Centre Amsterdam; GROUP by the Ethical Review Board of the University Medical Centre Utrecht). Informed consent was obtained from all participants and their parents/guardians if they were under the age of 16 years. The study procedures of each of the six included studies are detailed in the Supplemental Material.

This study's research aim, hypotheses, and planned analyses were preregistered on Open Science Framework (<https://osf.io/b258d/metadata/osf>). As reported in the pre-registration, some smaller subsets of the data that were included in the current study have previously been analyzed and published with the purpose to analyze age or sex effects on trust (Fett, Gromann, et al., 2014; Lemmers-Jansen et al., 2017; Lemmers-Jansen, Fett, Shergill, et al., 2019). However, the analysis in the current study is unique in that it examines age and sex differences in different aspects of trust toward humans and computers from early adolescence to late adulthood in the largest sample so far, thus contributing more reliable and extended insights.

Measures

Multi-Round Trust Game with a Trustworthy and Untrustworthy Game Partner. The trust game (Berg et al., 1995) is a well-established experimental paradigm which is regularly used in adolescent (Fett, Shergill, et al., 2014; Lemmers-Jansen, Fett, Shergill, et al., 2019; Van den Bos et al., 2012), as well as adult samples (Holm & Nystedt, 2005; Sutter & Kocher, 2007). The game simulates the back-and-forth flow of a social interaction, which makes it possible to measure levels of interpersonal trust while participants are interacting with others, making the game a suitable operationalization for our research aim.

All participants took part in one or more conditions of the same trust games (Berg et al., 1995) that were played with preprogrammed counterparts (except COGPID and DECOP, explained below). In all our studies, participants were in the investor role. The trust games consisted of 20 experimental rounds. In each experimental trust game round, participants were asked to transfer an (integer) amount between £/€0 and £/€10 to the trustee (i.e., game partner). The transferred amount was tripled. The subsequent trustee repayment depended on the preprogrammed computer algorithm (see Supplemental Material for an explanation of the computer algorithm). A round ended with a display of the partner's repayment and total earning of that round to the participant. At the beginning of a new round, participants received £10 again, meaning there were no cumulative totals.

Each round started with an investment cue of £10 (2 s). The following investment period required the participant to move a cursor with their index fingers to select a number from 0 to 10, always starting at 5 (max. 4 s). The invested amount was shown (2 s), followed by a waiting period with a bar slowly filling itself with dots (2–4 s) and a fixation cross (500 milliseconds). The trustee's response was then displayed (3 s), followed by the totals (3–5 s, depending on the trustee's response). The round ended with a fixation cross (500 milliseconds). In total, each round lasted 18.5 s. See Figure 1 for a visual display of our trust game design.

Participants were never informed about the nature of the partner's trustworthiness (i.e., whether the partner was trustworthy or not) and had to learn about the trustworthiness of the trustee through their repayments. However, participants were informed that they would play a game with a human or computer counterpart (see Supplemental Material for an explanation of the task instructions regarding human and computer partners). There were three possible trust game conditions. In the first, participants believed they played a human partner but actually played a computer algorithm which was programed to be highly trustworthy, henceforth called **trustworthy human partner**. In the second condition, participants again believed they played a human partner but actually played an algorithm which was programed to be

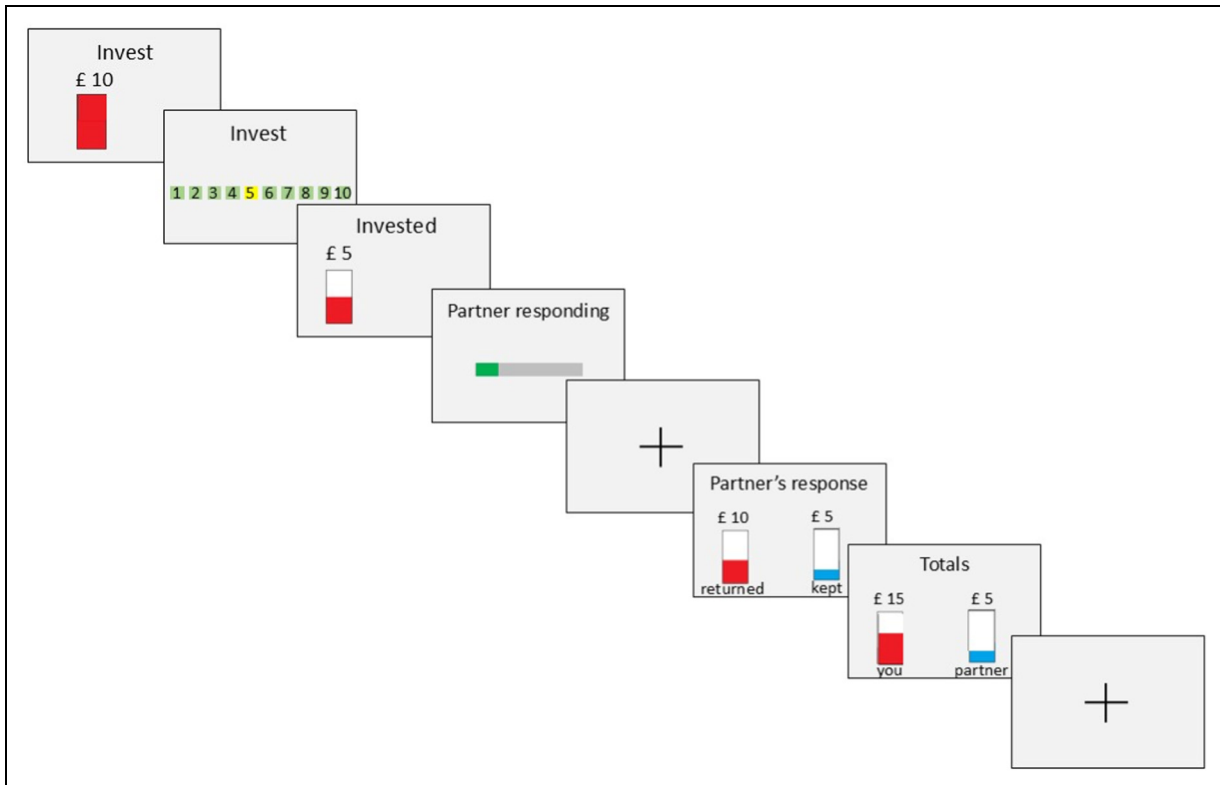


Figure 1. The trust game. Note: This is an example of a round in the trust game. The participant receives £10 and invests £5. The partner receives the tripled amount of which £5 is kept and £10 is returned to the participant.

highly untrustworthy, henceforth called *untrustworthy human partner*. In the third, the computer condition, participants were truthfully informed that they would play a game with a computer. Unknown to the participants the underlying algorithm was the same as the one used for the trustworthy human partner, henceforth called *trustworthy computer partner*.

In sum, trustee behavior in all three conditions was based on a preprogrammed algorithm, which in case of the trustworthy human partner and computer partner made trustworthy repayments, and which in case of the untrustworthy human partner, behaved untrustworthily. The specific returns in each trial occurred in a probabilistic way. See the Supplemental Material for more details on the respective algorithms.

The exceptions to the above description were the COGPID and DECOP studies. DECOP used comparable trust games, but the return algorithms differed from the other projects. In COGPID participants were playing the trust game with real human partners. As such, data from COGPID and DECOP could only be used in the analysis on initial trust toward human partner, where the trustee returns could not yet have impacted on investor trust.

Initial trust. Initial trust was operationalized as the investment made in the first round of a trust game. We investigated initial trust in two conditions: (1) the first

investment in the multi-round trust game with the first (presumed or real) human counterpart and (2) the first investment in the multi-round trust game with a computer partner, where available. Table 2 shows the conditions that were included in the respective projects.

The adaptation of trust. In all conditions (see Table 2), the adaptation of trust was operationalized as the change in investments over 20 trust game rounds of a multi-round trust game. This concerns the repeated measures of investments within person and was done so for all three conditions.

General Cognitive Ability. We report general cognitive ability, where available, to further describe our sample and to identify possible factors that could impact trust behavior. In the TAP and DECOP projects, the English version of the vocabulary subtest of the Wechsler Abbreviated Scale of Intelligence (WASI) was used as indicator of general cognitive ability in adolescents (Wechsler, 1999). In the Trust project, the English version of the Wechsler Adult Intelligence Scale (WAIS-III) was used (Wechsler, 1981). The COGPID and Beslist projects used the Dutch version of the WAIS-III vocabulary test. *T*-scores of the WASI were converted to scaled scores (mean = 10, SD = 3) for comparability (Deutsch Lezak et al., 2012). The GROUP

Table 2. The Conditions and Analyses That Were Included for the Six Projects.

	Initial Trust Human Partner	Initial Trust Computer Partner	Adaptation of Trust Trustworthy Human Partner	Adaptation of Trust Trustworthy Computer Partner	Adaptation of Trust Untrustworthy Human Partner	Location of Data Collection
TAP	x	x	x	x	x	UK*
Trust	x	x	x	x	x	UK*
DECOP	x	x				UK*
COGPID	x					NL
Beslist	x		x		x	NL*
GROUP	x		x			NL*

Note: 'x' means that the condition was included in the study. In the COGPID project, participants played with a real human counterpart. *The asterisk indicates that for some or all the participants, the trust game was administered in an MRI scanner. TAP = Trust in Adolescence Project; DECOP = Decision Making and Context Processing in Psychosis; COGPID = Cognitive Processes in Decision making; GROUP = Genetic Risk and Outcome in Psychosis project.

project did not collect data on cognitive functioning with a vocabulary measure.

Age. Age is a continuous predictor variable. Additional sensitivity analyses were conducted using a model including the main effect of age by using age groups as predictor (instead of a continuous variable) as previously done by Sutter and Kocher (2007). The age categories were based on The Office for National Statistics (2023) (15 years and under, 16–24 years, 25–34 years, 35–49 years, 50–64 years, 65 and over).

Sex. Sex is a categorical predictor variable (male, female).

Collection Mode. The trust game was administered in the MRI scanner for all participants in the Beslist and GROUP projects, and to a subset of participants in the TAP, Trust, and DECOP projects. The participants in the COGPID project completed the trust game during a behavioral session. Therefore, collection mode (MRI versus behavioral study) was added as a categorical control variable to all analyses.

Procedure

The current study pooled data from six different projects that used the trust game. In the following section, we provide a short explanation of the testing procedures (more detailed explanations are included in the Supplemental Material). Participants and, where applicable, their legal guardians provided informed consent.

All participants completed demographic questionnaires at the beginning of the testing sessions which also comprised additional measures (not analyzed in the current study), and then either played the trust games in an MRI scanner or on a computer or laptop in a testing room. TAP, Trust and DECOP testing sessions took place at the Institute of Psychiatry, Psychology and Neuroscience at King's College London, with MRI scanning at the Centre of

Neuroimaging Sciences. The COGPID project only included behavioral testing and was conducted in the Beelab of the Faculty of Economics at Maastricht University. The testing sessions and scanning for the Beslist project was an fMRI study and took place at the Spinoza Centre of the University of Amsterdam. For the GROUP add-on project, testing and scanning took place the Academic Medical Centre in Amsterdam. Upon completion of the testing sessions, participants were remunerated for their time. For details on the individual testing procedures per project see the Supplemental Material or previous publications - Fett, Shergill, et al. (2014) for TAP, Gromann et al. (2013) for Trust, Hanssen et al. (2022) for DECOP, Fett et al. (2012) for COGPID, Lemmers-Jansen, Fett, Hanssen, et al. (2019) for Beslist, and Gromann et al. (2014) for GROUP.

Statistical Analyses

Five analyses were run according to partner type (i.e., human counterpart versus computer partner), the trustworthiness level of the partner (trustworthy versus untrustworthy partner), and for initial trust and the adaptation of trust over repeated game rounds, separately. For each analysis, linear and quadratic age-related effects were tested. Model selection was based on AIC, BIC, and nested model comparisons. F-tests were used to compare nested linear regression models, whereas likelihood-ratio tests were used for multi-level models. All analyses were performed controlling for collection mode (MRI study versus a behavioral study). We conducted sensitivity analyses to explore the effect of estimated general cognitive ability on trust (except for the GROUP project). Furthermore, sensitivity analyses were conducted using age groups as predictor. If outliers in trust were present, sensitivity analyses were run without these participants (and results are reported only in cases when the effects of age and sex on trust changed).

1. Age and sex effects on initial trust toward a human partner: A regression model was used to

test age and sex effects on initial trust when participants played with a (presumed or real) human partner. The main model included the first trust game investment (outcome variable) and the two-way interaction between age and sex, and the main effect of age and sex (predictors).

2. **Age and sex effects on initial trust toward a computer partner:** A regression model was used to test age and sex effects on initial trust when participants played with a computer partner. The main model consisted of the investment in the first round (outcome variable) and the two-way interaction between age and sex, and the main effect of age and sex (predictors).
3. **Age and sex effects on the adaptation of trust in repeated interactions with a trustworthy human partner:** A multi-level regression model was run to test age and sex effects on trust in repeated interactions with a trustworthy human partner. The adaptation of trust refers to the change in investments throughout the game (i.e., repeated measures within person), requiring a multi-level modeling approach. The main model included the investments in repeated rounds (outcome variable) and the three-way interaction between age, sex, and round, and the related two-way interactions and main effects (predictors). A random intercept and a random slope for round were included.
4. **Age and sex effects on the adaptation of trust in repeated interactions with a trustworthy computer partner:** A multi-level regression model was used to test age and sex effects on trust in repeated interactions with a trustworthy computer partner. The main model included the investments in repeated rounds (outcome variable) and the three-way interaction between age, sex, and round, and the related two-way interactions and main effects (predictors). A random intercept and a random slope for round were included.
5. **Age and sex effects on the adaptation of trust in repeated interactions with a untrustworthy human partner:** A multi-level regression model was used to test age and sex effects on trust in repeated interactions with an untrustworthy human partner. The main model included the investments (outcome variable) and the three-way interaction between age, sex, and round, including the related two-way interactions and main effects (predictors). A random intercept and a random slope for round were included.
6. **Additional analysis:** Using multi-level analyses, we examined if initial trust (by examining the effect of condition) and changes in trust over repeated trust game interactions (by examining the interaction between condition and round) differ for interactions

Table 3. The Results of the Model Fit for Analysis 1 and 2 (Initial Trust).

Analysis	1		2A		2B	
	b	p	b	p	b	p
Intercept	6.7	<.01	5.26	<.01	5.29	<.01
Age linear × sex	6.2	.22	2.2	.68		
Age quadratic × sex	8.81	.09	2.26	.67		
Age linear	0.71	.84	0.57	.9	2.12	.39
Age quadratic	−8.47	.02	−3.41	.42	−1.84	.46
Sex	−0.37	.2	0.62	.1	0.6	.11
Collection mode	−0.13	.66	0.66	.08	0.64	.08

Note: Analysis 1: human partner – inclusion of the two-way interactions.

Analysis 2A: computer partner – inclusion of the two-way interactions.

Analysis 2B: computer partner – removal of the two-way interactions.

with a (presumed) human partner versus computer partner.

Results

Analysis 1: Initial Trust Toward a Human Partner

In the main model, the interactions between age (linear and quadratic terms) and sex were not significant (see analysis 1 in Table 3; $N = 304$, mean age = 26.6, standard deviation = 11.9, range = 13–58, 52% female). The quadratic main effect of age was significant and negative (see Table 3 and Figure 2). The linear effects of age and sex were non-significant. Two participants scored as outliers; however, analyses without these outliers did not change the direction or strength of the associations. Sensitivity analyses with age as a categorical variable also showed non-significant associations (see Table S4 in the Supplemental Material). Cognitive ability was significantly and positively associated with higher initial trust (see Table S4 in the Supplemental Material). Sensitivity analyses including cognitive ability did not change the effects of age and sex on trust.

Analysis 2: Initial Trust Toward a Computer Partner

In the main model, the interaction between age (linear and quadratic terms) and sex was not significant (see analysis 2A in Table 3; $N = 189$, mean age = 24.2, standard deviation = 10.8, range = 13–53, 55% female). The results of the nested model comparison indicated that the model including the interaction between age and sex ($AIC = 884$, $BIC = 910$) did not significantly improve model fit compared with the model without the interaction ($AIC = 880$, $BIC = 900$), $F(2, 182) = 0.17$, $p = .84$. The results of the model including only the main effects showed that associations with age (linear and quadratic terms) and sex were non-significant (see analysis 2B in Table 3). Sensitivity

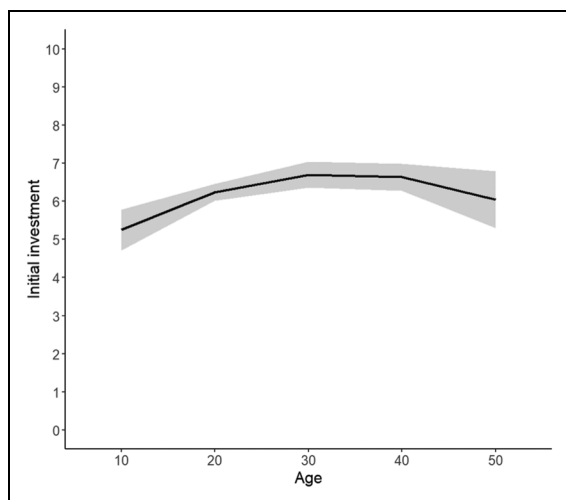


Figure 2. The quadratic age effect for initial trust toward a human partner.

analyses with the inclusion of cognitive ability showed no change in the effect of age and sex on trust and the main effect of cognitive ability was not significant (see Table S5 in the Supplemental Material). In the sensitivity analyses with age as a categorical variable, all associations remained non-significant (see Table S5 in the Supplemental Material).

Analysis 3: The Adaptation of Trust in Repeated Interactions with a Trustworthy Human Partner

In the main model, the three-way interaction between age (linear and quadratic terms), sex, and trust game round was not significant (see analysis 3A in Table 4; $N=220$, mean age = 22.5, standard deviation = 9.3, range = 13–55, 49% female). The nested model comparison indicated that including the three-way interaction did not significantly improve model fit (AIC = 19,707, BIC = 19,815) compared with the model without the three-way interaction (AIC = 19,703, BIC = 19,799), $\chi^2(2)=0.34$, $p=.85$. The results of the model including the two-way interactions between age (linear and quadratic terms) and sex, age (linear and quadratic terms) and game round, and sex and game round, and the related main effects showed that only the interaction between the quadratic term of age and sex was significant (see analysis 3B in Table 4). The nested model comparison indicated that including the three two-way interactions did not significantly improve model fit (AIC = 19,703, BIC = 19,799) compared with the model including only main effects (AIC = 19,701, BIC = 19,765), $\chi^2(5)=7.90$, $p=.16$. We therefore also fitted a model with only the main effects of age (linear and quadratic terms), sex, and round as predictors (see analysis 3C in Table 4). Sensitivity analyses with the inclusion of cognitive ability showed no

change in the effect of age and sex on trust; however, cognitive ability had a positive, significant effect on trust (see Table S6 in the Supplemental Material). Furthermore, the results of the sensitivity analyses with age as a categorical variable showed non-significant associations (see Table S6 in the Supplemental Material).

Analysis 4: The Adaptation of Trust in Repeated Interactions with a Trustworthy Computer Partner

In the main model, the three-way interaction between age (linear and quadratic terms), sex, and game round was not significant (see analysis 4A in Table 4; $N=140$, mean age = 20.6, standard deviation = 8.8, range = 13–49, 49% female). The nested model comparison indicated that including the three-way interaction did not significantly improve model fit (AIC = 12,552, BIC = 12,653) compared with the model without the three-way interaction (AIC = 12,549, BIC = 12,638), $\chi^2(2)=1.61$, $p=.45$. The results of the model including the two-way interactions showed that the interactions between age (linear and quadratic terms) and sex, age (linear and quadratic terms) and game round, and sex and game round were also non-significant (see analysis 4B in Table 4). The nested model comparison indicated that including the three two-way interactions did not significantly improve model fit (AIC = 12,549, BIC = 12,638) compared with the model including only main effects (AIC = 12,545, BIC = 12,605), $\chi^2(5)=6.1$, $p=.3$. Subsequently, we also fitted a model with only the main effects of age (linear and quadratic terms), sex, and game round as predictors (see analysis 4C in Table 4). Sensitivity analyses with the inclusion of cognitive ability did not change the effects of age and sex on trust, but results showed that higher cognitive ability was significantly and positively associated with higher trust (see Table S7 in the Supplemental Material). The results of the sensitivity analyses with age as a categorical variable showed a significant difference between age group 1 (15 years and under) and group 4 (35–49 years; see Table S7 in the Supplemental Material).

Analysis 5: The Adaptation of Trust in Repeated Interactions with a Untrustworthy Human Partner

The main model showed that the three-way interaction between age (linear and quadratic terms), sex, and game round was not significant (see analysis 5A in Table 4; $N=186$, mean age = 20.7, standard deviation = 7.7, range = 13–49, 48% female). The nested model comparison indicated that including the three-way interaction did not significantly improve model fit (AIC = 18,033, BIC = 18,139) compared with the model without the three-way

Table 4. The Result of the Model Fit for Analyses 3, 4 and 5 (the Adaptation of Trust Over Repeated Interactions).

Analysis	3A			3B			3C			4A			4B			4C			5A			5B		
	b	p		b	p		b	p		b	p		b	p		b	p		b	p		b	p	
Random effects																								
Intercept	2.84			2.83			2.88			3.11			3.11			3.08			2.96			2.99		
Round	0.01			0.01			0.01			0.01			0.01			0.01			0.01			0.01		
Fixed effects																								
Intercept	6.54	<.01		6.54	<.01		6.49	<.01		5.63	<.01		5.63	<.01		5.69	<.01		5.22	<.01		5.23	<.01	
Round × age linear × sex	0.56	.61								1.16	.25								1.75	.16				
Round × age quadratic × sex	-0.27	.81								0.47	.64								-2.3	.07				
Age linear × sex	-20.26	.25		-16.14	.31					-15.34	.43		-5.62	.75					-8.54	.66		8.5	.58	
Age quadratic × sex	34.72	.05		32.75	.04					2.46	.89		6.4	.69					17.16	.36		-5.23	.72	
Round × age linear	0.51	.51		0.79	.15					-0.06	.94		0.5	.32					-2.32	.01		-1.39	.03	
Round × age quadratic	0.17	.83		0.03	.95					0.75	.29		0.94	.06					1.64	.06		0.47	.46	
Round × sex	-0.01	.62		-0.01	.62					-0.02	.24		-0.02	.24					0.02	.36		0.02	.36	
Age linear	24.4	.05		22.34	.06		19.29	.02		25.86	.05		21.19	.09		22.33	.01		27.32	.04		18.31	.13	
Age quadratic	-34.33	<.01		-33.3	.01		-15.72	.05		-18.26	.15		-19.83	.1		-8.62	.28		-15.42	.25		-3.99	.74	
Round	0.04	<.01		0.04	<.01		0.04	<.01		0.07	<.01		0.07	<.01		0.06	<.01		-0.13	<.01		-0.13	<.01	
Sex	-0.01	.97		-0.01	.97		-0.05	.84		0.73	.04		0.73	.04		0.56	.07		-0.43	.17		-0.43	.17	
Collection mode	0.48	.08		0.48	.08		0.6	.02		0.47	.17		0.47	.17		0.53	.09		0.23	.4		0.22	.04	

Note: For random effects, the variance is reported. Model 3A: trustworthy human partner – inclusion of the three-way interaction. Model 3B: trustworthy human partner – removal of the three-way interaction. Model 3C: trustworthy human partner – the main effects of age, sex, and round. Model 4A: trustworthy computer partner – inclusion of the three-way interaction. Model 4B: trustworthy computer partner – removal of the three-way interaction. Model 4C: trustworthy computer partner – the main effects of age, sex, and round. Model 5A: untrustworthy human partner – inclusion of the three-way interaction. Model 5B: removal of the three-way interaction.

interaction ($AIC = 18,035$, $BIC = 18,128$), $\chi^2(2) = 5.48$, $p = .06$. The results of the model including the two-way interactions between age (linear and quadratic terms) and sex, age (linear and quadratic terms) and game round, and sex and game round (and the related main effects, see analysis 5B in Table 4) showed a significant, negative two-way interaction between age (linear term) and game round, indicating that the decrease of trust throughout the game became stronger with age (see Figure 3). Sensitivity analyses with the inclusion of cognitive ability showed no change in the effect of age and sex on trust and the effect of cognitive ability was not significant (see Table S8 in the Supplemental Material). Furthermore, the results of the sensitivity analyses with age as a categorical variable showed non-significant associations (see Table S8 in the Supplemental Material).

Additional Analysis: Initial and Adaptation of Trust When Playing with a Human Partner Versus a Computer Partner

Additionally, we examined if initial trust differed when participants were informed that they were playing with a human partner compared to when they were informed that they played with a computer partner. The results showed that participants' initial investment was significantly higher when they believed that they were playing with a human partner (mean investment: 6.42) compared to when they believed that they were playing with a computer partner (mean investment: 5.93) ($t(242) = 2.13$, $p = .03$). The adaptation of trust (i.e., increase in investments) when believing that the game was played with a human partner was less steep compared to when playing with a computer partner (while participants in fact were playing two trust games with the identical preprogrammed "trustworthy" algorithm) ($t(6970) = 2.66$, $p = .01$).

Discussion

This study investigated age and sex differences in the levels of initial trust, in trust during repeated interactions with a trustworthy or untrustworthy counterpart, and toward human and computer partners. The main results showed a quadratic effect of age on initial trust toward presumed human counterparts; however, no such effect was present when participants were explicitly instructed that they were playing with computers. Over repeated trust game interactions with untrustworthy partners, decreases in trust were stronger with increasing age. In contrast, no age-related effects were observed for the levels of trust over repeated interactions with a trustworthy partner. We did not find any significant sex differences in initial trust or trust during repeated interactions with the different game

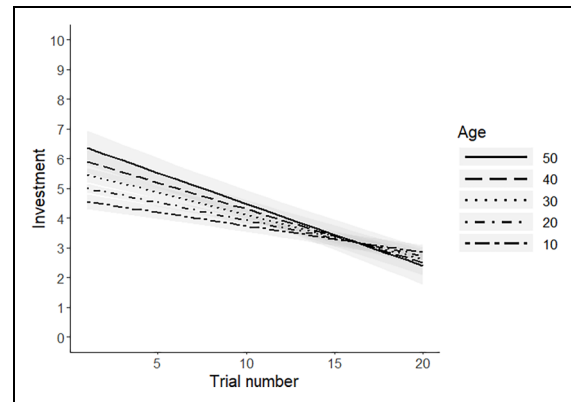


Figure 3. A significant interaction between age and game round in repeated interactions with an untrustworthy human partner.

partners. Overall, the findings suggest that regardless of sex, trust toward unknown others is the highest in mid-life and that individuals respond to others' untrustworthy behavior in an increasingly tit-for-tat manner as they get older. Longitudinal research is needed to corroborate whether these effects are indeed developmental.

Initial Trust Toward Human and Computer Partners

The present findings showed support for quadratic age-related patterns of initial trust in interactions with a human partner, with increasing trust from age 13 to 40, and a decrease thereafter. These results are largely in line with study findings of Sutter and Kocher (2007), who used a one-shot trust game design (i.e., a trust game existing of only one trial) in which participants believed that they were interacting with other individuals. Their results showed that initial trust increased from childhood to adulthood after which it stabilized in older age. The age-related quadratic pattern of trust fits with the practical 'stranger-danger' phenomenon, which refers to warning and making children and older adults cautious of the dangers of interacting (and trusting) unknown others (Moran et al., 1997).

In contrast, our results did not support higher age-related initial trust when participants were informed that they interacted with a computer counterpart. However, we found that participants had higher mean levels of initial trust when they thought their game partner was human as compared to when they knew they interacted with a computer. Overall, our results support that the nature of the counterpart influences participants' trust, highlighting the importance for comparisons and conclusions based on different trust game studies (also see Houser et al., 2010; Jin et al., 2025; Johnson & Mislin, 2011). Our findings may be of relevance to broader discussions surrounding the growing use of AI, as AI is increasingly taking over human roles

during business-related social interactions (e.g., in customer services and checkouts in stores). Even though technology is modernized with features to resemble human theory of mind capabilities, and humans may respond by attributing mental states to computers (Thellman et al., 2022, 2023; Wang & Goel, 2022), our results point to possible differences in patterns of behavior between trust-based interactions with computers and those with human partners. Perhaps trust is too complex, requiring people to be vulnerable, which may raise their cautiousness when interacting with computers. This can especially be the case in situations in which people interact with AI partners where privacy or the sharing of personal data plays a role, as for example during interactions in medical settings (e.g., registration desks in hospitals or pickup counters in pharmacies), where computers increasingly substitute human roles. However, these results should be interpreted with caution, as trust-based interactions with AI agents may depend on individuals' attitudes toward AI, which were not measured in the present study. Such attitudes would be particularly relevant in the computer condition, where they may influence whether behavior in this condition reflects trust-related processes or, at least in part, a more general tendency toward risk-taking. Future research should continue to examine how trust differs during interactions with computers and humans, extend investigations to other paradigms and everyday situations, and consider the role of individuals' attitudes toward AI (Li et al., 2024; Reinecke et al., 2025; Yang & Rau, 2024).

Trust During Repeated Interactions with Human and Computer Partners

The only significant effect of age that we found during repeated trust game interactions was a stronger decrease in trust during the trust game with untrustworthy game partners. The absence of age effects in trust during trustworthy interaction was consistent with earlier work in which only subtle age differences were found (Telga & Lupiáñez, 2021). In line with our results, previous studies also found that age and learning effects during repeated trust game interactions were more pronounced with untrustworthy compared to trustworthy others (Fareri et al., 2012; Sijtsma, Lee, et al., 2023). Furthermore, Fett, Gromann, et al. (2014) and Lemmers-Jansen et al. (2017), of which data was included in the current study, found age effects during untrustworthy interactions (as opposed to trustworthy interactions), indicating that these earlier findings hold in the analysis based on the larger pooled dataset. Perhaps age comes with increasing experience with untrustworthy interactions in daily life and the associated negative consequences (i.e., getting betrayed), and how to manage these (Fareri et al., 2012), explaining the

stronger age-related decreases in trust when interacting with untrustworthy others.

Sex Differences

A previous meta-analysis including adult samples showed that males have higher levels of baseline trust (i.e., similar to initial trust where trusting decisions are made without knowledge about the counterpart) than females (Van den Akker et al., 2020), whereas findings in adolescent samples were mixed (Derks et al., 2014; Lemmers-Jansen et al., 2017; Sijtsma, Lee, et al., 2023). Social role theory, evolutionary theory, and theories on sex differences in risk perception and expectations suggest women are less trusting than men (Borghans et al., 2009; Coltrane, 2007; Dohmen et al., 2011; Eagly & Wood, 2011; Van den Akker et al., 2020; Wood & Eagly, 2012). However, our null findings regarding sex differences in trust do not support these theories, calling for future studies to extend this line of research. Differences in findings could be due to utilization of different versions of the trust game across the different studies. For example, the meta-analysis of Van den Akker et al. (2020), analyzed one-shot trust games. Consequently, they have no opportunity to build a social relationship with their partner, resulting in a more uncertain and potentially riskier decision environment. It is therefore possible that sex differences in trust behavior are more pronounced in one-shot trust games. However, this hypothesis remains speculative and should be examined in future research by directly comparing one-shot and multi-round versions of the trust game.

Limitations

In the current study, we pooled data from six different trust game projects which allowed us to examine age- and sex-related differences in trust from early adolescence to late mid-adulthood. While this paper gives novel insights, the findings need to be interpreted in the light of some limitations. First, it is important to consider that developmental patterns can only be investigated with longitudinal data, and that the current findings are based on cross-sectional data. Nevertheless, the age-related pattern of adaptive trust that we found toward an untrustworthy partner is in line with an earlier longitudinal study in adolescents (Sijtsma, Lee, et al., 2023: [do note that the age range in this study was 11–16 years]). Future research should examine developmental patterns of trust using larger and more generalizable samples and preferably longitudinal designs. Second, it is important to note that the current study included relatively more individuals under 30 than above. This may have resulted in less precise estimates of effects in mid-adulthood, meaning that future research is

needed to expand and corroborate these findings. Third, it is possible that further changes in trust happen at an older age, outside the age range we investigated here. This is suggested by work of Bailey and Leon (2019) who found that older adults were more trusting toward interaction partners with untrustworthy reputations than young adults. Fourth, information indicating whether participants in the TAP study belonged to the same classes in their respective schools were not available, and potential dependencies could not be modeled. However, a previous study by our group, using similar trust game data collected among schoolchildren, found that the intraclass correlation within classes was sufficiently low that class-level clustering did not need to be taken into account (Sijtsma, van Buuren, et al., 2023). Fifth, although the trust game is a well-established experimental paradigm entailing an interactive aspect of social trusting behavior between two individuals, the design has limitations. The trust game does not capture all forms of trust, e.g., generalized trust or trust in institutions, which are measured by self-report questionnaires. However, the trust game design is still highly valued because it measures actual trust behavior while people are engaged in a back-and-forth social interaction. Finally, although results based on the European social survey showed that levels of generalized and social trust in the UK and the Netherlands are similar (Glatz & Eder, 2020; Reeskens & Hooghe, 2008), cultural norms, for example regarding implicit or explicit communication style, might still impact trusting behavior (Rheu et al., 2021). For our sample, sensitivity analyses did not show national differences in trust (see Supplemental Material); however, we did not collect data on cultural values and norms. We advise future research to examine the possible impact of culture on trust within social interactions.

Conclusion

To conclude, the results of the current study suggest that age-related associations with trust from early adolescence to late mid-adulthood are specific to initial trust and adaptive reductions in trust during interactions with untrustworthy others. They further suggest that interacting with humans as opposed to computers matters for the level of expressed trust, which has important implications in a world that relies increasingly on AI.

Acknowledgements

We would like to express our sincere gratitude to Prof. Dr. Lydia Krabbendam for her invaluable scientific guidance during the development of this manuscript.

ORCID iDs

H Sijtsma  <https://orcid.org/0000-0002-1820-0559>

ILJ Lemmers-Jansen  <https://orcid.org/0000-0001-9044-6468>

M Van Buuren  <https://orcid.org/0000-0002-9330-5617>

A-K Fett  <https://orcid.org/0000-0003-0282-273X>

Ethical Approval and Informed Consent Statement

The projects were approved by the local research ethics committees (TAP by London-Surrey Borders; Trust by Barking and Havering; DECOP by London-Harrow; COGPID by the Medical Research Ethics Committee of the University Medical Centre Maastricht; Beslist by the Medical Research Ethics Committee of the Vrije Universiteit Medical Centre Amsterdam; GROUP by the Ethical Review Board of the University Medical Centre Utrecht). Written informed consent was obtained from all participants and their parents/guardians if they were under the age of 16 years.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data Availability Statement

During the time when data was collected, participants did not provide consent to public archiving of the research data; therefore, the data are not stored in a public repository.

Supplemental Material

Supplemental material for this article is available online.

References

- Andrews, J. L., Ahmed, S. P., & Blakemore, S.-J. (2021). Navigating the social environment in adolescence: The role of social brain development. *Biological Psychiatry*, 89(2), 109–118. <https://doi.org/10.1016/j.biopsych.2020.09.012>
- Bailey, P. E., & Leon, T. (2019). A systematic review and meta-analysis of age-related differences in trust. *Psychology and Aging*, 34(5), 674–685. <https://doi.org/10.1037/pag0000368>
- Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, 10(1), 122–142. <https://doi.org/10.1006/game.1995.1027>
- Borghans, L., Heckman, J. J., Golsteyn, B. H. H., & Meijers, H. (2009). Gender differences in risk aversion and ambiguity aversion. *Journal of the European Economic Association*, 7(2-3), 649–658. <https://doi.org/10.1162/JEEA.2009.7.2-3.649>
- Buchan, N. R., Croson, R. T. A., & Solnick, S. (2008). Trust and gender: An examination of behavior and beliefs in the investment game. *Journal of Economic Behavior & Organization*, 68(3-4), 466–476. <https://doi.org/10.1016/j.jebo.2007.10.006>
- Coltrane, S. (2007). Engendering children. In G. Handel (Ed.), *Childhood socialization* (pp. 279–310). AldineTransaction.

- Crone, E. A., & Fuligni, A. J. (2020). Self and others in adolescence. *Annual Review of Psychology*, 71(1), 447–469. <https://doi.org/10.1146/annurev-psych-010419-050937>
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
- Derks, J., Lee, N. C., & Krabbendam, L. (2014). Adolescent trust and trustworthiness: Role of gender and social value orientation. *Journal of Adolescence*, 37(8), 1379–1386. <https://doi.org/10.1016/j.adolescence.2014.09.014>
- Deutsch Lezak, M., Howieson, D. B., Bigler, E. D., et al. (2012). *Neuropsychological assessment*. Oxford University Press.
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3), 522–550. <https://doi.org/10.1111/j.1542-4774.2011.01015.x>
- Dorough, A. R., & Glöckner, A. (2019). A cross-national analysis of sex differences in prisoner's dilemma games. *British Journal of Social Psychology*, 58(1), 225–240. <https://doi.org/10.1111/bjso.12287>
- Duncan, C., Sganzerla, L., Kaltwasser, L., & Dziobek, I. (2025). Trust learning in the repeated trust game: A meta-analytic study. *British Journal of Psychology*, 117(3), 1026–1055. <https://doi.org/10.1111/bjop.70045>
- Eagly, A. H., & Wood, W. (2011). Social role theory. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (pp. 458–476). Sage Publications Ltd.
- Erikson, E. H. (1963). *Childhood and society*. WW Norton & Company.
- Fareri, D. S., Chang, L. J., & Delgado, M. R. (2012). Effects of direct social experience on trust decisions and neural reward circuitry. *Frontiers in Neuroscience*, 6, 148. <https://doi.org/10.3389/fnins.2012.00148>
- Fett, A.-K. J., Gromann, P. M., Giampietro, V., Shergill, S. S., & Krabbendam, L. (2014). Default distrust? An fMRI investigation of the neural development of trust and cooperation. *Social Cognitive and Affective Neuroscience*, 9(4), 395–402. <https://doi.org/10.1093/scan/nss144>
- Fett, A.-K. J., Shergill, S. S., Gromann, P. M., Dumontheil, I., Blakemore, S.-J., Yakub, F., & Krabbendam, L. (2014). Trust and social reciprocity in adolescence – A matter of perspective-taking. *Journal of Adolescence*, 37(2), 175–184. <https://doi.org/10.1016/j.adolescence.2013.11.011>
- Fett, A.-K. J., Shergill, S. S., Joyce, D. W., Riedl, A., Strobel, M., Gromann, P. M., & Krabbendam, L. (2012). To trust or not to trust: The dynamics of social interaction in psychosis. *Brain*, 135(3), 976–984. <https://doi.org/10.1093/brain/awr359>
- Glatz, C., & Eder, A. (2020). Patterns of trust and subjective well-being across Europe: New insights from repeated cross-sectional analyses based on the European social survey 2002–2016. *Social Indicators Research*, 148(2), 417–439. <https://doi.org/10.1007/s11205-019-02212-x>
- Gromann, P., Shergill, S., De Haan, L., Meewis, D. G. J., Fett, A.-K. J., Korver-Nieberg, N., & Krabbendam, L. (2014). Reduced brain reward response during cooperation in first-degree relatives of patients with psychosis: An fMRI study. *Psychological Medicine*, 44(16), 3445–3454. <https://doi.org/10.1017/S0033291714000737>
- Gromann, P. M., Heslenfeld, D. J., Fett, A.-K., Joyce, D. W., Shergill, S. S., & Krabbendam, L. (2013). Trust versus paranoia: Abnormal response to social reward in psychotic illness. *Brain*, 136(6), 1968–1975. <https://doi.org/10.1093/brain/awt076>
- Hall, H. K., Milleer, P. M. R., Summers, M. J., & Isbel, B. (2021). Longitudinal research on perspective taking in adolescence: A systematic review. *Adolescent Research Review*, 6, 125–150. <https://doi.org/10.1007/s40894-021-00150-9>
- Hanssen, E., Van Buuren, M., Van Atteveldt, N., Lemmers-Jansen, I. L. J., & Fett, A.-K. J. (2022). Neural, behavioural and real-life correlates of social context sensitivity and social reward learning during interpersonal interactions in the schizophrenia spectrum. *Australian & New Zealand Journal of Psychiatry*, 56(1), 59–70. <https://doi.org/10.1177/00048674211010327>
- Haselhuhn, M. P., Kennedy, J. A., Kray, L. J., Van Zant, A. B., & Schweitzer, M. E. (2015). Gender differences in trust dynamics: Women trust more than men following a trust violation. *Journal of Experimental Social Psychology*, 56, 104–109. <https://doi.org/10.1016/j.jesp.2014.09.007>
- Holm, H., & Nystedt, P. (2005). Intra-generational trust—a semi-experimental study of trust among different generations. *Journal of Economic Behavior & Organization*, 58(3), 403–419. <https://doi.org/10.1016/j.jebo.2003.10.013>
- Houser, D., Schunk, D., & Winter, J. (2010). Distinguishing trust from risk: An anatomy of the investment game. *Journal of Economic Behavior & Organization*, 74(1-2), 72–81. <https://doi.org/10.1016/j.jebo.2010.01.002>
- Jin, S., Yan, S., & Zhang, X. (2025). Measuring trust across countries: Inconsistencies between experiments and surveys. *Economics Letters*, 248, 112213. <https://doi.org/10.1016/j.econlet.2025.112213>
- Johnson, N. D., & Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology*, 32(5), 865–889. <https://doi.org/10.1016/j.joep.2011.05.007>
- Korver, N., Quee, P. J., Boos, H. B., Simons, C. J. P., & de Haan, L. (2012). Genetic Risk and Outcome of Psychosis (GROUP), a multi site longitudinal cohort study focused on gene–environment interaction: Objectives, sample characteristics, recruitment and assessment methods. *International Journal of Methods in Psychiatric Research*, 21(3), 205–221. <https://doi.org/10.1002/mpr.1352>
- Krabbendam, L., Sijtsma, H., Crone, E. A., & van Buuren, M. (2024). Trust in adolescence: Development, mechanisms and future directions. *Developmental Cognitive Neuroscience*, 69, 101426. <https://doi.org/10.1016/j.dcn.2024.101426>
- Krueger, F., & Meyer-Lindenberg, A. (2019). Toward a model of interpersonal trust drawn from neuroscience, psychology, and

- economics. *Trends in Neurosciences*, 42(2), 92–101. <https://doi.org/10.1016/j.tins.2018.10.004>
- Larsen, B., & Luna, B. (2018). Adolescence as a neurobiological critical period for the development of higher-order cognition. *Neuroscience & Biobehavioral Reviews*, 94, 179–195. <https://doi.org/10.1016/j.neubiorev.2018.09.005>
- Lee, N. C., Jolles, J., & Krabbendam, L. (2016). Social information influences trust behaviour in adolescents. *Journal of Adolescence*, 46(1), 66–75. <https://doi.org/10.1016/j.adolescence.2015.10.021>
- Leković, V. (2012). Trust as an institutional factor of economic success. *Ekonomski horizonti*, 14(2), 63–75. <https://doi.org/10.5937/ekonhor1202063L>
- Lemmers-Jansen, I. L. J., Fett, A.-K. J., Hanssen, E., Veltman, D. J., & Krabbendam, L. (2019). Learning to trust: Social feedback normalizes trust behavior in first-episode psychosis and clinical high risk. *Psychological Medicine*, 49(5), 780–790. <https://doi.org/10.1017/S003329171800140X>
- Lemmers-Jansen, I. L. J., Fett, A.-K. J., Shergill, S. S., van Kesteren, M. T. R., & Krabbendam, L. (2019). Girls-boys: An investigation of gender differences in the behavioral and neural mechanisms of trust and reciprocity in adolescence. *Frontiers in Human Neuroscience*, 13, 257. <https://doi.org/10.3389/fnhum.2019.00257>
- Lemmers-Jansen, I. L. J., Krabbendam, L., Veltman, D. J., & Fett, A.-K. J. (2017). Boys vs. girls: Gender differences in the neural development of trust and reciprocity depend on social context. *Developmental Cognitive Neuroscience*, 25, 235–245. <https://doi.org/10.1016/j.dcn.2017.02.001>
- Lewicki, R. J., Tomlinson, E. C., & Gillespie, N. (2006). Models of interpersonal trust development: Theoretical approaches, empirical evidence, and future directions. *Journal of Management*, 32(6), 991–1022. <https://doi.org/10.1177/0149206306294405>
- Li, Y., Wu, B., Huang, Y., & Luan, S. (2024). Developing trustworthy artificial intelligence: Insights from research on interpersonal, human-automation, and human-AI trust. *Frontiers in Psychology*, 15, 1382693. <https://doi.org/10.3389/fpsyg.2024.1382693>
- Ma, I., Westhoff, B., & van Duijvenvoorde, A. C. K. (2020). The cognitive mechanisms that drive social belief updates during adolescence. *BioRxiv*. <https://doi.org/10.1101/2020.05.19.105114>
- Marangunić, N., & Granić, A. (2015). Technology acceptance model: A literature review from 1986 to 2013. *Universal Access in the Information Society*, 14(1), 81–95. <https://doi.org/10.1007/s10209-014-0348-1>
- Misztal, B. A. (1996). *Trust in modern societies: The search for the bases of social order*. Blackwell Publishers.
- Moran, E., Warden, D., Macleod, L., Mayes, G., & Gillies, J. (1997). Stranger-danger: What do children know? *Child Abuse Review: Journal of the British Association for the Study and Prevention of Child Abuse and Neglect*, 6(1), 11–23. [https://doi.org/10.1002/\(SICI\)1099-0852\(199703\)6:1<11::AID-CAR286>3.0.CO;2-G](https://doi.org/10.1002/(SICI)1099-0852(199703)6:1<11::AID-CAR286>3.0.CO;2-G)
- Murtin, F., Fleischer, L., Siegerink, V., Aassve, A., Algan, Y., Boarini, R., & Vallve, R. H. (2018). Trust and its determinants: Evidence from the trustlab experiment. *Organisation for Economic Cooperation and Development*, 2, 1–74. <https://doi.org/10.1787/869ef2ec-en>
- Newton, K. (2001). Trust, social capital, civil society, and democracy. *International Political Science Review*, 22(2), 201–214. <https://doi.org/10.1177/0192512101222004>
- Nussenbaum, K., & Hartley, C. A. (2019). Reinforcement learning across development: What insights can we draw from a decade of research? *Developmental Cognitive Neuroscience*, 40, 100733. <https://doi.org/10.1016/j.dcn.2019.100733>
- Reeskens, T., & Hooghe, M. (2008). Cross-cultural measurement equivalence of generalized trust: Evidence from the European Social Survey (2002 and 2004). *Social Indicators Research*, 85(3), 515–532. <https://doi.org/10.1007/s11205-007-9100-z>
- Reinecke, M. G., Kappes, A., Porsdam Mann, S., Savulescu, J., & Earp, B. D. (2025). The need for an empirical research program regarding human-AI relational norms. *AI and Ethics*, 5(1), 71–80. <https://doi.org/10.1007/s43681-024-00631-2>
- Rheu, M., Shin, J. Y., Peng, W., & Huh-Yoo, J. (2021). Systematic review: Trust-building factors and implications for conversational agent design. *International Journal of Human-Computer Interaction*, 37(1), 81–96. <https://doi.org/10.1080/10447318.2020.1807710>
- Sijtsma, H., Lee, N., Braams, B., Hollarek, M., Walsh, R. J., van Buuren, M., & Krabbendam, L. (2023). The development of adolescent trust behavior. *Journal of Experimental Child Psychology*, 231, 105653. <https://doi.org/10.1016/j.jecp.2023.105653>
- Sijtsma, H., van Buuren, M., Hollarek, M., Walsh, R. J., Lee, N. C., Braams, B. R., & Krabbendam, L. (2023). Social network position, trust behavior, and neural activity in young adolescents. *Neuroimage*, 268, 119882. <https://doi.org/10.1016/j.neuroimage.2023.119882>
- Simpson, J. A., & Vieth, G. (2021). Trust and psychology: Psychological theories and principles underlying interpersonal trust. In F. Krueger (Ed.), *The neurobiology of trust* (pp. 15–35). Cambridge University Press.
- Stolle, D. (2002). Trusting strangers – the concept of generalized trust in perspective. *Österreichische Zeitschrift für Politikwissenschaft*, 31(4), 397–412. <https://doi.org/10.15203/ozp.814.vol31iss4>
- Sutter, M., & Kocher, M. G. (2007). Trust and trustworthiness across different age groups. *Games and Economic Behavior*, 59(2), 364–382. <https://doi.org/10.1016/j.geb.2006.07.006>
- Sydnor, V. J., Larsen, B., Bassett, D. S., Alexander-Bloch, A., Fair, D. A., Liston, C., Mackey, A. P., Milham, M. P., Pines, A., Roalf, D. R., Seidlitz, J., Xu, T., Raznahan, A., & Satterthwaite, T. D. (2021). Neurodevelopment of the association cortices: Patterns, mechanisms, and implications for psychopathology. *Neuron*, 109(18), 2820–2846. <https://doi.org/10.1016/j.neuron.2021.06.016>
- Telga, M., & Lupiáñez, J. (2021). Older and younger adults perform similarly in an iterated trust game. *Frontiers in*

- Psychology*, 12, 747187. <https://doi.org/10.3389/fpsyg.2021.747187>
- Thellman, S., Ageskär, O., Allander, I., Degerstedt, M., Hyland, O., Nguyen, P., Nääs, H., & Wickman, N. (2023). A survey of attributions and preferences regarding higher-order mental states in artificial agents. *Proceedings of the 11th international conference on human-agent interaction* (pp. 97-104).
- Thellman, S., De Graaf, M., & Ziemke, T. (2022). Mental state attribution to robots: A systematic review of conceptions, methods, and findings. *ACM Transactions on Human-Robot Interaction*, 11(4), 1–51. <https://doi.org/10.1145/3526112>
- The Office for National Statistics. (2023). *Age classifications: Census 2021*. Retrieved February 3, from <https://www.ons.gov.uk/census/census2021dictionary/variablesbytopic/demographyvariables/census2021age/classifications>
- Uslaner, E. M. (2002). *The moral foundations of trust*. Cambridge University Press.
- Valle, A., Massaro, D., Castelli, I., & Marchetti, A. (2015). Theory of mind development in adolescence and early adulthood: The growing complexity of recursive thinking ability. *Europe's Journal of Psychology*, 11(1), 112–124. <https://doi.org/10.5964/ejop.v11i1.829>
- Van de Groep, S., Meuwese, R., Zanolie, K., Güroğlu, B., & Crone, E. A. (2020). Developmental changes and individual differences in trust and reciprocity in adolescence. *Journal of Research on Adolescence*, 30, 192–208. <https://doi.org/10.1111/jora.12459>
- Van den Akker, O. R., Van Assen, M. A., Van Vugt, M., & Wicherts, J. M. (2020). Sex differences in trust and trustworthiness: A meta-analysis of the trust game and the gift-exchange game. *Journal of Economic Psychology*, 81, 102329. <https://doi.org/10.1016/j.joep.2020.102329>
- Van den Bos, W., Van Dijk, E., & Crone, E. A. (2012). Learning whom to trust in repeated social interactions: A developmental perspective. *Group Processes & Intergroup Relations*, 15(2), 243–256. <https://doi.org/10.1177/1368430211418698>
- Van der Graaff, J., Branje, S., De Wied, M., Hawk, S., Van Lier, P., & Meeus, W. (2014). Perspective taking and empathic concern in adolescence: Gender differences in developmental changes. *Developmental Psychology*, 50(3), 881–888. <https://doi.org/10.1037/a0034325>
- Wang, Q., & Goel, A. K. (2022). Mutual theory of mind for human-AI communication. *arXiv preprint arXiv:2210.03842*.
- Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised*. Psychological Corporation.
- Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. Harcourt Assessment.
- Westhoff, B., Molleman, L., Viding, E., van den Bos, W., & van Duijvenvoorde, A. C. K. (2020). Developmental asymmetries in learning to adjust to cooperative and uncooperative environments. *Scientific Reports*, 10, 21761. <https://doi.org/10.1038/s41598-020-78546-1>
- Wood, W., & Eagly, A. H. (2012). Biosocial construction of sex differences and similarities in behavior. In J. M. Olson & M. P. Zanna (Eds.), *Advances in experimental social psychology* (pp. 55–123). Elsevier Academic Press.
- Yang, M., & Rau, P.-L. P. (2024). Trust building with artificial intelligence: Comparing with human in investment behaviour, emotional arousal and neuro activities. *Theoretical Issues in Ergonomics Science*, 25(5), 593–614. <https://doi.org/10.1080/1463922X.2023.2290019>

Author Biographies

Hester Sijtsma is a Postdoctoral Researcher. Her research focuses on the development of social cognition, social relationships, and wellbeing in adolescents.

ILJ Lemmers-Jansen is an Associate Professor in Clinical Developmental and Neuropsychology. Her research evolved from basic fMRI research in early psychosis, focusing on social interactions, to more applied research into social functioning and social networks. Currently, one line of her research applies qualitative investigations into the development of psychosis.

M Van Buuren is an Associate Professor in Clinical Developmental Psychology. Her research is aimed at the development of social cognition and self-concept in adolescence. Specifically, her research is focused on the underlying brain mechanisms and the association with social functioning and psychopathology.

A-K Fett is a Reader in Clinical and Biological Psychology. She holds a PhD in the developmental neuro-cognitive mechanisms underlying social cognition and functioning in health and psychosis. Her research focuses on the mechanisms that shape social functioning, including trust, loneliness, and social isolation, in both the general population and individuals with psychotic disorders and other mental health conditions.