

Author Manuscript

For Open Access version of the published paper see:

<https://doi.org/10.1080/10447318.2022.2085400>

The General Attitudes towards Artificial Intelligence Scale (GAAIS): Confirmatory Validation and Associations with Personality, Corporate Distrust, and General Trust

Astrid Schepman and Paul Rodway

School of Psychology, University of Chester.

Author note

Astrid Schepman <https://orcid.org/0000-0002-7407-362X>

Paul Rodway <https://orcid.org/0000-0002-7667-6782>

There are no conflicts of interest to declare

Correspondence concerning this article should be addressed to Dr Astrid Schepman, School of Psychology, University of Chester, Parkgate Road, Chester, Cheshire, CH1 4BJ, United Kingdom. Email: a.schepman@chester.ac.uk

Abstract

Acceptance of Artificial Intelligence may be predicted by individual psychological correlates, examined here. Study 1 reports confirmatory validation of the General Attitudes towards Artificial Intelligence Scale (GAAIS) following initial validation elsewhere. Confirmatory Factor Analysis confirmed the two-factor structure (Positive, Negative) and showed good convergent and divergent validity with a related scale. Study 2 tested whether psychological factors (Big Five personality traits, corporate distrust, and general trust) predicted attitudes towards AI. Introverts had more positive attitudes towards AI overall, likely because of algorithm appreciation. Conscientiousness and agreeableness were associated with forgiving attitudes towards negative aspects of AI. Higher corporate distrust led to negative attitudes towards AI overall, while higher general trust led to positive views of the benefits of AI. The dissociation between general trust and corporate distrust may reflect the public's attributions of the benefits and drawbacks of AI. Results are discussed in relation to theory and prior findings.

Keywords: Artificial Intelligence; Human-AI interaction; Personality; Individual Differences; Trust; Corporations

Introduction

Background

Society will be profoundly impacted by Artificial Intelligence (AI) over the next few decades (e.g. Makridakis, 2017; Olhede & Wolfe, 2018; Vesnic-Alujevic et al., 2020). Unlike computing devices and applications for individual use introduced over the last few decades, decisions regarding the adoption of AI do not rest wholly with individuals (Brownsword & Harel, 2019; Misra et al., 2020). Individuals have less choice over the adoption of AI than over the adoption of a laptop computer or smartphone, because other people, namely those working for large corporations or for governments, will make many of the decisions to introduce the AI technologies (Chen & Wen, 2021; Jones et al., 2018). This difference in user control impacts strongly on people's attitudes towards AI, as compared to older information technologies, with different perceived advantages, concerns, and risks (e.g. Anderson et al., 2018; Carrasco, et al., 2019; Cave et al., 2019; Edelman; 2019; Fast & Horvitz; 2017; Royal Society Working Group, 2017, Stephanidis et al., 2019; Triberti et al., 2021; Zhang & Dafoe, 2019). It is therefore important to gauge attitudes towards AI specifically, and to study what the psychological correlates of attitudes towards AI are.

Our primary aim was to report on new research examining these psychological correlates. We did this using an instrument that had been initially validated, the General Attitudes Towards Artificial Intelligence Scale (GAAIS, Schepman & Rodway, 2020), but that had not yet undergone confirmatory validation. Initial validation had identified two subscales: Positive and Negative. Confirmatory validation of this factor structure of the GAAIS was an important precursor, and performing this was a further important aim of the current research.

Technology and personality

A common source of individual differences in behaviors or attitudes is personality. Personality traits reflect consistent attitudes and behavioral patterns in individuals. A widely-used classification system for personality traits is the Big Five (Costa & McCrae, 1992; McCrae & Costa, 2008). This consists of five dimensions, often abbreviation as OCEAN, namely Openness to experience / open-mindedness (defined by Costa and McCrae as intellect, imagination and culture), Conscientiousness (will to achieve), Extraversion (urgency), Agreeableness (the opposite of antagonism), and Neuroticism / negative emotionality (emotional instability). Their construct validity and the replicability of their predictive powers are well-established (Soto, 2019). Because personality can predict many aspects of attitudes in a range of domains (e.g. McManus et al., 2004; Milfont, & Sibley, 2012), it was important to establish whether it had predictive value in relation to attitudes towards AI.

Prior findings: Personality and technology

Personality influences people's technology acceptance, but not always in the same way. The personality dimensions that correlate with technology evaluations, and the direction of the correlations, can vary depending on the technology domain and type of attitude measured (e.g. acceptance, perceived usability, intention to use, etc.). This becomes evident when examining prior research, from which we selected the research that was most relevant to AI. For example, Devaraj et al., (2008) explored links between personality traits and the components of the Technology Acceptance Model (Davis, 1989) as applied to collaborative technology. They found that perceived usefulness showed positive associations with

agreeableness, and negative associations with neuroticism, and that conscientiousness and extraversion acted as significant positive moderating variables. Interestingly, the technology domain had a social function in this study and this may have played a role in the observed associations. Smartphone ownership and usage were positively associated with extraversion and agreeableness (Lane & Manner, 2011), most likely also because of the social uses that smartphones provide. In contrast, though, Zhou and Lu (2011) found no association between perceived usefulness of mobile commerce and extraversion (in their work measured as talkativeness, energy and passion). Further research by Svendsen et al. (2013) also linked personality to components of the Technology Acceptance Model (Davis, 1989). In Svendsen et al.'s (2013) study the technology domain was a cross-platform application for handling music and pictures, including sharing with friends. Svendsen et al. found that technology acceptance beliefs, and the behavioral intention to use technology were positively associated with extraversion, that openness to experience was related to perceived ease of use, and that emotional stability (the reverse of negative emotionality / neuroticism) was positively related to behavioral intentions to use technology. As with Devaraj et al. (2008), the social aspect of this application may have led to a positive association with extraversion.

Barnett et al. (2015), probed interactions with learning technology (lecture notes, communications) and associations with personality. In replication, but also in contrast to some of the findings described above, Barnett et al. (2015) found a positive association between perceived or actual use and conscientiousness, and a *negative* association between perceived or actual use and neuroticism, as well as (unexpectedly) extraversion, with no significant roles for openness or agreeableness. An interesting difference in relation to earlier findings is that the association with extraversion was negative in this study. This may be because the main role of the technology was providing information, and the technology served less of a social function.

In the broad domain of everyday technology products (e.g. Google search, Microsoft Word, a microwave oven, a can opener, etc.) Kortum and Oswald (2018) found that openness to experience and agreeableness were positive predictors of technology usability ratings. Here, a different combination of personality dimensions came to the fore as predictors of technology evaluations, potentially due to the statistical smoothing effect of the different functions of the wide range of technologies, and possibly due to the focus on usability rather than other types of evaluation.

As observed, links between specific personality traits and technology can depend on the type and function of the technology involved and the aspects of the technology that were evaluated. For example, technology whose main function is to facilitate users showing themselves to audiences (e.g. TikTok) tends to be liked by people who are extravert (e.g. Meng & Leung, 2021), because such technologies support their personal needs and drives. More introverted people may instead prefer technologies that allow them to navigate situations without unnecessary social interactions, such as self-service check-outs (e.g. Lee et al., 2010).

Associations between specific technologies / technology-related behaviors and personality traits have also been observed for the other personality dimensions. For example, Shropshire et al. (2015) found that agreeableness and conscientiousness were positively related to the use of information security behaviors, and Qu et al. (2020) showed positive associations between acceptance of self-driving cars and the traits of openness and extraversion, with neuroticism showing negative associations.

An important take-home message from this sample of prior research is that there is clear evidence of links between personality traits and measures of technology acceptance, but our discussion of the prior literature has demonstrated that associations between personality

and attitudes towards technology are complex and context-dependent. Technologies with different functions may give rise to differing association patterns with personality traits. For this reason, it is important to obtain data that are specific to each technology.

An important aim of our research was to examine the link between general attitudes towards Artificial Intelligence and personality traits. This precise link has not been examined previously and, as argued, cannot be extrapolated from prior research, due to the context-dependence and complexity of the association patterns observed in prior research.

Hypotheses regarding associations between attitudes towards AI and personality

Our personality-related hypotheses link the two subscales of the GAAIS (Positive and Negative) to personality traits. The Positive GAAIS measures positive attitudes towards AI, including perceptions of utility (e.g. economic opportunities, improved performance), desired use (e.g. at work), and positive emotions (e.g. excitement, being impressed). The Negative GAAIS measures Negative attitudes, featuring concerns about AI (e.g. unethical use, making errors) and negative emotions (e.g. discomfort, finding AI sinister). The GAAIS is scored in such a way that for both subscales higher scores indicate more positive attitudes towards AI. Our chosen personality measure was the Big Five Inventory-2 Short Form (BFI-2-S; Soto & John, 2017). The higher the score, the stronger the presence of the named construct.

Because there was no specific research examining associations between general attitudes towards AI and personality traits, our hypotheses were based on general reasoning and prior research (notwithstanding some variations in prior research findings). We tested the hypotheses via their significance as coefficients in the full model of two hierarchical multiple regression analyses, one for each subscale of the GAAIS. Full details follow, but the

hierarchical analysis included a block of demographics, a block of the five personality traits, and a block of the two trust measures.

Our first hypothesis was that open-mindedness (Intellectual Curiosity, Aesthetic Sensitivity, Creative Imagination, as defined by Soto & John, 2017) would be a positive predictor of the Positive GAAIS, which captures excitement regarding AI. It would seem reasonable that a technology as innovative as AI might be viewed positively by open-minded individuals. As discussed, this trait has been linked to some aspects of technology and technology adoption before (e.g. Kortum & Oswald, 2018; Qu et al, 2020; Svendsen et al, 2013, but see Barnett et al. 2015 for a non-association). Thus, $H_{open-mindedness}$: *Open-mindedness is a significant positive predictor of the Positive GAAIS.*

We also hypothesized that negative emotionality (Anxiety, Depression, and Emotional Volatility; Soto & John, 2017) would be a negative predictor of the Negative GAAIS, with the Negative GAAIS reflecting concerns about AI, and low scores on the Negative GAAIS indicating high levels of concern. This link has been found with other major innovative technologies that carry a potential risk (e.g. autonomous vehicles, see Qu et al, 2020), and more general technologies (Barnett et al., 2015; Devaraj et al., 2008; Svendsen et al., 2013). Thus, $H_{negative\ emotionality}$: *Negative emotionality is a significant negative predictor of the Negative GAAIS.*

AI can enhance productivity, and both subscales of the GAAIS were therefore hypothesized to associate positively with conscientiousness (Organization, Productiveness, Responsibility; Soto & John, 2017). Prior research has shown that this personality trait to be positively associated with a range of technology evaluations (e.g. Barnett et al, 2015; Devaraj et al, 2008; Shropshire et al, 2015). Thus, $H_{conscientiousness}$: *Conscientiousness is a significant positive predictor of the Positive and Negative GAAIS.*

It could also be hypothesized that attitudes towards AI may be more positive in those who are more socially compliant, respectful, and trusting, which are aspects of agreeableness (Compassion, Respectfulness, Trust; Soto & John, 2017). This notion gives rise to a hypothesis of a positive association between both GAAIS subscales and agreeableness. As discussed, a related association was found by Devaraj et al. (2008), Lane and Manner (2011), Shropshire et al. (2015), and Kortum and Oswald (2018), but not by Barnett et al (2015). Thus, $H_{\text{agreeableness}}$: *Agreeableness is a significant positive predictor of the Positive and Negative GAAIS.*

In light of contradictory prior findings on the link between extraversion (Sociability, Assertiveness, Energy Level; Soto & John, 2017) and technology acceptance, key aspects of AI were carefully considered in the hypothesis formulation, and two contrasting hypotheses were derived. Based on the findings of algorithm appreciation (Logg, Minson, & Moore, 2019), and the idea that AI may replace people in some routine tasks, introverts may have more favorable attitudes towards AI than extraverts (see also Barnett et al, 2015; Lee et al, 2020). Thus: $H_{\text{extraversion-neg}}$: *Extraversion is a significant negative predictor of the Positive and Negative GAAIS.*

A contrasting hypothesis was derived from prior findings with other technologies, which found positive associations between extraversion and technology evaluations (e.g. Devaraj et al., 2008; Lane & Manner, 2011; Qu et al., 2020; Svendsen et al., 2013). Based on these prior findings, extraverts may have had positive views of many new technologies, because engaging with them and talking about them may have allowed them to gain social attention (Ashton et al., 2002). Such a perspective gave rise to the hypothesis that there would be a positive association between the GAAIS and extraversion. Thus, $H_{\text{extraversion-pos}}$: *Extraversion is a significant positive predictor of the Positive and Negative GAAIS.*

Artificial Intelligence and Trust

Artificial Intelligence raises issues of trust, and this is a much-researched aspect of AI (see e.g. Araujo et al., 2020; Rheu et al., 2021; Siau & Wang, 2018, for recent examples). In brief, findings to date show that the general public seems willing to accept AI, sometimes even preferring it to humans (e.g. Logg, et al., 2019), while at the same time having apprehensions about important aspects of AI and how it is used (e.g. Yokoi et al., 2021; Schepman & Rodway, 2020).

Trust has been studied as a predictor of risk perception in relation to the acceptance of new technologies. In a recent review, Siegrist (2021) argued that trust indicates a willingness to accept the vulnerability of having risks controlled by a third party. Siegrist reviews findings of this relationship in the context of e.g. nuclear power, climate change, or gene technology, where end users are likely to have no choice but to trust a third party, because their own knowledge is not sufficient to calculate and manage the risks, and because they have no power to manage the risk themselves. The level of trust they exhibit is related to their rating of the values of those entrusted with the risk management, in line with the Salient Value Similarity Model (SVS, Earle & Cvetkovich, 1995). A very similar situation is likely to apply with respect to AI. For our purposes, the concept of trust towards AI, or other forms of technology or automation requires further decomposition, because it is multi-faceted, and different aspects may link differently to attitudes towards AI.

Aspects of trust and their links to attitudes towards AI

The first aspect of trust we consider is a sense of reliance on the AI. This aspect of trust has been researched for many years, not just in relation to AI, but also other types of automation (see e.g. Muir & Moray, 1996). Reliance-based trust is confidence that the technology will deliver on the task for which it was designed. This may include elements of functionality, consistency, and safety. Autonomous vehicles may be life-enhancing because they may preserve the mobility of those who can no longer drive to due old age or poor health (Charness et al., 2018), but users have to be willing to entrust their lives to this technology (Liu & Liu, 2021), and this is based on confidence in its accurate functioning (Kaur & Rampersad, 2018). Similarly, AI used in health settings raises issues of confidence around its performance (Asan et al., 2020; Baldauf et al., 2020). If the AI provides erroneous outcomes, lives could be at stake. According to Ryan (2020) the use of the term “trust” in this context is a misnomer, because trust is inherently tied to interpersonal relations between animate beings who experience emotions (see also DeCamp & Tilburt, 2019). However, as is evident from the citations above, the term “trust” with this sense is in wide use, notwithstanding Ryan’s suggestions. We have embedded this aspect of trust in AI in the GAAIS (e.g. via the item “I think artificially intelligent systems make many errors”).

A further aspect of trust is linked to explainability (see e.g. EU, 2020; Schmidt, et al., 2020; Stanton & Jensen, 2021). This involves creating AI that is explainable by design, whose workings can be understood by humans, so that humans can stay in control. According to the EU (2020), making the AI designs less of a black box will engender trust. They suggest that this may involve less reliance on algorithms and more “human-guided symbolic learning”. Smith (2021) observed that AI opacity can also cause issues with legal liabilities, while Walmsley (2021) advocated for multiple forms of AI transparency. In all, this aspect of

trust is about knowing what the AI does, so that it can be controlled by humans. The GAAIS captures the dimension of controllability of AI via the item “Artificial Intelligence might take control of people” and the more emotional items “I find Artificial Intelligence sinister” and “I think Artificial Intelligence is dangerous” (all from the negative GAAIS).

A further dimension of trust is impacted by the fact that Artificial Intelligence applications are often controlled by large corporations, which may harvest user data to power their AI applications and create profits, without the user being able to give informed consent. Privacy, data protection, and ethical use are important elements of trust in that respect (Stephanidis et al., 2019; Ikkatai, et al., 2022). Drawing a participant sample in Taiwan, Chen and Wen (2021) found a positive association between trust in AI and trust in corporations. Interestingly, in their research the object of trust was the scientific AI community (e.g. “I trust the AI science community to do what is right”), and in that respect it differed from our measure, which measures trust towards AI as a more abstract construct. In our research, this aspect of trust formed part of the Negative GAAIS via the items “Organisations use Artificial Intelligence unethically” and “Artificial Intelligence is used to spy on people”.

In addition to corporations, humans in general can also be objects of trust. All AI applications have been designed by humans. Humans remain important in the ongoing functioning of machine learning, which on the face of it has the capacity to function with less human input (Deng et al., 2020). Because humans are behind every AI application, general trust in people may also be an important determinant of general attitudes towards AI, as it had been with older technologies (Muir & Moray, 1996). After all, the humans who designed the applications are first and foremost still humans, even though their AI products are inanimate, and even though the designers and programmers may work for large corporations which form trust-objects in their own right, as explained above. People’s general trust in people may

transfer to those designing the AI, and it may therefore impact on people's attitudes towards AI.

Hypotheses regarding associations between attitudes towards AI and trust

As explained, some aspects of trust were embedded in the GAAIS, alongside the other attitudinal dimensions that constitute the GAAIS. In addition, we examined to what extent corporate distrust was associated with the GAAIS. We selected an individual trait measure of trust towards corporations. This measure was named "Corporate Distrust" (Adams, et al., 2010; more details follow in the Method of Study 2). The higher the score, the stronger the distrust of corporations. We hypothesized that Corporate Distrust would show a negative association with both subscales of the GAAIS. Thus, $H_{\text{corporate distrust}}$: *Corporate Distrust is a negative predictor of the Positive and Negative GAAIS.*

We also selected a measure of general trust towards people, namely the General Trust Scale (Yamagishi & Yamagishi, 1994; more details in the Method for Study 2). The higher the score on this scale, the higher the participant's general trust in people. Because trust in people in general may transfer to the people who created AI, we hypothesized that scores on the General Trust Scale would show positive associations with both subscales of the GAAIS. Thus, $H_{\text{general trust}}$: *General trust is a positive predictor of the Positive and Negative GAAIS.*

Study 1: Attitudes towards AI and their measurement: Confirmatory validation of the GAAIS

Study 1 had the aim of confirming the factor structure of the General Attitudes towards Artificial Intelligence Scale via Confirmatory Factor Analysis, following initial

validation (Schepman & Rodway, 2020). The initial validation process had involved an Exploratory Factor Analysis (EFA) of data from 100 participants. There had been 32 items reflecting positive and negative attitudes towards AI, ranging from practical benefits, to emotional reactions, and concerns. During the EFA process, the scale was reduced to 20 items. At initial validation, two factors had been identified: Positive and Negative.

A further element of validation had been to establish convergent and discriminant validity against existing scales. In Schepman and Rodway (2020) the GAAIS had been validated against the Technology Readiness Index (TRI; Lam, et al., 2008). This is a scale that measures readiness for consumer technologies, and has been widely used across its various iterations, with the original version (Parasuraman, 2000) having over 3,000 citations on Google Scholar. The updated version of the TRI (Lam, et al., 2008) used in Schepman and Rodway (2020) was partly chosen for its relative brevity. It had 18 items across four subscales, while the original version (Parasuraman, 2000) had 36, which was somewhat long in the context of the study. The GAAIS had shown good convergent and discriminant validity with the TRI at initial validation.

At initial validation, the GAAIS was also validated against exemplars of specific applications of AI that had been assembled from reports in UK newspapers. Participants rated their comfortableness with AI carrying out featured tasks (e.g. “Helping detect life on other planets” and “Providing psychotherapy for patients with phobias”), as well as the perceived capability of AI compared to humans. This showed a pattern of significant correlations between ratings for comfortableness with AI for specific applications and both subscales of the GAAIS. Slightly weaker correlations were shown for the perceived capability ratings and the GAAIS.

In all, the data reported in Schepman and Rodway (2020) served to provide initial validation for the GAAIS, with a focus on the conceptual design, the initial discovery of the factor structure via EFA, the elimination of non-loading or cross-loading items, the testing of construct validity against the TRI and against perceptions of specific AI applications. These were all essential elements of initial scale validation.

However, the process was not complete, because it was important to confirm the factor structure that was discovered using EFA based on the initial validation sample. This is done by drawing a fresh sample of data and subjecting that to Confirmatory Factor Analysis (CFA). This ensures that the factor structure was not specific to the initial sample. This is a prerequisite for the use of the scale to test substantive hypotheses such as our own. Performing this CFA was an important aim of Study 1. We also tested whether convergent and discriminant validity against the TRI replicated in a new and larger sample.

Method

Ethics

This study and Study 2 were reviewed and approved by the School of Psychology Ethics Committee at our institution (Ethics approval codes ASPR300620 AND ASPR250620, respectively) and complied with the British Psychological Society's (2014) Code of Human Research Ethics (2nd edition).

Recruitment, participants and demographic information

We included data from 304 participants, recruited via Prolific.co, an online participant recruitment platform based in the UK. Prolific has been assessed as being a high-quality

crowdsourcing platform in comparison to alternative platforms, as measured via attention checks and the replication of known effects (Peer, et al., 2017). Participants were adult UK residents aged 18 or over. There were 151 women, 151 men, 1 other, 1 prefer not to answer, with a mean age of 35.7 years ($SD = 13.2$ years, range 18-76). Data from three additional people were not included: one male due to incomplete data: 40% of GAAIS responses missing), and a further male, and a female (due to failing two of the three attention checks). None of the participants had participated in the Schepman and Rodway (2020) study. The sample in Schepman and Rodway was targeted at employees, but in this study, that restriction was not used, to enhance generalisability of the findings.

Participants' highest educational qualifications were: 0.3% no formal education; 14.1% GCSE or equivalent (secondary school qualification usually taken aged 16); 28.3% A-Level or equivalent (pre-university secondary school qualification usually taken aged 18); 0.3% Higher National Diploma; 40.8% Bachelor's degree or equivalent; 12.2% Master's degree or equivalent; 2.6% Doctoral degree or equivalent; 0.3% apprenticeship level 3; 0.3% apprenticeship; 0.7% preferred not to answer.

To assess the potential impact of computer expertise (as seen e.g. in Zhang & Dafoe, 2019), we asked participants to self-rate their computer expertise. None of the users in this sample chose "Hardly ever use the computer and do not feel very competent"; 1.3% rated themselves as "Slightly below average computer user, infrequently using the computer, using few applications"; 47% replied "Average computer user, using the internet, standard applications etc."; 36.8% chose "User of specialist applications but not an IT specialist"; 8.2% indicated "Considerable IT expertise short of full professional qualifications"; and 6.6% declared themselves to be "Professionally qualified computer scientist or IT specialist". The wide diversity of occupations in our sample can be seen in the data file available via the Supplementary Materials.

Measures and Procedure

Data were collected via Qualtrics in mid-March 2021 and mid-June 2021. Participants gave their informed consent. They were then presented with the General Attitudes towards Artificial Intelligence Scale (GAAIS), which was shown item by item, in a single random order with one attention check item to check whether the questions had been read. Items can be seen in Table 1 and in the Appendix, where the scale instructions can also be found.

After the GAAIS, participants completed the Technological Readiness Index (TRI), with the same item order as in Lam et al. (2008, Table 2 therein), prefixed by the instruction “*The next (and final) scale is about technology in general*”. For both instruments, response options were: Strongly Disagree, Somewhat Disagree, Neutral, Somewhat Agree, Strongly Agree, Prefer not to answer (the latter for ethical reasons), and participants were asked to indicate one answer per question. Two further attention checks were embedded in the TRI. Participants were paid a small financial reward in line with Prolific tariffs.

Results

Missing data treatment and scoring

Missing data were rare (0.33% for the GAAIS, 0.13% for the TRI). Missing data points were replaced with the grand mean for each scale. The Positive GAAIS, and the Innovativeness and Optimism subscales of the TRI were scored 1 to 5, with 1 = Strongly disagree, through 3 = Neutral, to 5 = Strongly agree. The items on the Negative GAAIS, and the Discomfort and Insecurity subscales of the TRI, were reverse-scored 1 to 5 (1 = Strongly agree; 5 = Strongly disagree). Thus, higher scores on each subscale represent more positive attitudes.

Confirmatory Factor Analysis of the GAAIS

To confirm the two-factor structure of the GAAIS, which had been identified during its initial validation, a Confirmatory Factor Analysis was run on JASP (JASP Team, 2020), in turn running lavaan (Rosseel, 2012). The estimator was DWLS (diagonally weighted least squares) because the item-level data were ordinal (Li, 2016). For this reason, multivariate normality was irrelevant and therefore not assessed. The Kaiser-Meyer-Olkin test yielded a Measure of Sampling Adequacy of 0.9, which is classed as excellent, and Bartlett's test of sphericity showed a lack of an identity correlation matrix, $X^2 = 2408.12$, $df = 190$, $p < .001$.

Based on Schepman and Rodway (2020), Factor 1 in the confirmatory model was the Positive GAAIS (12 items), and Factor 2 was the Negative GAAIS (8 items), and factors were allowed to correlate. The Chi-Square for the fit of this two-factor model was significant, $X^2 = 223.08$, $df = 169$, $p = .003$, with significance suggesting an imperfect fit. However, it is known that with larger sample sizes, Chi-Square can be oversensitive to small deviations, thus we interpret this finding against other major fit indices. The X^2 / df ratio was 1.32, with < 2 showing a good fit. Further standard fit indices also showed a good fit, CFI = .987; TLI = .986; SMRS = 0.065; RMSEA = 0.032, 90% CI [0.019, 0.044], $p = .997$. The factor covariance was 0.492, 95% CI [0.455, 0.528], $SE = 0.019$, $z = 26.215$, $p < .001$ and the correlation between the two factors was $r = 0.397$, $p < .001$.

The ECVI (Expected cross validation index) for this two-factor model was 1.007. We compared this to an alternative model with only one factor comprising all items, which showed an ECVI of 2.308, which was much higher (with lower values indicating a better fit). Thus, a two-factor model was supported.

All these patterns confirmed the factor structure identified at initial validation (Schepman & Rodway, 2020). Factor loadings are shown in Table 1. Negative item 6 (“I think artificially intelligent systems make many errors.”) had a relatively modest loading in this sample (.301 standardized, which is only just above the .3 minimum loading recommended in Field, 2013). However, its tentative removal had minimal impact on the overall model fit. In addition, the item had importance in the construct, because, as discussed in the Introduction, it reflected reliance-based trust in AI, which is an important aspect of attitudes towards AI, contributing to the breadth and comprehensiveness of the construct. It was also felt that the item may be of use in item-by-item analyses that future scale users may want to conduct, which may be of use in relation to vignettes that describe AI fiascos, or in relation to international or longitudinal research. Furthermore, the item had a somewhat higher factor loading at initial validation (0.47), suggestion some sample-specific variation. For all those reasons, we decided to retain it, rather than make a post-hoc modification to the scale based on this specific data set.

--- Insert Table 1 about here ---

Internal Consistency GAAIS and Technology Readiness Index, scale means and SDs

We produced Cronbach alpha values for all subscales to check their internal consistency. For the Positive GAAIS $\alpha = 0.88$ (12 items), and for the Negative GAAIS $\alpha = 0.82$ (8 items), both reflecting good internal consistency. The Innovativeness subscale of the TRI showed $\alpha = 0.85$, Optimism $\alpha = 0.67$, Discomfort $\alpha = 0.74$, Insecurity $\alpha = .074$. All TRI subscales showed acceptable to good internal consistency, particularly given the small numbers of items on each subscale (6 for Insecurity, 4 for the rest).

The Positive GAAIS had a scale mean of 3.60 ($SD = 0.68$) and the Negative GAAIS a slightly lower mean, $M = 3.10$, $SD = 0.76$. These means were a good match to those observed in Schepman and Rodway (2020), where the positive GAAIS had also shown a mean of 3.60, negative GAAIS 2.93, indicating good anchoring of the GAAIS to prior data. The TRI Innovativeness subscale had a mean of 3.47, $SD = 1.04$, Optimism $M = 4.11$, $SD = 0.66$, Discomfort $M = 3.03$, $SD = 0.87$, Insecurity $M = 3.22$, $SD = 0.77$. Here, again, the means were close to the Schepman & Rodway (2020) means of 3.66, 4.07, 3.02, and 3.12, respectively, providing further evidence of good anchoring to previously observed means in this new sample.

GAAIS and TRI: Correlations and Regressions to check for replication of convergent and discriminant validity

To check for convergent and discriminant validity between the subscales of the GAAIS and the subscales of the Technology Readiness Index, two separate sets of association measures were produced.

The first set consisted of the bivariate correlations between the four subscales of the TRI and the two subscales of the GAAIS (see Table 2). The Positive and Negative GAAIS correlated significantly with all subscales of the TRI, but the Positive GAAIS more strongly with the positive TRI subscales (Innovativeness and Optimism) and the Negative GAAIS more strongly with the negative TRI subscales (Discomfort and Insecurity), indicating good specificity. The correlations were not so strong as to suggest that the TRI and the GAAIS were measuring identical constructs.

The second set arose from Multiple linear regression analyses (jamovi project, 2021). We examined the significance of the TRI subscales as predictors when combined in the same

model, for enhanced inferential statistical testing rigor. The four subscales of the Technology Readiness Index were entered simultaneously as independent variables (predictors) into two multiple linear regression models, once with the Positive GAAIS as the dependent variable (criterion / outcome), and once with the Negative GAAIS as the dependent variable. The significance of each coefficient was of focal interest, but we also report the overall models, as these may be of interest to readers. First, assumptions were checked. For the Positive GAAIS the Shapiro-Wilk test for normality of the residuals, $SW = .971$, $p < .001$, showed non-normality of residuals; Breusch-Pagan for homoskedasticity, $BP = 2.14$, $p = .71$, n.s.; All VIF < 1.5 for collinearity; Durbin-Watson = 1.79, $p = 0.072$, n.s., for autocorrelation; Cook's distance: $M = 0.00408$, $SD = 0.0103$, $Min = 2.21e-7$, $Max = 0.118$, all < 1 , for outliers. For the Negative GAAIS, Shapiro-Wilk for normality of the residuals, $SW = .984$, $p = .002$, also showing non-normality of residuals; Breusch-Pagan for homoskedasticity, $BP = 9.74$, $p = 0.045$ showing heteroskedasticity; All VIF < 1.5 for collinearity; Durbin-Watson = 2.05, $p = 0.66$, n.s., for autocorrelation; Cook's distance: $M = 0.00435$, $SD = 0.0120$, $Min = 1.15e-9$, $Max = 0.173$, all < 1 , for outliers). Although deviations from the normal distribution of the residuals were detected for both subscales, and heteroskedasticity for the Negative GAAIS, wild bootstrap versions of the analysis (SPSS 27, 2000 samples, residuals, bias corrected accelerated) yielded the same patterns of coefficient significance, which was our main focus in this analysis. Therefore, the Ordinary Least Squares versions of the multiple regression analyses are reported, to report a stable analysis outcome and avoid the reduction in reproducibility of the analyses that are inherent in the resampling techniques of bootstrapping.

The overall multiple regression model with the four TRI subscales predicted 38.9% of the variance for the Positive GAAIS, $F(4, 299) = 47.6$, $p < .001$, and 20.4% for the Negative GAAIS, $F(4, 299) = 19.1$, $p < .001$. As for the focally relevant coefficients (Table 2), the

Positive GAAIS was significantly predicted by three subscales of the Technology Readiness Index, namely Innovativeness, Optimism, and Discomfort, with the positive subscale Optimism showing the highest standardised Beta value. Optimism, and the two negative TRI subscales Discomfort and Insecurity were significant predictors of the Negative GAAIS, with the two negative TRI subscales showing the highest Beta values. Overall, there was significant prediction from the TRI to the GAAIS, showing convergent validity. There was some specificity such that the Positive GAAIS was more strongly predicted by the positive subscales of the TRI, and the Negative GAAIS by the negative subscales of the TRI. There was also stronger prediction of the Positive GAAIS by the Optimism subscale of the TRI, which reflects societally-based attitudes towards technology, than by the Innovativeness subscale, which reflects more individually-based attitudes. This was also observed in Schepman and Rodway (2020). The associations between the TRI and the GAAIS were not so strong that it would be reasonable to assume that the GAAIS and the TRI measured identical constructs, thus providing evidence of discriminant validity, building on the pattern found at initial validation.

---- insert Table 2 about here ---

Study 1 Discussion

The main aim of Study 1 was to confirm the factor structure of the GAAIS, and this aim was achieved. The Confirmatory Factor Analysis showed excellent fit indices, and the Positive and Negative GAAIS showed strong internal consistency. Convergent and discriminant validity against the Technology Readiness Index was also replicated and confirmed. This confirmation of the factor structure and the convergent and discriminant

validity further validated the GAAIS, and established that the GAAIS was suitable as a tool to measure attitudes towards artificial intelligence. Study 2 examined personality traits, corporate distrust, and general trust as predictors of attitudes towards Artificial Intelligence, allowing us to test the hypotheses set out in the Introduction using the validated GAAIS. Because we were drawing a further fresh sample to test these substantive hypotheses, we also took the opportunity to re-confirm the factor structure with a further CFA for enhanced confidence, and to test the impact of a minor change in response anchor phrasing.

Study 2: Prediction of GAAIS by Personality and Trust Measures

This study was focused on testing the hypotheses regarding prediction of the GAAIS by personality and trust measures that were set out in the Introduction.

Method

Participants

Data were collected in early July 2021. There were 300 UK-based participants, adults over 18, recruited via Prolific.co. They had not taken part in Study 1, nor in the study reported in Schepman and Rodway (2020). There were 147 female participants, 151 male, 1 other, and 1 prefer not to answer. Data from four further participants were discarded due to a

failure on two or more of four attention checks. Mean age was 32.97 years ($SD = 11.86$, range 18 – 69; six participants did not provide their age).

Education levels were .3% no formal education [coded as 0 for analysis purposes]; 7.3% GCSE or equivalent [1]; 31% A-level or equivalent [2]; 42% Bachelor's degree or equivalent [3]; 14.7% Master's degree or equivalent [4]; 3.7% Doctoral degree or equivalent [5]; 1% other [due to lack of further specification treated as missing, coded as 999].

Self-rated computer expertise was .3% “Hardly ever use the computer and do not feel very competent” [coded as 0]; 1.3% “Slightly below average computer user, infrequently using the computer” [1]; 41.3% “Average computer user, using the internet, standard applications etc.” [2]; 36.3% “User of specialist applications but not an IT specialist” [3]; 12% “Considerable IT expertise short of full professional qualifications” [4]; 8.3% “Professionally qualified computer scientist or IT specialist” [5], and .3% prefer not to answer [999, missing]. Occupations can be seen in the data file via the Supplementary Information.

Measures and Procedure

The GAAIS was changed from Study 1 in one respect: The anchor “somewhat disagree” was replaced with “disagree”, and “somewhat agree” was replaced with “agree”. This was because analysis of the (rare) attention check errors of Study 1 showed the potential for “strongly” and “somewhat” to be confused when read quickly. The updated anchors eliminated the strong visual similarity between the relevant anchors.

We selected the Big Five Inventory-2 Short Form (BFI-2-S; Soto & John, 2017) to measure personality traits. We selected this thirty-item scale because of its balance between brevity and accuracy. In this scale the Big Five personality traits were named as Extraversion, Agreeableness, Conscientiousness, Negative Emotionality, and Open-Mindedness.

In addition, we used the Corporate Distrust Scale (Adams et al., 2010). This scale had 13 items measuring the unitary construct of distrust towards corporations (example item: “Corporations are driven by greed”). The scale was the only instrument designed to capture distrust towards corporations in general, which was why it was an optimal measurement in the context of AI, where the end user may not be certain of which corporation(s) may have provided the AI that they may be exposed to. At its validation the Corporate Distrust Scale correlated positively with Organizational Cynicism (Pugh, et al., 2003), which measured cynicism towards specific organizations. It also correlated with a range of other related measures, though not so strongly as to make the Corporate Distrust Scale redundant.

We also chose the General Trust Scale (Yamagishi & Yamagishi, 1994) to measure trust in people in general. This scale has six items (example item: “Most people are trustworthy”), and forms a unitary construct that measures general trust in other people. The scale has been widely used, and its validation article (Yamagishi & Yamagishi, 1994) had over 2800 citations on Google Scholar. Many alternative methods of measuring trust exist, but these were deemed to be either too specific in their trust objects (e.g. romantic partners, specific state institutions), too long, or consisting of too many unnecessary factors for our purposes.

We included four attention checks across all scales. After giving their informed consent, participants completed the scales in the order listed above. They received a small financial reward upon completion.

Results

Missing data treatment and scoring

The GAAIS was scored as for Study 1, with higher scores on each subscale representing more positive attitudes towards Artificial Intelligence. Other scales were scored 1-5 in line with their published scoring methods, which included reverse-scoring for half the items on the Big Five Inventory-2-S. For the personality subscales and the distrust and trust scales, the higher the score, the stronger the presence of the named construct. Missing data were once again rare and replaced with grand means for the scale / subscale: 0.17% for the GAAIS, 0.06% for Extraversion, 0.11% for Agreeableness; 0.06% for Conscientiousness; 0.11% for Negative Emotionality, 0.6% for Open-Mindedness, 0.25% for Corporate Distrust. There were no missing data for General Trust.

Replication of Confirmatory Factor Analysis of the GAAIS

We took the opportunity to replicate the Confirmatory Factor Analysis of the GAAIS on this new data set for enhanced reliability and to ensure that the revised response options were valid. Assumption checks showed Kaiser-Meyer-Olkin Measure of Sampling Adequacy was .89, and Bartlett's test of Sphericity $X^2 = 2068.03$, $df = 190$, $p < .001$. The main CFA showed that the Chi-Square for the CFA model fit was not significant, thus showing a good fit, $X^2 = 174.67$, $df = 169$, $p = .347$, as did the other fit indices, CFI = .998, TLI = .998, SMRS = 0.059, RMSEA = 0.011, 90% CI [0, 0.029], $p = 1$. The factor covariance was 0.57, 95% CI [0.529, 0.611], $SE = 0.021$, $z = 27.351$, $p < .001$ and the correlation between the two factors was $r = 0.48$, $p < .001$. The ECVI (Expected cross validation index) for this two-factor model was 0.862, and that for an alternative one-factor model comprising all items was

1.700, which was again much higher, with lower values indicating a better fit. These data once more confirmed the two factors, namely positive and negative, identified in Schepman & Rodway (2020) and in Study 1. The fit indices were all even stronger in Study 2. Factor Loadings are shown in Table 3. Note that the standardized loading for item 6 was 0.367, which was higher than in Study 1.

---- insert Table 3 about here ----

Internal Consistency GAAIS and Technology Readiness Index and scale means and SDs

Cronbach alpha values were calculated to check for internal consistency. For the Positive GAAIS, $\alpha = .85$ (12 items), for the Negative GAAIS $\alpha = .82$ (8 items). For the Big Five personality traits from the Big Five Inventory-2-S (6 items each): extraversion $\alpha = .74$, agreeableness $\alpha = .72$, conscientiousness $\alpha = .74$, negative emotionality $\alpha = .85$, open-mindedness $\alpha = .72$, each acceptable. There were high Cronbach alphas for or the trust measures, corporate distrust, $\alpha = .94$, and general trust, $\alpha = .85$.

The Positive GAAIS had a scale mean of 3.61 ($SD = .60$) and the Negative GAAIS, $M = 3.14$, $SD = .71$. The means were similar to those obtained for Study 1, again showing good anchoring. Of the personality traits, agreeableness had a mean of 3.68, $SD = .69$, conscientiousness $M = 3.56$, $SD = .74$, extraversion $M = 2.99$, $SD = .81$, negative emotionality $M = 2.87$, $SD = .95$, and open-mindedness $M = 3.72$, $SD = .74$. For the trust measures, corporate distrust had a mean of 3.85, $SD = .73$, and general trust, $M = 3.64$, $SD = .68$.

Hierarchical Multiple Linear Regression for Positive GAAIS: Overall model

To test the hypotheses and to evaluate the impact of demographic variables, a three-block (demographics, personality traits, trust measures) hierarchical multiple regression analysis was conducted with the Positive GAAIS as the outcome variable. Assumption checks were met: Shapiro-Wilk for normality of the residuals, $SW = .995$, $p = .497$, paired with a visual inspection of the Q-Q plot; Breusch-Pagan for homoskedasticity, $BP = 13.8$, $p = .244$; All $VIF < 1.5$ for collinearity; Durbin-Watson = 2.396 for autocorrelation; Cook's distance: $M = 0.00377$, $SD = 0.00602$, $Min = 1.36e-7$, $Max = 0.0366$, all < 1 , for outliers.

In the first block (Model 1), the demographic predictors were entered namely Gender (0 = Male, 1 = Female), Age in years, Education (categories scored 0 – 5 ordinally, with a higher number reflecting a higher education level), Computer expertise (0 – 5 ordinally, with a higher number reflecting greater computer expertise). These variables accounted for a significant proportion of variance, $R^2 = .143$, $F(4, 286) = 11.96$, $p < .001$. In second block, the Big Five personality subscales were added (Model 2), providing significantly better prediction, $R^2_{change} = 0.45$, $F_{change}(5, 281) = 3.13$, $p < .001$. In the third and final block (Model 3) the prediction was improved significantly by adding corporate distrust and general trust, $R^2_{change} = 0.52$, $F_{change}(2, 279) = 9.59$, $p < .001$. Model 3 accounted for 24.1% of the variance, which was significant, $F(11, 297) = 8.09$, $p < .001$. The significance patterns and estimates for the coefficients for model 3 are shown in Table 4. The coefficients of Model 3 were used to test our hypotheses.

---- Insert Table 4 about here ----

Coefficients: Predicting the Positive GAAIS from Personality, Corporate Distrust, and General Trust

Three demographic coefficients were significant. Greater age predicted more negative attitudes on the positive GAAIS, as did female gender, while greater computer expertise led to more positive attitudes.

In relation to $H_{\text{extraversion-pos}}$ and $H_{\text{extraversion-neg}}$, extraversion showed a significant negative association with the Positive GAAIS, which meant introverted people showed more positive attitudes towards positive aspects of Artificial Intelligence than extraverted people, providing support for $H_{\text{extraversion-neg}}$. None of the other Big Five personality traits predicted the Positive GAAIS in the context of the other predictor variables.

As for the trust measures, $H_{\text{corporate distrust}}$ was supported by significant negative prediction of the Positive GAAIS by corporate distrust. This indicated that people who showed high levels of distrust of corporations had more negative attitudes towards the positive aspects of AI. $H_{\text{general trust}}$ was supported by significant positive prediction of the Positive GAAIS by general trust. People with greater trust towards other people showed more positive attitudes towards the positive aspects of Artificial Intelligence. The hypotheses are more fully interpreted in the Discussion.

Hierarchical Multiple Linear Regression for the Negative GAAIS: Overall model

The same three-block hierarchical regression was conducted to predict the Negative GAAIS, which reflected attitudes towards the negative aspects of Artificial Intelligence. The reader is reminded that these were reverse-scored, so the higher the scores on the Negative GAAIS, the more positive and forgiving the attitudes towards these negative aspects.

In this analysis, the Breusch-Pagan assumption check was significant, $BP = 21.1, p = 0.033$, suggesting some heteroskedasticity. The other assumption checks showed good compliance, $SW = .998, p = .97$, for normality of the residuals, alongside an inspection of the Q-Q plot; all $VIF < 1.5$ for collinearity, Durbin-Watson = 2.30 for autocorrelation, Cook's distance $M = 0.00391, SD = 0.00664, Min = 4.80e-8, Max = 0.0426$, all < 1 , for outliers.

To examine the impact of the Breusch-Pagan significance on the model, a wild bootstrap version of the analysis was conducted (SPSS 27, 2000 samples, residuals, bias corrected accelerated) and its output was compared to the output of the standard Ordinary Least Squares (OLS) analysis. There were no differences in the patterns of coefficient significance between these two analyses. We therefore report the OLS analysis, in part to report the outcomes for the positive and negative GAAIS on the same basis, and in part so that to avoid a reduction in analysis reproducibility that is inherent in bootstrapping.

The first block (Model 1), demographics, did not predict a significant proportion of variance, $R^2 = .025, F(4, 286) = 1.86, p = .118$. Adding the personality variables in the second block (Model 2) significantly improved the model's prediction, $R_{change} = .099, F_{change}(5, 281) = 6.37, p < .001$. Adding the third block, which consisted of the two trust measures (Model 3), improved the model's prediction significantly again, $R_{change} = .099, F_{change}(2, 279) = 17.699, p < .001$. Model 3 predicted 22.3% of the variance, which was a significant percentage, $F(11, 279) = 7.29, p < .001$. The significance patterns and estimates for the coefficients in Model 3 are displayed in Table 5.

--- insert Table 5 about here ---

Coefficients: Associations between the Negative GAAIS and Personality, Corporate Distrust, and General Trust

In Model 3, the demographic coefficient of age was again significant, with higher age leading to more negative attitudes towards the negative aspects of Artificial Intelligence. Higher levels of computer expertise were significantly associated with more forgiving attitudes on the Negative GAAIS.

Of the personality variables, extraversion was a strong negative predictor, with introverts having more forgiving attitudes towards the negative aspects of AI, an association that had also been seen in the Positive GAAIS, and further supporting $H_{\text{extraversion-neg}}$. The Negative GAAIS was significantly positively predicted by two further personality variables, namely agreeableness (supporting $H_{\text{agreeableness}}$), and conscientiousness (supporting $H_{\text{conscientiousness}}$), with both showing positive associations. This means that people who had high levels of agreeableness and people who had high levels of conscientiousness showed more positive attitudes towards the negative aspects of AI than people who measured lower on those traits.

Of the trust measures, only corporate distrust was a significant negative coefficient in the prediction of the Negative GAAIS, supporting $H_{\text{corporate distrust}}$, meaning that those with a high level of distrust in corporations expressed more negative attitudes regarding the negative aspects of AI. The pattern of hypotheses and their inferential statistical support is summarized in Table 6.

--- Insert Table 6 About here ---

Discussion

Confirmatory validation of the GAAIS

The two studies with large samples both further validated the General Attitudes towards Artificial Intelligence Scale via Confirmatory Factor Analysis. The two-factor structure (Positive and Negative GAAIS) was confirmed each time, with good model fits. Both subscales showed convergent and discriminant validity against the expected subscales of the Technology Readiness Index (Lam et al., 2018) in Study 1. Study 2 showed that alternative anchors were also suitable. Confirmation of the factor structure, and establishing convergent and discriminant validity were important findings, because they confirmed that the scale was suitable as a measurement tool to obtain substantive findings, which was the focal aim of Study 2.

Personality and attitudes towards AI

The strongest association between specific Big Five personality traits and the GAAIS was the negative association between extraversion and both subscales of the GAAIS, supporting the hypothesis ($H_{\text{extraversion-neg}}$) that introverts have more positive attitudes towards AI. This hypothesis was generated via the theoretical notion and empirical evidence for algorithm appreciation (Logg et al., 2019). This term refers to a preference for algorithms over humans in informing decisions. In our data introverts had a stronger tendency to have positive attitudes towards the positive aspects of AI and were more forgiving towards the negative aspects of AI. This could be related to their lower liking for social interaction. Perhaps introverts are more enthusiastic about AI because it is linked to the notion that it can provide functions without the need to involve people, explaining the association with the Positive GAAIS. They might therefore also be more willing to tolerate the negative aspects of

AI, explaining the association with the Negative GAAIS. Prior empirical support for a negative association between technology and extraversion was found by Barrett et al. (2015), who probed learning technology, which could also reduce the need for social interaction. This also chimes with recent findings by Rheu et al. (2021), who found that AI-driven social chatbots were trusted more by introverted people if they refrained from engaging in small talk.

The findings ran counter to some association patterns found with other technologies (e.g. Devaraj et al, 2008; Meng & Leung, 2021; Qu et al, 2020; Svendsen et al., 2013), which had led to the opposing hypothesis ($H_{\text{extraversion-pos}}$) regarding extraversion and attitudes towards AI, namely of a positive association. This hypothesis was not supported by our data. AI differs from the technologies that were examined in the other cited studies, because it may not provide opportunities for extraverts to gain “social attention”, defined as “a tendency to behave in ways that attract or hold social attention and also to enjoy those behaviors” (Ashton et al, 2002, p. 246). This is because AI is often embedded and hidden from the user, even where it might be used socially (e.g. Facebook or dating apps). Users may not associate these types of social technologies with AI because they might not be aware of the AI that is used to provide the functionality. More research on the negative association between extraversion and attitudes towards AI would allow for a deeper exploration of this interesting link. It would provide useful information on the acceptability of AI in various societal and business settings.

The Negative GAAIS was predicted by conscientiousness, but the Positive GAAIS was not. People who were more conscientious had more positive attitudes in relation to the negative aspects of AI. Thus, our hypothesis ($H_{\text{conscientiousness}}$) that both subscales would show a positive association with conscientiousness was only partly supported. Conscientiousness has been positively associated with technological innovations in other domains. For example,

Ardebili and Rickertsen (2020) found a positive association between conscientiousness and attitudes towards genetically modified (GM) salmon and GM salmon feed (see also Barnett et al, 2015; Devaraj et al, 2008; Shropshire et al, 2015). It is interesting to observe that conscientious people were more tolerant of this innovation, which is considered somewhat controversial by some people. It may be that a forgiving attitude to perceived drawbacks of technology is an aspect of conscientiousness that explains the specific link between this trait and the Negative GAAIS, but not the Positive GAAIS. Conscientiousness has also shown negative associations with technology acceptance. For example, Charness et al (2018) showed a negative association between conscientiousness and autonomous vehicle acceptance, though they interpreted this as a reluctance on the part of conscientious people to have autonomous vehicles replace their own driving achievements (see also Hong et al., 2020). Further research may lead to more fine-grained interpretations of this interesting observed link between conscientiousness and AI attitudes.

The Negative GAAIS showed a positive association with agreeableness, which only partly supported our hypothesis ($H_{\text{agreeableness}}$) that both subscales would show this positive association. This pattern differs somewhat from some prior findings, which showed links between positive aspects of technology and agreeableness (e.g. Devaraj et al, 2008; Kortum & Oswald, 2018; Shropshire et al, 2015), though the association was not found in Barnett et al (2015), who examined learning technology. Our data showed that the association between technology evaluations and agreeableness was not only impacted by the type of technology and the types of ratings, but also by the polarity (positive or negative) of the content of the evaluative questions. Important facets of agreeableness are social compliance and trust, and these facets may have led to the pattern observed. Trusting and socially compliant people may be less likely to object to the negative aspects of AI, but there may be no specific link

between agreeableness and the positive aspects of AI, because other personality traits may drive attitudes towards the more positive aspects of AI.

The hypothesis that the Positive GAAIS would be positively associated with open-mindedness ($H_{\text{open-mindedness}}$) did not receive statistical support. As discussed, this measure has been found to correlate with some measures related to technology evaluations in the past (Kortum & Oswald, 2018; Qu et al., 2020; Svendsen et al. 2013), but not always consistently (Barnett et al, 2015). It is possible that open-mindedness as measured by the BF-2-S places emphasis on creativity and imagination (4 items), with less weight placed on general intellectual curiosity (2 items), and may therefore not associate readily with the utility and excitement aspects of the Positive GAAIS. However, it is also possibly that even with a greater weight on general intellectual curiosity (i.e. a liking for deep thinking and abstract ideas) the significant association would not be forthcoming, because the GAAIS is oriented to practical everyday aspects of AI, and not the interesting philosophical aspects of AI (e.g. Cappelen & Dever, 2021) that might have associated more strongly with open-mindedness. Such an interpretation may also apply to previous instances where measures related to new technology did not associate with open-mindedness (e.g. Kortum & Oswald, 2018; Qu et al, 2020, Svendsen et al, 2013) because open-mindedness may not automatically extend to the entire technology domain, but instead may associate only with specific aspects of technology (see also Barnett et al, 2015; Rauschnabel et al., 2015).

There was no statistical support for the hypothesis that the Negative GAAIS would show a negative association with negative emotionality ($H_{\text{negative emotionality}}$). In other words, people who were more “neurotic” did not show more negative attitudes towards the negative aspects of AI. Instead, there was no significant association. In the Big Five Inventory-2-S, the constituent facets of negative emotionality include anxiety, depression, and emotional volatility. It is possible that these specific constituent facets do not determine whether people

feel that e.g. AI will be used to spy on people, or that AI might take over, because these negative emotions may be directed towards issues of greater personal concern (e.g. relationships, life issues etc.). Such general life concerns might not transfer to other domains, such as technology. This would explain the lack of association between the Negative GAAIS and negative emotionality. Further research would be interesting, because for some technologies this association does appear to emerge (Barnett et al, 2015; Devaraj et al, 2008; Qu et al, 2020; Svendsen et al, 2013). It would be interesting to tease apart what aspects of technology and what aspects of AI might be significantly associated with negative emotionality.

Trust and AI

The GAAIS showed strong negative associations with corporate distrust, supporting $H_{\text{corporate distrust}}$ fully. The stronger the corporate distrust, the more negative the attitudes towards AI. This applied equally to the Positive and Negative GAAIS. The general trust scale was significantly positively associated with the Positive GAAIS (the more trust in people, the more positive the attitudes towards the positive aspects of AI). However, this scale was not significantly associated with the Negative GAAIS, overall partly supporting $H_{\text{general trust}}$.

It is interesting that general trust appears to be associated with the positive aspects only. This may be because the positive aspects of AI may be attributed, at least partly, to the efforts of individual people, but the negative aspects may be blamed more specifically on corporations, with less “culpability” attributed to individuals. Those who harbored strong corporate distrust felt negative not only about the drawbacks of AI, as measured by the Negative GAAIS, but also felt less warmly towards the positive aspects of AI. Perhaps a strong sense of suspicion of corporations may give some people a negative view of AI

altogether. To combat this, balancing computer automation against human control may optimize perceived AI trustworthiness (Shneiderman, 2020).

Building on the Salient Value Similarity (SVS) Model (Earle & Cvetkovich, 1995; Siegrist, 2021), perceived salient value similarity has been identified as a potential moderator of trust in technology and AI contexts. Yokoi et al (2021), and Yokoi and Nakayachi (2021) found that if a user perceives shared values with an agent or organization offering AI or technology, e.g. in automated vehicles or in medical settings, the user is more likely to trust that agent or organization to manage risk on their behalf. Because AI can present many risks, studying this important moderator in future research on trust in AI and general attitudes towards AI is important.

In relation to trust in organizations, much has been written about AI ethics and corporate social responsibility, and it has been suggested that ethical frameworks are important in promoting this responsibility. However, in recent work Lauer (2021) suggested that AI ethics should emerge from more global organizational ethics in each company that uses AI, rather than from ethical frameworks. Studying links between ethics, trust, and AI acceptance is of great importance as further AI applications are released (Ikkatai et al., 2022). Our own data suggest that a genuine engagement with AI ethics at the corporate level could lead to greater trust in AI and greater acceptance of AI, to the mutual benefit of corporations and users.

Impact of demographic factors

Demographic factors were not our main focus and we did not generate specific hypotheses about these, but our observations were interesting and may be useful to some readers. They also largely echo other findings (e.g. Zhang & Dafoe's large US-based survey

from 2019), providing strong calibration of our data set to prior data, enhancing confidence in its validity.

One of our findings was that male respondents were more likely to express positive attitudes towards the positive elements of AI. A similar pattern was found in Zhang and Dafoe (2019). It is possible that this gendered pattern may reflect levels of engagement with or interest in AI, previously described as a gender gap (Fatemi, 2020).

Younger people were more likely to show positive attitudes on both subscales than older people. Once again, this was also observed in Zhang and Dafoe (2019). This may be reflective of a general association between age and technology acceptance as examined by e.g. Hauk et al., (2018), although Hauk et al. suggested some limitations in this association, finding that it applied more strongly to technology that was perceived to be less useful by older adults. More generally, it is important not to make overgeneralizations about associations between age and technology, because stereotype threats can inhibit the adoption of potentially useful technology by older adults (Mariano et al., 2021). The link between age and attitudes towards AI therefore needs further research.

Finally, general education levels had no impact on attitudes towards AI, unlike in Zhang and Dafoe (2019). However, greater computer expertise was associated with more positive attitudes towards AI on both scales, broadly echoing Zhang and Dafoe's (2019) findings, though with a different operationalization of expertise, with theirs being defined by the completion of computer science or engineering degrees. Given that some experts signaled negative predictions about the impact of AI on society (Anderson et al., 2018; Müller & Bostrom, 2016), it is interesting to see that expertise led to positive associations with both subscales of the GAAIS, with those possessing higher levels of computer expertise showing more favourable views than those with lower levels. It echoes other links observed between

scientific knowledge and acceptance of new technologies, e.g. in general biotechnology (Lusk & Rozan, 2005), nanotechnology, and animal cloning (Kim & Kim, 2015). However, there are also contradictory findings, e.g. from Connor and Siegrist (2010) who suggest experience rather than knowledge may be the driver of the association, and from Sturgis and Allum (2004), who suggest that attributions play an important role. Recent work suggests that the impact of specific AI training is of relevance in AI attitudes (e.g. Ehsan et al., 2021; Sit et al., 2020). However, relationships between knowledge and attitudes are complex and context-dependent. For that reason, further research would be useful.

Limitations and further work

Our work, inevitably, has limitations that would need to be addressed by future research. An important feature of the GAAIS is that it asks for attitudes towards AI in general. Different respondents may have had different connotations with the term AI, and they may have held different exemplars of AI applications in mind when they responded. To counter this, Schepman and Rodway (2020) had examined how well the GAAIS correlated with comfortableness with specific applications, namely applications involving big data / automatization vs. applications that involved human judgement. Comfortableness with these sets of specific applications showed strong correlations with the GAAIS, providing some reassurance with regard to this potential limitation. This should also be seen in the context of different cultures. In the UK context, mainstream media tend to report on opportunities and innovations afforded by AI, and investments in AI, with some reporting on ethical and legislative issues, e.g. privacy, and data protection. Fewer discussions on the use of AI in warfare reach prime-time or front-page headlines, and thus these issues may be less

prominent in the UK public's perceptions. Of course, such uses of AI are considered in publications for specialist audiences (e.g. Johnson, 2019), but the public are unlikely to read these. Research in other cultures (e.g. Japan: Ikkatai et al, 2022) has shown that the public is anxious about the use of AI in autonomous weaponry that can be used in combat. Ikkatai et al. also documented wider apprehensions about the risks and ethical sensitivities surrounding AI. It is possible that a complex interaction between geolocation and news reporting may contribute to cultural differences in public perceptions of AI, and these can be measured via the GAAIS using international samples. More research on this important topic is needed.

Our sample was balanced for gender, and spanned the age ranges, but future research with the GAAIS may be able to reach higher level of sample representativeness of the population in terms of age. Our mean age was around the mid-30s, while the median UK population age is slightly higher, at around 40.4 years (Office of National Statistics, 2020). Future research with representative samples is planned. By way of assessment of the data, our sample mean ages were 36.15 years, $SD = 10.35$, range 18 – 64, employees only (Schepman & Rodway, 2020), 35.7 years ($SD = 13.2$ years, range 18 – 76 (Study 1), and 32.97 years ($SD = 11.86$, range 18 – 69, showing mean ages that varied by a few years, yet with very stable results in terms of mean GAAIS scores. This may form an early indication that a slightly higher sample age may not impact strongly on the results.

Our current research examined personality traits and trust measures as individual predictors of attitudes towards AI, which was commensurate with the current stage of development of the GAAIS. In future, it would be beneficial to examine more dynamic links, because Elson et al., (2018) have shown that associations between personality traits (in their case extraversion) and measures of acceptance of technology (in their case confidence in agent recommendations) can change dynamically in response to situational factors. Examining dynamic changes in attitudes towards AI and potentially dynamic changes in

association between AI and other factors would form useful and interesting future research. Similarly, more fine-grained explorations of attitudes towards AI, corporate distrust, and trust would allow for further, deeper insights to build on the data reported here.

Conclusions

The GAAIS was shown to be a strong measurement instrument to capture general attitudes towards Artificial Intelligence, separating attitudes towards the positive aspects and negative aspects of AI, with the structure confirmed in two confirmatory factor analyses using data from two large samples. The GAAIS showed significant associations with trust. High levels of corporate distrust associated negatively with people's attitudes towards positive and negative aspects of Artificial Intelligence, although general trust towards people was more specifically associated with the positive aspects. Our data showed clearly that corporate trust is an important factor in AI acceptance, and it is important that companies using AI strive to do this in a socially responsible and ethical way. The GAAIS also showed clear associations with some Big Five personality traits. The most consistent effect was that introverts displayed more positive attitudes than extraverts on both subscales of the GAAIS. People who were conscientious showed more positive attitudes towards the negative aspects of AI, as did people who measured higher on agreeableness. However, negative emotionality and open-mindedness were not significant predictors of either subscale of the GAAIS. The finding that introverts are more positive towards AI is an important finding that is likely to have useful practical implications in the deployment and marketing of AI-based products. There is scope for additional research to examine the associations in more detail.

About the Authors

Astrid Schepman is a Senior Lecturer in Psychology at the University of Chester, UK, with a multidisciplinary educational background in Experimental Psychology and Linguistics. Her interests in technology include public perceptions of Artificial Intelligence, technology and emotion, and technology's interfaces with language.

Paul Rodway is a Senior Lecturer in Psychology at the University of Chester, UK, with an educational background in Experimental Psychology and Intelligent Systems. His technology interests include applications of Artificial Intelligence and their links to Individual Differences, and Cognitive Science.

References

- Adams, J. E., Highhouse, S., & Zickar, M. J. (2010). Understanding general distrust of corporations. *Corporate Reputation Review*, 13(1), 38-51.
<https://doi.org/10.1057/crr.2010.6>
- Anderson, J., Rainie, L., & Luchsinger, A. (2018). Artificial Intelligence and the Future of Humans. *Pew Research Center, December*. https://www.elon.edu/docs/e-web/imagining/surveys/2018_survey/AI_and_the_Future_of_Humans_12_10_18.pdf
- Araujo, T., Helberger, N., Kruijkemeier, S., & De Vreese, C. H. (2020). In AI we trust? Perceptions about automated decision-making by artificial intelligence. *AI & Society*, 35(3), 611-623. <https://doi.org/10.1007/s00146-019-00931-w>
- Ardebili, A. T., & Rickertsen, K. (2020). Personality traits, knowledge, and consumer acceptance of genetically modified plant and animal products. *Food Quality and Preference*, 80, 103825. <https://doi.org/10.1016/j.foodqual.2019.103825>
- Asan, O., Bayrak, A. E., & Choudhury, A. (2020). Artificial intelligence and human trust in healthcare: focus on clinicians. *Journal of medical Internet research*, 22(6), e15154. <https://doi.org/10.2196/15154>
- Ashton, M. C., Lee, K., & Paunonen, S. V. (2002). What Is the Central Feature of Extraversion? Social Attention Versus Reward Sensitivity. *Journal of Personality and Social Psychology*, 83(1), 245-252. <https://doi.org/10.1037/0022-3514.83.1.245>
- Baldauf, M., Fröhlich, P., & Endl, R. (2020). Trust Me, I'm a Doctor—User Perceptions of AI-Driven Apps for Mobile Health Diagnosis. In *19th International Conference on Mobile and Ubiquitous Multimedia*. New York, NY, USA: Association for

Computing Machinery, November 2020, pp. 167-178.

<https://doi.org/10.1145/3428361.3428362>

Barnett, T., Pearson, A. W., Pearson, R., & Kellermanns, F. W. (2015). Five-factor model personality traits as predictors of perceived and actual usage of technology. *European Journal of Information Systems*, 24(4), 374-390. <https://doi.org/10.1057/ejis.2014.10>

British Psychological Society (2014). *Code of Human Research Ethics (2nd edition)*.

Leicester, UK: British Psychological Society. <https://www.bps.org.uk/news-and-policy/bps-code-human-research-ethics-2nd-edition-2014>

Brownsword, R., & Harel, A. (2019). Law, liberty and technology: criminal justice in the context of smart machines. *International Journal of Law in Context*, 15(2), 107-125. <https://doi.org/10.1017/S1744552319000065>

Cappelen, H., & Dever, J. (2021). *Making AI Intelligible: Philosophical Foundations*. Oxford University Press.

Carrasco, M., Mills, S., Whybrew, A., & Jura, A. (2019). The Citizen's Perspective on the Use of AI in Government. BCG Digital Government Benchmarking. *Boston Consulting Group*. <https://www.bcg.com/publications/2019/citizen-perspective-use-artificial-intelligence-government-digital-benchmarking.aspx>

Cave, S., Coughlan, K., & Dihal, K. (2019). Scary robots': examining public responses to AI. In *Proceedings of the Second AAAI / ACM Annual Conference on AI, Ethics, and Society*. <https://doi.org/10.17863/CAM.35741>
<https://www.repository.cam.ac.uk/handle/1810/288453>

Charness, N., Yoon, J. S., Souders, D., Stothart, C., & Yehnert, C. (2018). Predictors of attitudes toward autonomous vehicles: The roles of age, gender, prior knowledge, and

personality. *Frontiers in psychology*, 9, 2589.

<https://doi.org/10.3389/fpsyg.2018.02589>

Chen, Y. N. K., & Wen, C. H. R. (2021). Impacts of Attitudes Toward Government and Corporations on Public Trust in Artificial Intelligence. *Communication Studies*, 72(1), 115-131. <https://doi.org/10.1080/10510974.2020.1807380>

Connor, M., & Siegrist, M. (2010). Factors influencing people's acceptance of gene technology: The role of knowledge, health expectations, naturalness, and social trust. *Science communication*, 32(4), 514-538.

<https://doi.org/10.1177%2F1075547009358919>

Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and the NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.

Davis, F.D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13, 319-340. <https://doi.org/10.2307/249008>

DeCamp, M., & Tilburt, J. C. (2019). Why we cannot trust artificial intelligence in medicine. *The Lancet Digital Health*, 1(8), e390. [https://doi.org/10.1016/S2589-7500\(19\)30197-9](https://doi.org/10.1016/S2589-7500(19)30197-9)

Deng, C., Ji, X., Rainey, C., Zhang, J., & Lu, W. (2020). Integrating Machine Learning with Human Knowledge. *Iscience*, 101656. <https://doi.org/10.1016/j.isci.2020.101656>

Devaraj, S., Easley, R. F., & Crant, J. M. (2008). Research note—how does personality matter? Relating the five-factor model to technology acceptance and use. *Information systems research*, 19(1), 93-105. <https://doi.org/10.1287/isre.1070.0153>

Earle, T. C., & Cvetkovich, G. T. (1995). *Social trust: Toward a cosmopolitan society*. Praeger Press.

Edelman (2019). Artificial Intelligence (AI) Survey.

https://www.edelman.com/sites/g/files/aatuss191/files/2019-03/2019_Edelman_AI_Survey_Whitepaper.pdf

Ehsan, U., Passi, S., Liao, Q. V., Chan, L., Lee, I., Muller, M., & Riedl, M. O. (2021). The who in explainable ai: How ai background shapes perceptions of ai explanations. *arXiv preprint arXiv:2107.13509*.

Elson, J. S., Derrick, D., & Ligon, G. (2018). Examining trust and reliance in collaborations between humans and automated agents. In *Proceedings of the 51st Hawaii International Conference on System Sciences (HICSS-51)*, 430 – 439. January, 2018.

URI: <http://hdl.handle.net/10125/49943> https://aisel.aisnet.org/hicss-51/cl/processes_and_technologies_for_team/6/ ISBN: 978-0-9981331-1-9

EU (2020). Transparent, Reliable and Unbiased Smart Tool for AI. Horizon Research Programme, 2020-2024. <https://cordis.europa.eu/project/id/952060>

Fatemi, F. (2020). Bridging the gender gap in AI. *Forbes*.

<https://www.forbes.com/sites/falonfatemi/2020/02/17/bridging-the-gender-gap-in-ai/>

Field, A. (2013). *Discovering Statistics using SPSS* (4th Edition). Sage.

Hauk, N., Hüffmeier, J., & Krumm, S. (2018). Ready to be a silver surfer? A meta-analysis on the relationship between chronological age and technology acceptance. *Computers in Human Behavior*, 84, 304-319. <https://doi.org/10.1016/j.chb.2018.01.020>

Hong, J. W., Wang, Y., & Lanz, P. (2020). Why is artificial intelligence blamed more? Analysis of faulting artificial intelligence for self-driving car accidents in

- experimental settings. *International Journal of Human–Computer Interaction*, 36(18), 1768-1774. <https://doi.org/10.1080/10447318.2020.1785693>
- Ikkatai, Y., Hartwig, T., Takanashi, N. & Yokoyama, H. H. (2022). Octagon Measurement: Public Attitudes toward AI Ethics, *International Journal of Human–Computer Interaction*. <https://doi.org/10.1080/10447318.2021.2009669>
- jamovi project (2021). jamovi (Version 1.6) [Computer Software]. Retrieved from <https://www.jamovi.org>
- JASP Team (2020). JASP (Version 0.13.1) [Computer software]. Retrieved from <https://jasp-stats.org/>
- Johnson, J. (2019). Artificial intelligence & future warfare: implications for international security. *Defense & Security Analysis*, 35:2, 147-169. <https://doi.org/10.1080/14751798.2019.1600800>
- Jones, M. L., Kaufman, E. & Edenberg, E. (2018). AI and the Ethics of Automating Consent. *IEEE Security & Privacy*, 16 (3), 64-72. <https://doi.org/10.1109/MSP.2018.2701155>
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2), 1-36. URL <http://www.jstatsoft.org/v48/i02/>
- Kaur, K., & Rampersad, G. (2018). Trust in driverless cars: Investigating key factors influencing the adoption of driverless cars. *Journal of Engineering and Technology Management*, 48, 87-96. <https://doi.org/10.1016/j.jengtecman.2018.04.006>
- Kim, S., & Kim, S. (2015). The role of value in the social acceptance of science-technology. *International Review of Public Administration*, 20(3), 305-322. <https://doi.org/10.1080/12294659.2015.1078081>

- Kortum, P., & Oswald, F. L. (2018). The impact of personality on the subjective assessment of usability. *International Journal of Human–Computer Interaction*, 34(2), 177-186. <https://doi.org/10.1080/10447318.2017.1336317>
- Li, C. H. (2016) Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, 48, 936–949. <https://doi.org/10.3758/s13428-015-0619-7>
- Lane, W., & Manner, C. (2011). The impact of personality traits on smartphone ownership and use. *International Journal of Business and Social Science*, 2(17), 22–28.
- Lam, S. Y., Chiang, J., & Parasuraman, A. (2008). The effects of the dimensions of technology readiness on technology acceptance: An empirical analysis. *Journal of interactive marketing*, 22(4), 19-39. <https://doi.org/10.1002/dir.20119>
- Lauer, D. (2021). You cannot have AI ethics without ethics. *AI and Ethics*, 1(1), 21-25. <https://doi.org/10.1007/s43681-020-00013-4>
- Lee, H. J., Cho, H. J. Xu, W. & Fairhurst, A. (2010). The Influence of Consumer Traits and Demographics on Intention to Use Retail Self-service Checkouts. *Marketing Intelligence and Planning*, 28: 46–58. <https://doi.org/10.1108/02634501011014606>
- Liu, P., & Liu, J. (2021). Selfish or Utilitarian Automated Vehicles? Deontological Evaluation and Public Acceptance. *International Journal of Human–Computer Interaction*, 1-12. <https://doi.org/10.1080/10447318.2021.1876357>
- Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, 151, 90-103. <https://doi.org/10.1016/j.obhdp.2018.12.005>

- Lusk, J. L., & Rozan, A. (2005). Consumer acceptance of biotechnology and the role of second generation technologies in the USA and Europe. *TRENDS in Biotechnology*, 23(8), 386-387. <https://doi.org/10.1016/j.tibtech.2005.05.012>
- Makridakis, S. (2017). The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms. *Futures*, 90, 46-60. <https://doi.org/10.1016/j.futures.2017.03.006>
- Mariano, J., Marques, S., Ramos, M. R., Gerardo, F., Cunha, C. L. D., Girenko, A., ... & de Vries, H. (2021). Too old for technology? Stereotype threat and technology use by older adults. *Behaviour & Information Technology*, 1-12.
<https://doi.org/10.1080/0144929X.2021.1882577>
- McCrae, R. R., & Costa, P. T., Jr. (2008). Empirical and theoretical status of the five-factor model of personality traits. In G. J. Boyle, G. Matthews, & D. H. Saklofske (Eds.), *The SAGE handbook of personality theory and assessment (Personality theories and models, Vol. 1*, pp. 273–294). Thousand Oaks, CA: Sage.
- McManus, I., Keeling, A. & Paice, E. (2004). Stress, burnout and doctors' attitudes to work are determined by personality and learning style: A twelve year longitudinal study of UK medical graduates. *BMC Medicine* 2, 29. <https://doi.org/10.1186/1741-7015-2-29>
- Meng, K. S., & Leung, L. (2021). Factors influencing TikTok engagement behaviors in China: An examination of gratifications sought, narcissism, and the Big Five personality traits. *Telecommunications Policy*, 45(7), 102172.
<https://doi.org/10.1016/j.telpol.2021.102172>
- Milfont, T. L., & Sibley, C. G. (2012). The big five personality traits and environmental engagement: Associations at the individual and societal level. *Journal of Environmental Psychology*, 32(2), 187-195.
<https://doi.org/10.1016/j.jenvp.2011.12.006>

- Muir, B. M., & Moray, N. (1996). Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation. *Ergonomics*, 39(3):429-60.
<https://doi.org/10.1080/00140139608964474>
- Müller, V. C. and Bostrom, N. (2016). Future progress in artificial intelligence: A survey of expert opinion, in Müller, V. C. (Ed.). *Fundamental Issues of Artificial Intelligence* pp. 553-571. Springer Synthese Library.
- Office of National Statistics (2020). National Statistics status for population estimates at date of last assessment 24 November 2020.
<https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/mid2020#age-structure-of-the-uk-population>
- Olhede, S.C. & Wolfe, P.J. (2018). The growing ubiquity of algorithms in society: Implications, impacts and innovations. *Philosophical Transactions of the Royal Society. A*, 376, p. 20170364. <https://doi.org/10.1098/rsta.2017.0364> .
- Parasuraman, A. (2000). Technology Readiness Index (TRI): A Multiple-Item Scale to Measure Readiness to Embrace New Technologies. *Journal of Service Research*, 2(4), 307–320. <https://doi.org/10.1177/109467050024001>
- Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, 70, 153-163. <https://doi.org/10.1016/j.jesp.2017.01.006>
- Pugh, S. D., Skarlicki, D. P., & Passell, B. S. (2003). After the fall: Layoff victims' trust and cynicism in re-employment. *Journal of Occupational and Organizational Psychology*, 76(2), 201–212. <https://doi.org/10.1348/096317903765913704>

- Rauschnabel, P. A., Brem, A., & Ivens, B. S. (2015). Who will buy smart glasses? Empirical results of two pre-market-entry studies on the role of personality in individual awareness and intended adoption of Google Glass wearables. *Computers in Human Behavior*, 49, 635-647. <https://doi.org/10.1016/j.chb.2015.03.003>
- 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>
- Royal Society Working Group. (2017). *Machine learning: the power and promise of computers that learn by example*. Technical report. <https://royalsociety.org/topics-policy/projects/machine-learning/>
- Ryan, M. (2020). In AI We Trust: Ethics, Artificial Intelligence, and Reliability. *Science and Engineering Ethics*, 26, 2749–2767. <https://doi.org/10.1007/s11948-020-00228-y>
- Siegrist, M. (2021). Trust and risk perception: A critical review of the literature. *Risk analysis*, 41(3), 480-490. <https://doi.org/10.1111/risa.13325>
- Sit, C., Srinivasan, R., Amlani, A., Muthuswamy, K., Azam, A., Monzon, L., & Poon, D. S. (2020). Attitudes and perceptions of UK medical students towards artificial intelligence and radiology: a multicentre survey. *Insights into imaging*, 11(1), 1-6. <https://doi.org/10.1186/s13244-019-0830-7>
- Schepman, A. & Rodway, P. (2020). Initial validation of the general attitudes towards Artificial Intelligence Scale. *Computers in Human Behavior Reports*, 1, 100014, <https://doi.org/10.1016/j.chbr.2020.100014>

- Schmidt, P., Biessmann, F., & Teubner, T. (2020). Transparency and trust in artificial intelligence systems. *Journal of Decision Systems*, 29(4), 260-278.
<https://doi.org/10.1080/12460125.2020.1819094>
- Shneiderman, B. (2020). Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human–Computer Interaction*, 36(6), 495-504.
<https://doi.org/10.1080/10447318.2020.1741118>
- Stephanidis, C., Salvendy, G., Antona, M., Chen, J. Y., Dong, J., Duffy, V. G., ... & Zhou, J. (2019). Seven HCI grand challenges. *International Journal of Human–Computer Interaction*, 35(14), 1229-1269. Seven HCI grand challenges. *International Journal of Human–Computer Interaction*, 35(14), 1229–1269.
<https://doi.org/10.1080/10447318.2019.1619259>.
- Sturgis, P., & Allum, N. (2004). Science in society: re-evaluating the deficit model of public attitudes. *Public understanding of science*, 13(1), 55-74.
<https://doi.org/10.1177%2F0963662504042690>
- Siau, K., & Wang, W. (2018). Building trust in artificial intelligence, machine learning, and robotics. *Cutter business technology journal*, 31(2), 47-53.
- Smith, H. (2021). Clinical AI: opacity, accountability, responsibility and liability. *AI & Society*, 36, 535–545. <https://doi.org/10.1007/s00146-020-01019-6>
- Soto, C. J., & John, O. P. (2017). Short and extra-short forms of the Big Five Inventory–2: The BFI-2-S and BFI-2-XS. *Journal of Research in Personality*, 68, 69-81.
<https://doi.org/10.1016/j.jrp.2017.02.004>

- Soto, C. J. (2019). How replicable are links between personality traits and consequential life outcomes? The life outcomes of personality replication project. *Psychological Science*, 30(5), 711-727. <https://doi.org/10.1177/0956797619831612>
- Stanton, B. and Jensen, T. (2021), *Trust and Artificial Intelligence*, NIST Interagency / Internal Report (NISTIR 8330), National Institute of Standards and Technology, Gaithersburg, MD. https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=931087
- Svendsen, G. B., Johnsen, J. A. K., Almås-Sørensen, L., & Vittersø, J. (2013). Personality and technology acceptance: the influence of personality factors on the core constructs of the Technology Acceptance Model. *Behaviour & Information Technology*, 32(4), 323-334. <https://doi.org/10.1080/0144929X.2011.553740>
- Triberti, S., Durosini, I., Lin, J., La Torre, D., & Galán, M. R. (2021). On the "Human" in Human-Artificial Intelligence Interaction. *Frontiers in psychology*, 12, 808995. <https://doi.org/10.3389/fpsyg.2021.808995>
- Vesnic-Alujevic, L., Nascimento, S., & Polvora, A. (2020). Societal and ethical impacts of artificial intelligence: Critical notes on European policy frameworks. *Telecommunications Policy*, 44(6), 101961. <https://doi.org/10.1016/j.telpol.2020.101961>
- Walmsley, J. (2021). Artificial intelligence and the value of transparency. *AI & Society*, 36, 585–595. <https://doi.org/10.1007/s00146-020-01066-z>
- Yamagishi, T., & Yamagishi, M. (1994). Trust and commitment in the United States and Japan. *Motivation and Emotion*, 18, 129–166. <https://doi.org/10.1007/BF02249397>
- Yokoi, R., Eguchi, Y., Fujita, T., & Nakayachi, K. (2021). Artificial intelligence is trusted less than a doctor in medical treatment decisions: Influence of perceived care and

value similarity. *International Journal of Human–Computer Interaction*, 37(10), 981-990. <https://doi.org/10.1080/10447318.2020.1861763>

Yokoi, R., & Nakayachi, K. (2021). The Effect of Value Similarity on Trust in the Automation Systems: A Case of Transportation and Medical Care. *International Journal of Human–Computer Interaction*, 37 (13), 1269-1282. <https://doi.org/10.1080/10447318.2021.1876360>

Zhang, B. & Dafoe, A. (2019). *Artificial intelligence: American attitudes and trends*. Center for the Governance of AI, Future of Humanity Institute, University of Oxford, January 2019. <https://governanceai.github.io/US-Public-Opinion-Report-Jan-2019/executive-summary.html>

Zhou, T., & Lu, Y. (2011). The effects of personality traits on user acceptance of mobile commerce. *International Journal of Human–Computer Interaction*, 27(6), 545-561. <https://doi.org/10.1080/10447318.2011.555298>

Appendix: The General Attitudes towards Artificial Intelligence Scale (GAAIS)**Schepman & Rodway**

Instructions for participants: We are interested in your attitudes towards Artificial Intelligence. By Artificial Intelligence we mean devices that can perform tasks that would usually require human intelligence. Please note that these can be computers, robots or other hardware devices, possibly augmented with sensors or cameras, etc. Please complete the following scale, indicating your response to each item. There are no right or wrong answers. We are interested in your personal views.

Response Options at presentation:

Strongly disagree, Somewhat disagree, Neutral, Somewhat agree, Strongly agree

Alternative response options at presentation:

Strongly disagree, Disagree, Neutral, Agree, Strongly agree

List of items:

The item order has been re-randomised and an attention check has been included, so that the scale is ready for use.

Subscale (not for display)	Number (not for display)	Item
Positive	1	For routine transactions, I would rather interact with an artificially intelligent system than with a human.
Positive	2	Artificial Intelligence can provide new economic opportunities for this country.
Negative	3	Organisations use Artificial Intelligence unethically.
Positive	4	Artificially intelligent systems can help people feel happier.
Positive	5	I am impressed by what Artificial Intelligence can do.
Negative	6	I think artificially intelligent systems make many errors.
Positive	7	I am interested in using artificially intelligent systems in my daily life.
Negative	8	I find Artificial Intelligence sinister.
Negative	9	Artificial Intelligence might take control of people.
Negative	10	I think Artificial Intelligence is dangerous.
Positive	11	Artificial Intelligence can have positive impacts on people's wellbeing.
Positive	12	Artificial Intelligence is exciting.
Attention Check	A	I would be grateful if you could select Strongly agree.
Positive	13	An artificially intelligent agent would be better than an employee in many routine jobs.
Positive	14	There are many beneficial applications of Artificial Intelligence.
Negative	15	I shiver with discomfort when I think about future uses of Artificial Intelligence.
Positive	16	Artificially intelligent systems can perform better than humans.
Positive	17	Much of society will benefit from a future full of Artificial Intelligence
Positive	18	I would like to use Artificial Intelligence in my own job.
Negative	19	People like me will suffer if Artificial Intelligence is used more and more.
Negative	20	Artificial Intelligence is used to spy on people

Scoring: Check compliance with the Attention Check, then discount it from the scoring.

Score items marked “Positive” as Strongly disagree = 1; (Somewhat) disagree = 2; Neutral = 3; (Somewhat) agree = 4; and Strongly agree = 5. Score the items marked “Negative” in reverse so that Strongly disagree = 5; (Somewhat) disagree = 4; Neutral = 3; (Somewhat) agree = 2; and Strongly agree = 1. Then take the mean of the positive items to form an overall score for the positive subscale, and the mean of the negative items to form the negative subscale. The higher the score on each subscale, the more positive the attitude. We do not recommend calculating an overall scale mean.