Initial validation of the General Attitudes towards Artificial Intelligence Scale

Astrid Schepman and Paul Rodway

Department of Psychology, University of Chester.

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

Acknowledgement: We are grateful to Dr Claudine Clucas for her advice on Exploratory Factor Analysis and Unidimensionality Assessment.

Abstract

A new General Attitudes towards Artificial Intelligence Scale (GAAIS) was developed. The scale underwent initial statistical validation via Exploratory Factor Analysis, which identified positive and negative subscales. Both subscales captured emotions in line with their valence. In addition, the positive subscale reflected societal and personal utility, whereas the negative subscale reflected concerns. The scale showed good psychometric indices and convergent and discriminant validity against existing measures. To cross-validate general attitudes with attitudes towards specific instances of AI applications, summaries of tasks accomplished by specific applications of Artificial Intelligence were sourced from newspaper articles. These were rated for comfortableness and perceived capability. Comfortableness with specific applications was a strong predictor of general attitudes as measured by the GAAIS, but perceived capability was a weaker predictor. Participants viewed AI applications involving big data (e.g. astronomy, law, pharmacology) positively, but viewed applications for tasks involving human judgement, (e.g. medical treatment, psychological counselling) negatively. Applications with a strong ethical dimension led to stronger discomfort than their rated capabilities would predict. The survey data suggested that people held mixed views of AI. The initially validated two-factor GAAIS to measure General Attitudes towards Artificial Intelligence is included in the Appendix.

Keywords: Artificial Intelligence, Psychometrics, Questionnaire, Index, Attitudes, Perception

Highlights

- The General Attitudes towards Artificial Intelligence Scale was validated
- Attitudes towards AI differ from traditional technology acceptance
- Comfortableness and capability for specific AI applications were measured
- AI for big data was rated higher than AI for complex human judgements
- Attitudes towards AI were affected by ethical judgements

1. Introduction

1.1. Background

Use of Artificial Intelligence (AI) is growing at a fast pace and permeates many aspects of people's daily lives, both in personal and professional settings (Makridakis, 2017; Olhede & Wolfe, 2018). People's general attitudes towards AI are likely to play a large role in their acceptance of AI. An important aim of our study was to develop a tool by which general attitudes toward AI could be measured in practical and research contexts, and to explore the conceptual aspects of such a tool. This took the form of an initial conceptual and statistical validation of a new scale. A further aim was to document current general attitudes and attitudes towards specific exemplars of AI applications.

To support our aims, we inspected recent literature to look for major themes that could inform the creation of our scale items. We first discuss qualitative studies. Anderson, Rainie, and Luchsinger (2018), asked 979 experts for their views on the following question "As emerging algorithm-driven artificial intelligence (AI) continues to spread, will people be better off than they are today?". The respondents' collective views were mixed, identifying both benefits (e.g. enhanced effectiveness) and threats (e.g. data abuse, job losses, threats to human agency). Cave, Coughlan, and Dihan (2019) examined AI narratives produced by a representative sample of the UK population. They quantified the incidence of "Hopes" (e.g. AI making life easier) and "Fears" (e.g. AI taking over or replacing humans). They found a preponderance of negative views, in which narratives featuring dystopian expectations of AI's future impact prevailed. In contrast, Fast and Horvitz (2017) analysed news reports in the *New York Times* on AI over three decades, and noted increased reporting from 2009, with a general increase in optimism in the reporting, yet also with marked increases in concerns (e.g. loss of control, ethical issues, impact on work). Together, these works suggested contrasting positive and negative themes, which were held by experts, the general public, and the media alike.

Recent large-scale quantitative surveys reported similar mixed views and echoed the same broad themes. In a survey of UK attitudes towards machine learning (Royal Society Working Group, 2017), the public perceived opportunities, but also expressed concerns regarding harm, impersonal experiences, choice restriction, and replacement. Zhang and Dafoe's (2019) survey of US citizens' attitudes towards AI, examined applications in wide use (e.g. Siri, Google), and future applications likely to impact widely on society (e.g. use of AI in privacy protection, cyber attack prevention, etc.). Their findings provided a mixture of support and concerns regarding AI. Overall, more participants (42%) supported AI than opposed it (22%), yet caution was expressed by 82%, who felt, for example, that robots should be managed carefully. Carrasco, Mills, Whybrew, and Jura's (2019) BCG Digital Government Benchmarking survey obtained similar data, with people being more accepting of AI for some applications (e.g. traffic optimisation), than for others (e.g. parole board decisions). Interestingly, Carrasco et al suggested that AI may have been preferred to humans in countries where trust in governments may be low. Preferences for AI over humans has also been observed in different context, related to expertise, in a phenomenon named "algorithm appreciation" (Logg, Minson, & Moore, 2017). Issues regarding employee displacement, ethics, and non-transparent decision making were among the public's concerns. Edelman (2019) identified similar themes, alongside concerns about AI exacerbating wealth inequalities, loss of intelligent behaviour in humans, increase in social isolation, and the threat of abuses of power by malevolent actors (e.g. using deepfake material to spread misinformation). Overall, recent large surveys reported a range of positive and negative attitudes towards AI, echoing the key themes of the qualitative studies.

Other studies in the literature explored more specific aspects of AI perceptions in more depth. A selection is discussed here. One perceived negative aspect of AI is potential job displacement. Frey and Osborne (2017) generated computerisability scores for 702 occupations, with many of those being highly computerisable. Chui, Manyika, and Miremadi (2016) carried out an analysis with a similar aim but a different methodology, and also identified a range of jobs at risk of automation, as did White, Lacey and Ardanaz-Badia (2019). Naturally, this may cause negative emotions towards AI. However, Granulo, Fuchs, and Puntoni (2019) found that, although people had negative emotions if they imagined other people's jobs being replaced by robots, they would feel less negative it if their own jobs were replaced by robots when compared to their jobs being replaced by other people. Together, these works suggest that jobs with highly predictable tasks may indeed be automated, so people's concerns for their future employment might be accurate.

As noted, AI can also trigger ethical concerns, as illustrated from Fenech, Strukelj, and Buston (2018), who showed divided views on the use of AI in medical diagnosis in a representative UK sample (45% for, 34% against, 21% don't know). Similar divisions applied to comfortableness with personal medical information being used in

AI (40% comfortable vs. 49% uncomfortable, 11% don't know). A majority was against the use of AI in tasks usually performed by medical staff, such as answering medical questions, suggesting treatments (17% for, 63% against, 20% don't know). Vayena, Blasimme, and Cohen (2018) explored what could be done in response to a majority of the UK public feeling uncomfortable with the use of AI and machine learning in medical settings. They concluded that trust in these applications needed to be promoted by data protection, freedom from bias in decision making, appropriate regulation, and transparency (see Barnes, Elliott, Wright, Scharine, & Chen, 2019; Sheridan, 2019; Schaefer, Chen, Szalma, & Hancock, 2016 for recent discussions on trust in AI in other contexts). In all, these studies illustrated comfortableness, emotional reactions, perceived capability, ethical considerations, and trust as important themes. They also showed the mixed pattern of views that emerged from the more global survey studies and qualitative studies. Altogether, many important positive and negative views of AI were identified in prior studies, and these have informed the generation of items used in our scale.

1.2. The present study: A scale and allied measures

Our study's aim was to conduct initial exploratory work towards a measurement tool with which general attitudes towards AI could be gauged in different contexts. Although instruments have been developed that measure people's acceptance of technology (e.g. Davis, 1989; Parasuraman & Colby, 2015), most of these do not focus on AI, whose acceptance may be different in key dimensions. Technology Acceptance (Davis, 1989) is a construct that focuses primarily on the user's willingness to adopt technology through a consumer choice. However, frequently, consumer choice is not a factor in the application of AI, because large organisations and governments may decide to adopt AI without consulting with their end users, who therefore have no choice but to engage with it. For this reason, traditional technology acceptance measures might not be ideal to measure attitudes towards AI.

A more recently developed general technology scale is the Technology Readiness Index. It was revised several times, but we focus on a version by Lam, Chiang, and Parasuraman (2008). This Index contains some elements that make it better placed to capture key aspects of AI, but it also has some elements that may be less suited. Lam et al.'s (2008) Technology Readiness Index has four subscales; Innovativeness, exemplified by a sample item “You keep up with the latest technological developments in your areas of interest”, Optimism, e.g. “Technology gives people more control over their daily lives”, Discomfort, e.g. “Sometimes you think that technology systems are not designed for use by ordinary people”, and Insecurity, e.g. “You do not consider it safe to do any kind of financial business online”. These subscales provide an interesting mixture of measures that correspond mostly to the individual user experience (Innovativeness, Discomfort), and measures that primarily capture reactions to technology being used more widely in society (Optimism, Insecurity). We used the Technology Readiness Index to test for convergent and discriminant validity with our new scale, hypothesising that there would be stronger associations of our measures with the societally-based than the individually-based subscales of the Technology Readiness Index, because AI is outside the end user's own control.

Additionally, in the second part of our study, we measured participants' views towards specific applications of AI. An important aim of this part of the study was to cross-validate the general attitudes using an independent contemporary objective measure. During the formation of general attitudes, the generalisations that people arrive at may be biased by cognitive heuristics (Sloman & Lagnado, 2005). This can be caused by overgeneralisations being based on too few instances. It can also be caused by generalisations not having been informed by specific instances, but, for example, by general media coverage. Both causes can make generalisations inaccurate. Asking individuals to make judgements about specific exemplars can help overcome this. Moreover, providing specific exemplars of a general technology is likely to facilitate the person in expressing views of that technology. This is because it may be easier to think of the implications. In addition, their views may form in less abstract and more concrete ways. In this part of the data it was not our aim to produce a scale, but to discover latent factors in the data to create composite measures for cross-validation purposes. Our reasoning was that convergence between the general and specific AI measures would strengthen confidence in the general scale. The survey data are also of more general interest as a gauge of current attitudes towards AI and its specific applications. Another important aim behind the discovery of latent factor structures in specific applications was that it would allow for important conceptual insights about any groupings in participants' perceptions.

2. Method

2.1. Ethics

The study was approved by the Department of Psychology Ethics Committee at our institution and complied with the British Psychological Society's (2014) Code of Human Research Ethics (2nd edition).

2.2. Recruitment, participants and demographic information

2.2.1. Recruitment

Data were collected in May 2019 via Prolific (https://www.prolific.co), an online participant database based in the UK. Participants were payed £1.75 shortly after completion.

2.2.2. Participants

Data from 100 participants were collected, 50 male, 50 female, who were non-students, residing in the UK and aged over 18. Data from one male participant were removed because he did not answer any of the 11 attention checks correctly (see Section 2.3.5), suggesting that the remaining questions may not have been read properly. We focused on workers, because they were likely to be affected by AI in both their personal sphere and their employment setting (Frontier Economics, 2018, Makridakis, 2017; Olhede & Wolfe, 2018), and therefore formed a useful dual-purpose sample. One participant had indicated employment in the Prolific sample filtering fields, but reported being unemployed at the time of the survey, the rest were (self)-employed.

2.2.3. Age, education, computer expertise.

The retained sample had a mean age of 36.15 years (SD = 10.25, range 20 – 64).

Their education levels and self-rated computer expertise are documented in Table 1.

--- insert Table 1 about here ---

Table 1: Education levels and self-rated computer expertise of the sample

Education		Computer Expertise	
Level	Frequency	Level (d)	Frequency
No formal education	0	Hardly ever use the computer and do not feel very competent	0
GCSE or equivalent (a)	14	Slightly below average computer user, infrequently using the computer, using few applications	1
A-level or equivalent (b)	30	Average computer user, using the internet, standard applications etc.	43
Bachelor’s degree or equivalent	34	User of specialist applications but not an IT specialist	37
Master’s degree or equivalent	17	Considerable IT expertise short of full professional qualifications	11
Doctoral degree or equivalent	2	Professionally qualified computer scientist or IT specialist	10
Other (c)	3		

Table 1 Notes:
a) GCSE is a General High School qualification usually taken at age 16
b) A-Level is a more specialised High School qualification, pre-university entry, usually taken at age 18
c) Professional qualifications, some in addition to those listed above
d) Some people chose two options, namely one both “Considerable IT expertise short of full professional qualifications”, and “User of specialist applications but not an IT specialist”, and two chose both “User of specialist applications but not an IT specialist” and “Average computer user, using the internet, standard applications etc.”, included in both frequency categories, explaining sum of 102.

2.2.4. Occupations

We asked for occupations via an open text box, which yielded 82 different labels and three missing responses. A large majority of the occupations were in the service sector, in line with the wider UK economy, where around 80% of employment and Gross Domestic Product is the service sector (Duquemin, Rabaiotti, Tomlinson, & Stephens, 2019). We observed occupations from a wide socio-economic range (e.g. cleaner, caretaker, linen assistant, sales assistant, security vs. academic, director, general practitioner, lawyer, vet), suggesting that our sample included representation from all strata. There was substantial representation from IT-related occupations. Table 2 shows all occupations to allow readers to gain fuller insight into the range.

--- Insert Table 2 about here ---

Table 2: Occupations named by participants

Academic	Cyber security specialist	Lab assistant	Revenue accountant
Account manager	Data analyst	Lawyer	Sales
Actress	Data entry	Linen assistant	Sales advisor
Administration and finance officer	Design engineer	Marketing manager	Sales assistant (2)
Administrator (4)	Designer	Mechanical engineer	Security
Armed security	Director (2)	Mortgage broker	Senior project officer
Assistant manager	Education consultant	Nurse	Systems administrator
Assurance team lead	Engineer	Nurse specialist	Software engineer (2)
Bank manager	Event manager (2)	Office admin assistant	Teacher (3)
Behaviour officer	Executive	Office administrator	Technical support
Builder	Finance assistant	Office manager	Technical trainer
Business	Finance officer	Online retailer	Technician
Careers adviser	Food retail	Operator	Transport coordinator
Caretaker	General practitioner	PA	Transport manager
Civil servant	Graphic designer	Photographer	Vet
Cleaner (2)	Investment manager	Property management	Waitress
Clerk	IT (2)	Receptionist (3)	Warehouse clerk
Commercial assistant	IT analyst	Residential support worker	Warehouse supervisor
Compliance manager	IT supervisor	Restaurant manager	Web designer
Business consultant	IT technician (2)	Retail assistant	Writer (3)
Customer service			

Table 2 Note: Occupations in alphabetical order, with occupations named more than once showing the number of occurrences.

2.3. Measures

2.3.1. Overview

In this section we describe the design of three new measures. We also briefly outline one validated measure chosen from the literature.

2.3.2. General attitudes towards Artificial Intelligence

A variety of items reflecting manifestations of attitudes towards AI were generated, and subsequently evaluated by the authors for coverage, fit, clarity of expression, and suitability for a wide audience. We generated items that reflected the positive and negative themes identified from the literature (Section 1.1), creating 16 positive items (opportunities, benefits, positive emotions), and 16 negative items (concerns and negative emotions). It was important that the statements captured attitudes towards AI in general terms, abstracting away from specific applications, settings, or narrow time windows. Example items included “There are many beneficial applications of Artificial Intelligence” “Artificial Intelligence is exciting” (positive), “I think artificially intelligent systems make many errors” “I shiver with discomfort when I think about future uses of Artificial Intelligence” (negative). Trust was captured in e.g. “Artificial Intelligence is used to spy on people”, “I would entrust my life savings to an artificially intelligent investment system”. All items were phrased to be suitable for responses to a five-point Likert scale with the anchors strongly/somewhat (dis)agree and neutral.

2.3.3. Specific AI applications for comfortableness and capability ratings

To create a set of specific applications of AI for participants to rate, we gathered news stories that reported recent developments in artificial intelligence. The stories were sourced by searching for “Artificial Intelligence” on the websites of three quality UK newspapers (*The Guardian, The Independent, The Financial Times*) in late February 2019. Hits were classed as relevant if they described specific applications of AI. We used our judgement to exclude stories that overlapped with others, or that may be ethically problematic by being potentially distressing to participants. This process yielded 42 news stories, 14 from each newspaper. We produced brief one-line summaries of the tasks that the artificially intelligent systems were able to perform, and these formed items in the study. The items can be found in Appendix A, alongside URLs linking to the source newspaper articles.

2.3.4. Technology Readiness Index

We selected a validated scale to measure attitudes towards technology, namely the Technology Readiness Index, opting for a short version with 18 items (Lam, Chiang, & Parasuraman, 2008). This scale is psychometrically strong and well-used. It has four subscales (Innovativeness, Optimism, Discomfort, and Insecurity). The scale has been shown to predict user interactions with technology products, with its subscales having separate predictive power. Innovativeness and Discomfort are more closely related to individual user experiences, and Optimism and Insecurity more to the use of technology in society.

2.3.5. Attention checks

To assure the quality of the data, we used 11 attention checks embedded throughout all questionnaires. In some, a particular response was requested e.g. “We would be grateful if you could select somewhat comfortable”, with such items varying in their phrasing and requested responses. In the scales that used agreement responses we used factual questions by way of attention checks. Participants could agree or disagree with these (e.g. “You believe that London is a city”; “A chair is an animal”).

2.4. Procedure

Participants gave their informed consent. As part of the general consent, the following information was given: "*This study investigates people's perceptions of Artificial Intelligence (computing-based intelligent systems). We ask you to rate your views on artificially intelligent systems and technology more generally. At the end, you have the option of adding brief comments. There are no right or wrong answers. We are interested in your personal views.*" Other informed consent features were more general and complied with general British Psychological Society Ethical Guidelines.

Participants then completed each questionnaire in turn via JISC Online Surveys software. We used built-in data checks to ensure each question had exactly one answer, to minimise missing data. A "prefer not to answer" option was available. We told participants that there would be attention checks.

We issued separate instructions for each scale, and there were varying response options. For the General Attitudes towards Artificial Intelligence we stated: "*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.*" Response options were left-to-right "strongly disagree; somewhat disagree; neutral; somewhat agree; strongly agree”. Items were in the same random order for each participant.

For the specific applications, we first asked “*You will see a series of brief statements of tasks that artificially intelligent systems (AI) may be able perform. Please rate how comfortable you would feel with Artificial Intelligence performing each task*.” Response options were, left-to-right: “very uncomfortable; somewhat uncomfortable; neutral; somewhat comfortable; very comfortable”. After all the items were rated for comfortableness, we stated “*We will show you the same items again, but this time please rate how CAPABLE you perceive Artificial Intelligence to be compared to humans.*” Response options were, left-to-right: “AI much less capable than humans; somewhat less; equally capable; somewhat more; AI much more capable than humans”. Items were in the same random order for each participant, and the same order for comfortableness and capability.

Our final scale was the Technology Readiness Index, as presented in Lam et al. (2008, Table 2 therein) in the same order or presentation, with the brief instruction “*on the next screen, there are some questions about your technology use in general. Please complete the following scale, indicating your response to each item*”. Response options were, left-to-right, “strongly disagree; somewhat disagree; neutral; somewhat agree; strongly agree”.

After that, there was an optional open comments text box, allowing for brief comments up to 300 characters. Few respondents made use of this option, and comments largely echoed the themes of the main questionnaires, so there is no further report of these data. Finally, a debrief screen provided brief further information about the study, the general sources of the news stories for the application items, and sources of support in the unlikely event this was needed. The entire procedure including ethics processes and debriefing took participants just under 19 minutes on average.

3. Results

3.1. Data preparation and treatment of missing quantitative data

Because of the use of technical settings to minimise missing data, the only missing data were cases in which participants had chosen “prefer not to answer”. Use of this option was relatively rare, with overall 136 data points of 13266 or 1% missing. To ready the data for analysis, verbal labels constituting the answer provided were changed to numerical values 1 to 5, with leftmost options 1, rightmost options 5 in the first instance (see Section 2.4). Missing data points were replaced with the grand mean for the relevant block, rounded to the nearest integer, in all cases 3 (“neutral”). Rounding to the nearest integer was chosen in preference to exact values to avoid minor fractional discrepancies in means when some data were scored as unreversed in some analyses, and reversed in others. In practice, means were only a fraction removed from these rounded integers, and in light of the small proportion of missing data this rounding had minimal impact.

3.2. Overview of analyses

We present data from the General Attitudes towards AI questions first, followed by data from the specific applications of AI, for which comfortableness and perceived capability were measured. For each subset of the data, a series of analytic techniques were used. Fine-grained frequency data are presented, because these are likely of interest for those working in AI. They also calibrate our findings to those from other surveys. We then report Exploratory Factor Analysis and allied statistics. For the General Attitudes the Factor Analysis was used to validate the scale. For the ratings of specific applications, the aim was not to produce a scale, but to find factors to aid understanding, support dimension reduction, and produce composite measures to cross-validate the GAAIS. Full data are available via the Supplementary Materials.

3.3. General Attitudes towards Artificial Intelligence

3.3.1. Descriptive statistics: Frequencies

We report frequency categories of agreement visually, at this stage in unreversed form to aid interpretability. To ensure that the visualisations were interpretable, we combined (dis)agreement from the "strongly" and "somewhat" levels, retained the neutral category, and plotted the frequencies of categories in Figure 1 (positive statements) and Figure 2 (negative statements).

--- Insert Figures 1 and 2 about here ---

As can be seen, participants endorsed some positive statements with high frequency, e.g. that there would be many beneficial applications of AI, but participants were less ready to declare AI to be better than humans at complex decisions. In the negative items, many felt that AI might threaten job security, but few instinctively disliked AI or found it sinister.

Figure 1: Frequencies of responses to positive statements in the General Attitudes to Artificial Intelligence questionnaire

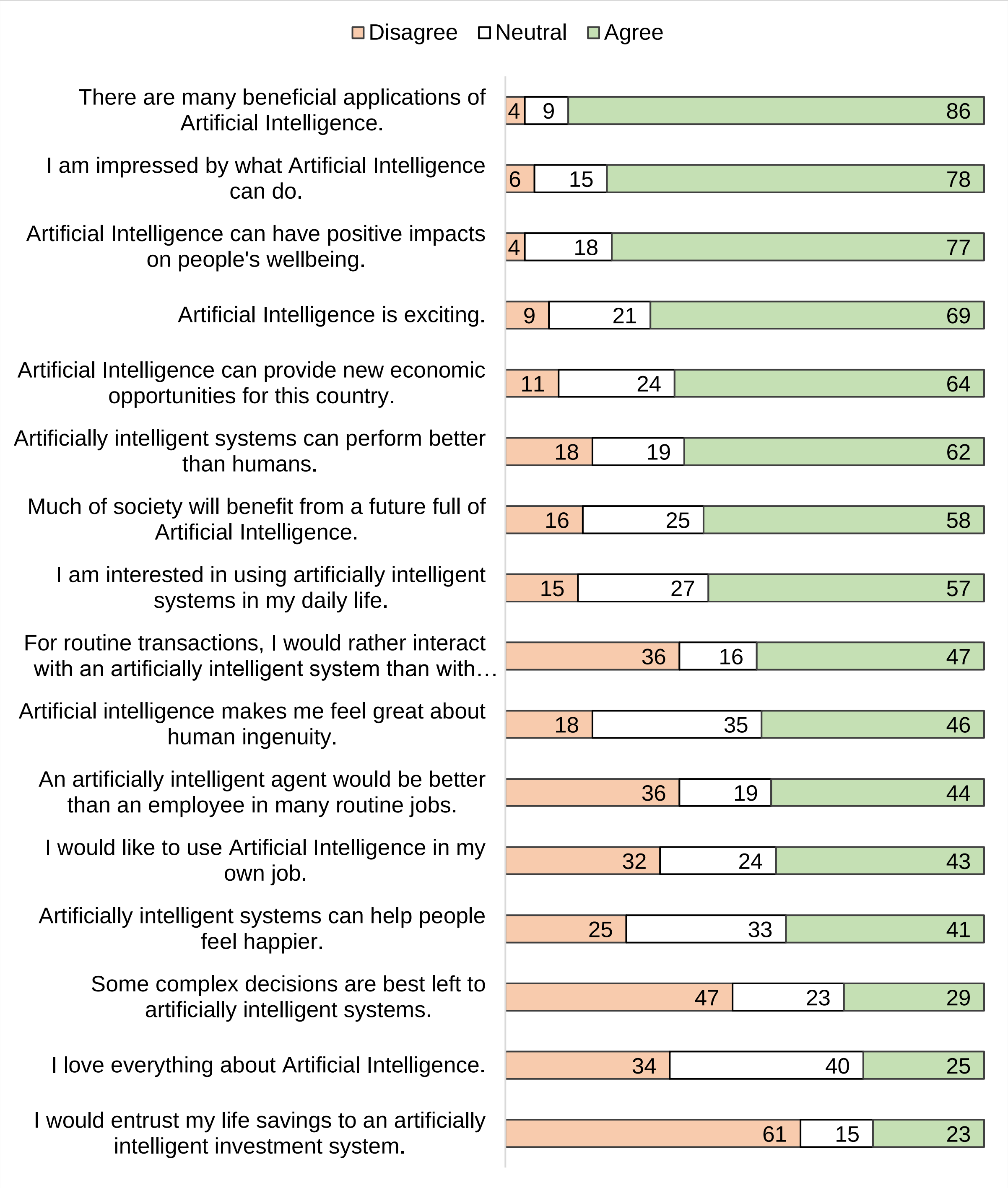


Figure 1 Note: Disagreement and agreement combine the “somewhat” and “strongly” categories of (dis)agreement. Disagreement is presented in orange at the left of the bars, neutral in white, centrally, and agreement in green as the rightmost part of the bars. N = 99, and bars contain raw frequencies. The last word in the truncated item starting “For routine transactions…” is “…humans”.

Figure 2: Frequencies of responses to negative statements in the General Attitudes to Artificial Intelligence questionnaire

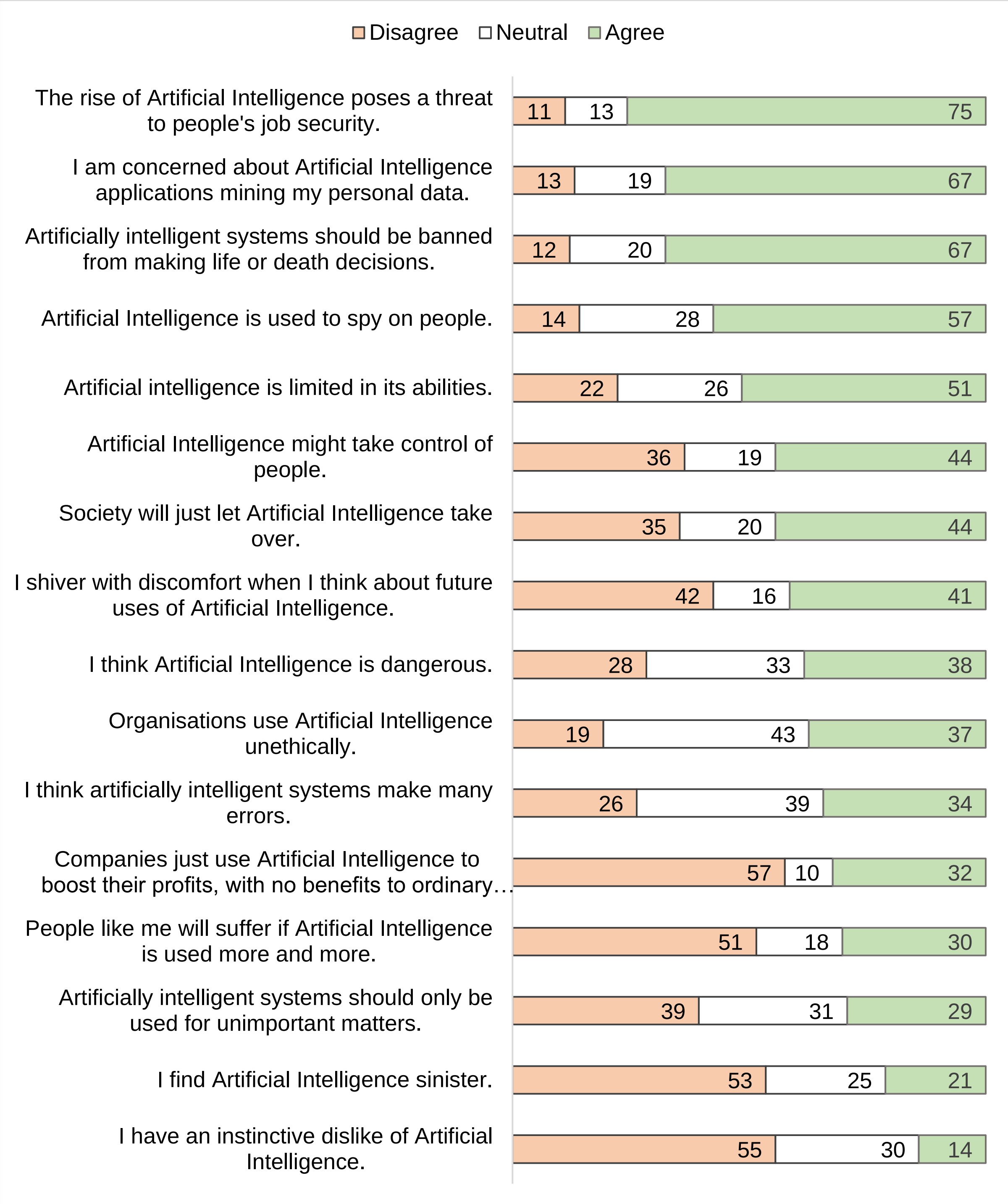


Figure 2 Note: Disagreement and agreement combine the "somewhat" and "strongly" categories of (dis)agreement. Disagreement is presented in orange at the left of the bars, neutral in white, centrally, and agreement in green as the rightmost part of the bars. N = 99, and bars contain raw frequencies. The last word in the truncated item starting "Companies just…" is "…people".

3.3.2. Exploratory Factor Analysis and internal consistency

We used Exploratory Factor Analysis to examine factors, and to test whether dimension reduction and the creation of composite subscales was supported. This process suggested two subscales along our a priori factors (positive and negative). We conducted internal consistency analyses for the two ensuing composite measures using Cronbach's alpha. Before the Exploratory Factor Analysis, we first reverse-scored the negative items, because all items needed the same polarity for this analysis. We then examined the item correlation matrix, and identified item pairs that were in multiple very low correlations with other items and had high associated p-values ($p > .7$), removing 7 items. The remaining 25 items were entered into an Exploratory Factor Analysis on Jamovi (Jamovi Project, 2019; R Core Team, 2018; Revelle, 2019), with Minimum Residuals as the extraction method, and promax as the rotation method, the latter chosen due to an expectation of correlated factors. Items with loadings of < .4 were suppressed. Based on parallel analysis, two factors were extracted. In this initial solution, there were four items that had low factor loadings (< .4), and one item that cross-loaded on both factors approximately evenly. These five items were removed, leaving 20 items. A final Exploratory Factor Analysis was run on the 20 items that were retained. Assumption checks for the final two-factor EFA model showed a significant Bartlett's test of Sphericity $\chi^2 = 817$, df = 190, $p < .001$, showing a viable correlation matrix that deviated significantly from an identity matrix. The Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO MSA) overall was .86, indicating amply sufficient sampling. The final model had twelve items that loaded onto factor 1, i.e. positive attitudes towards AI, and eight that loaded onto factor 2, i.e. negative views of AI. Hereby, the positivity and negativity of the items assumed during their creation was statistically supported, giving the factor structure good construct validity. In this solution the first factor accounted for 25.6% of the variance, and the second for 15.5%, cumulatively 41.6%. Model fit measures showed a RMSEA of .0573, 90% CI [.007, .068], TLI of .94, and the model test χ^2 = 182, df = 151, p = .046. These are acceptable fit measures. The final loadings are presented in Table 3.

--- insert Table 3 about here ---

Supported by the analyses reported, we created two subscales by taking the mean of the final retained items loading onto the relevant factors, namely positive attitudes towards AI (α = .88) and negative attitudes towards AI (α = .83). The two factors showed a factor correlation of .59, supporting the choice of the (oblique) promax rotation.

To evaluate whether there was a general attitudinal factor comprising both the negative and positive subscales, we used software entitled “Factor” (Lorenzo-Seva & Ferrando, 2019) to assess the unidimensionality of the set of 20 items retained following EFA. We ran a pure bifactor exploratory model with Maximum Likelihood extraction and promax rotation. Despite a different extraction method, the same factors were re-identified. The closeness to unidimensionality for a tentative general factor showed Unidimensional Congruence (UniCo) = 0.672, much lower than the .95 cut-off, and Explained Common Variance (ECV) = 0.482, much lower than the .85 cut-off, suggesting a lack of unidimensionality, and thus suggesting an overall scale mean should not be constructed.

Table 3: Factor loadings from the Exploratory Factor Analysis of General Attitudes towards Artificial Intelligence data

Item	Pos	Neg	U	IRC	Mean	SD
I am interested in using artificially intelligent systems in my daily life	0.78		0.43	0.64	3.56	1.03
There are many beneficial applications of Artificial Intelligence	0.77		0.40	0.68	4.22	0.82
Artificial Intelligence is exciting	0.76		0.49	0.59	3.91	1.00
Artificial Intelligence can provide new economic opportunities for this country	0.70		0.48	0.64	3.75	1.01
I would like to use Artificial Intelligence in my own job	0.66		0.54	0.59	3.13	1.24
An artificially intelligent agent would be better than an employee in many routine jobs	0.60		0.66	0.50	3.08	1.17
I am impressed by what Artificial Intelligence can do	0.60		0.63	0.53	4.13	0.89
Artificial Intelligence can have positive impacts on people's wellbeing	0.58		0.69	0.47	3.97	0.76
Artificially intelligent systems can help people feel happier	0.57		0.74	0.41	3.19	0.92
Artificially intelligent systems can perform better than humans	0.54		0.62	0.58	3.55	1.03
Much of society will benefit from a future full of Artificial Intelligence	0.49		0.63	0.57	3.55	1.03
For routine transactions, I would rather interact with an artificially intelligent system than with a human	0.47		0.79	0.39	3.15	1.22
I think Artificial Intelligence is dangerous		0.75	0.51	0.47	2.86	1.04
Organisations use Artificial Intelligence unethically		0.74	0.52	0.47	2.71	0.97
I find Artificial Intelligence sinister		0.65	0.45	0.63	3.42	1.09
Artificial Intelligence is used to spy on people		0.64	0.67	0.32	2.35	1.00

I shiver with discomfort when I think about future uses of Artificial Intelligence	0.62	0.43	0.66	3.06	1.34
Artificial Intelligence might take control of people	0.48	0.78	0.35	2.90	1.22
I think artificially intelligent systems make many errors	0.47	0.73	0.43	2.90	0.95
People like me will suffer if Artificial Intelligence is used more and more	0.41	0.59	0.60	3.23	1.20

Table 3 Note: Loadings for the retained 20 items, with factor loadings onto the positive (Pos) and negative (Neg) components, uniqueness (U, i.e. 1 minus Communality), item-rest correlation (IRC), mean, and standard deviation (SD). Note that negative items were reverse-scored in this analysis.

3.3. Technology Readiness Index: Internal Consistency checks

We checked the internal consistency of the pre-validated Technology Readiness Index as it applied to our sample. We first reverse-scored the appropriate items (i.e. the Discomfort and Insecurity subscales) and observed internal consistency metrics as follows: Innovation, α = .87, Optimism, α = .81, Discomfort, α = .74, Insecurity, α = .77, all acceptable to good, supporting dimension reduction to the pre-validated subscales by calculating means across relevant items.

3.4. Overall subscale means

Subscale means and SDs are in Table 4. Participants showed above neutral attitudes towards AI for the positive subscale, with the negative subscale averaging slightly below neutral. Our sample showed a reasonable match on the Technology Readiness Index to the values reported by Lam et al. (2008, Table 3 therein), with modest deviations, suggesting good anchoring of our sample to prior samples. The more positive aspects of technology (Innovativeness and Optimism) showed clearly positive means, the negative aspects (Discomfort, Insecurity) were also positive, but only just above neutral.

Table 4: Means and Standard Deviations for composite measures

	Mean	SD
General Attitudes towards AI		
Positive General Attitudes towards AI	3.60	0.67
Negative General Attitudes towards AI	2.93	0.75
Technology Readiness Index		
Innovativeness	3.66	1.00
Optimism	4.07	0.79
Discomfort	3.02	0.91
Insecurity	3.12	0.86

Table 4 Note: Based on reverse-scoring of negative scales, so the higher the score, the more positive the attitude, regardless of the initial polarity of the items.

3.5. Convergent and discriminant validity: Correlation and regression Analyses

We computed Pearson's correlations between the subscales of the General Attitudes towards Artificial Intelligence, and the subscales of the Technology Readiness Index. Correlations served an exploratory descriptive purpose, with their p-values only being provided for reference, but not for hypothesis evaluation. Correlation coefficients and their p-values can be seen in Table 5. Our more specific aim was to test the prediction that the Technology Readiness Index subscales that reflected technology in wider society would be more predictive of attitudes towards AI than the individually-based subscales of Technology Readiness Index. To do this on a more stringent footing than by a large number of correlations, we used multiple linear regression. Using Jamovi, we entered data from our newly created General Attitudes towards AI subscales, positive and negative in turn, as the criterion (dependent) variables, and the four subscales of the Technology Readiness Index were entered as predictor (independent) variables. Each multiple regression analysis was preceded by assumption checks, namely an autocorrelation test, collinearity check, inspection of the Q-Q plot of residuals, and residuals plots. All assumptions were met. Our primary interest was in discovering whether scores on our new General Attitudes towards AI subscales were significantly and uniquely predicted by scores on the technology readiness subscales. We report the F and p from ANOVAs testing the unique significant contribution for each predictor in Table 5. These regression analyses confirmed that Technology Readiness Index measures based on individual experiences (Innovativeness, Discomfort) did not show significant unique contributions to the subscales of the General Attitudes towards AI, while the Technology measures corresponding more closely to the use of technology in society (Optimism, Insecurity) did. Our positive General Attitudes towards AI subscale was significantly predicted by a positive subscale of the Technology Readiness Index (Optimism) only, and the negative attitudes towards AI additionally by a negative subscale (Insecurity). This supports our prediction, and underlines the need for our new measure that captures the aspects of AI that older measures of technology acceptance do not capture precisely. The pattern in these data provide evidence of convergent validity as well as discriminant validity of our new scale and subscales.

Table 5: Associations between the Technology Readiness Index and General Attitudes towards Artificial Intelligence Scale

		Innovativeness	Optimism	Discomfort	Insecurity
Positive General Attitudes towards AI	*r*	0.42	0.58	0.20	0.22
	p	<. 001	<. 001	0.051	0.029
	F	1.91	22.12	0.15	0.22
	p	0.17	< .001	0.696	0.643
Negative General Attitudes towards AI	*r*	0.27	0.44	0.27	0.43
	p	0.008	<. 001	0.007	<. 001
	F	0.08	7.19	0.32	9.94
	p	0.773	0.009	0.576	0.002

Table 5 Note: Correlations (r, p), and ANOVA tests (F, p).Technology Readiness Index subscales are listed on the top row, and our newly constructed subscales for General Attitudes towards Artificial Intelligence Scale are listed in the leftmost column, N = 99. The p-values for the correlations are based on two-tailed tests with alpha at .05. F and p are from the multiple regression's ANOVA for the factors, calculated with type 3 Sums of Squares, with dfs 1, 94. Please be reminded that all negative items on both scales were reverse-scored, so the higher a score the more positive the attitude.

3.6. Specific Applications of AI: Comfortableness and perceived capability

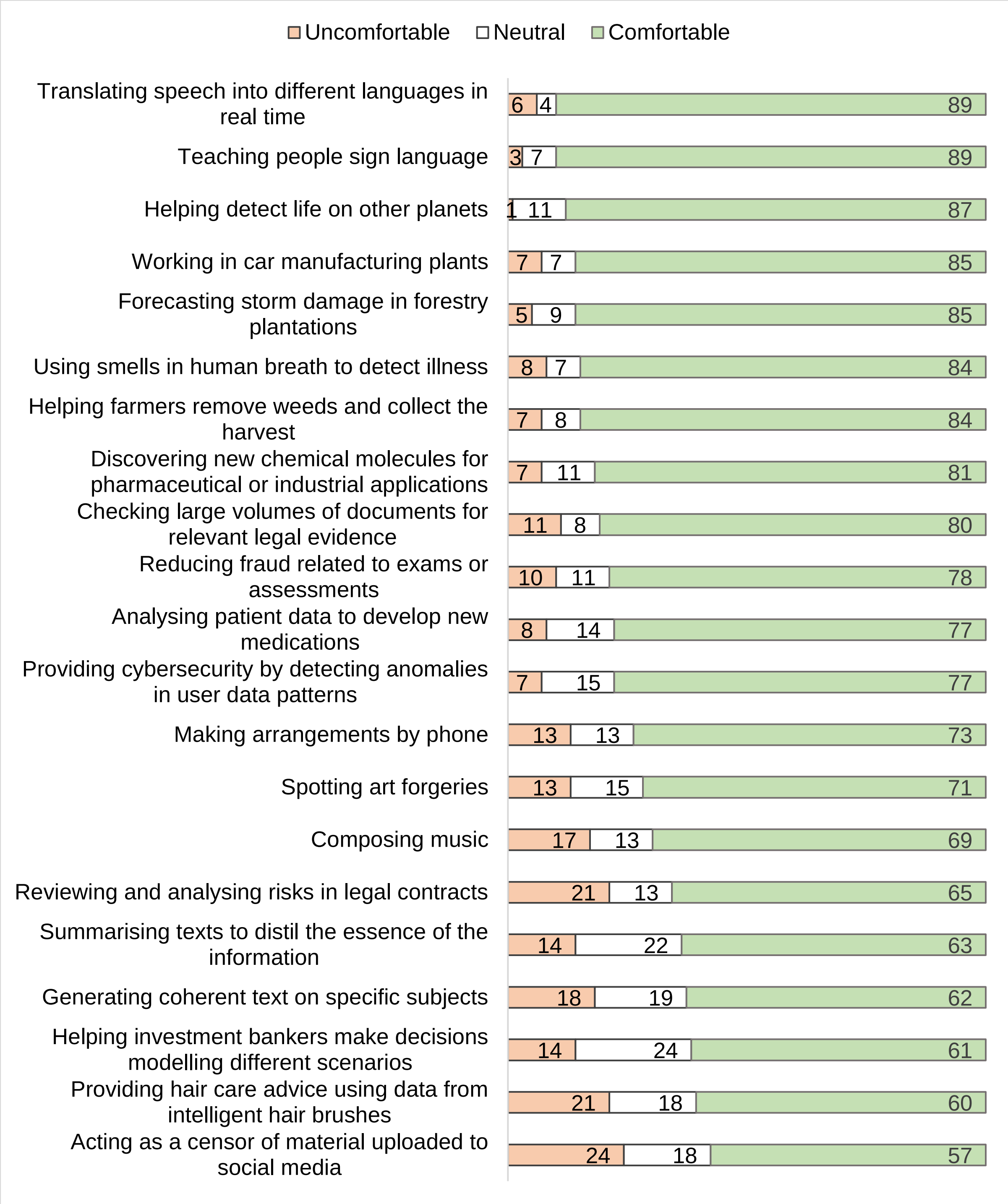
3.6.1. Descriptive Statistics: Frequencies

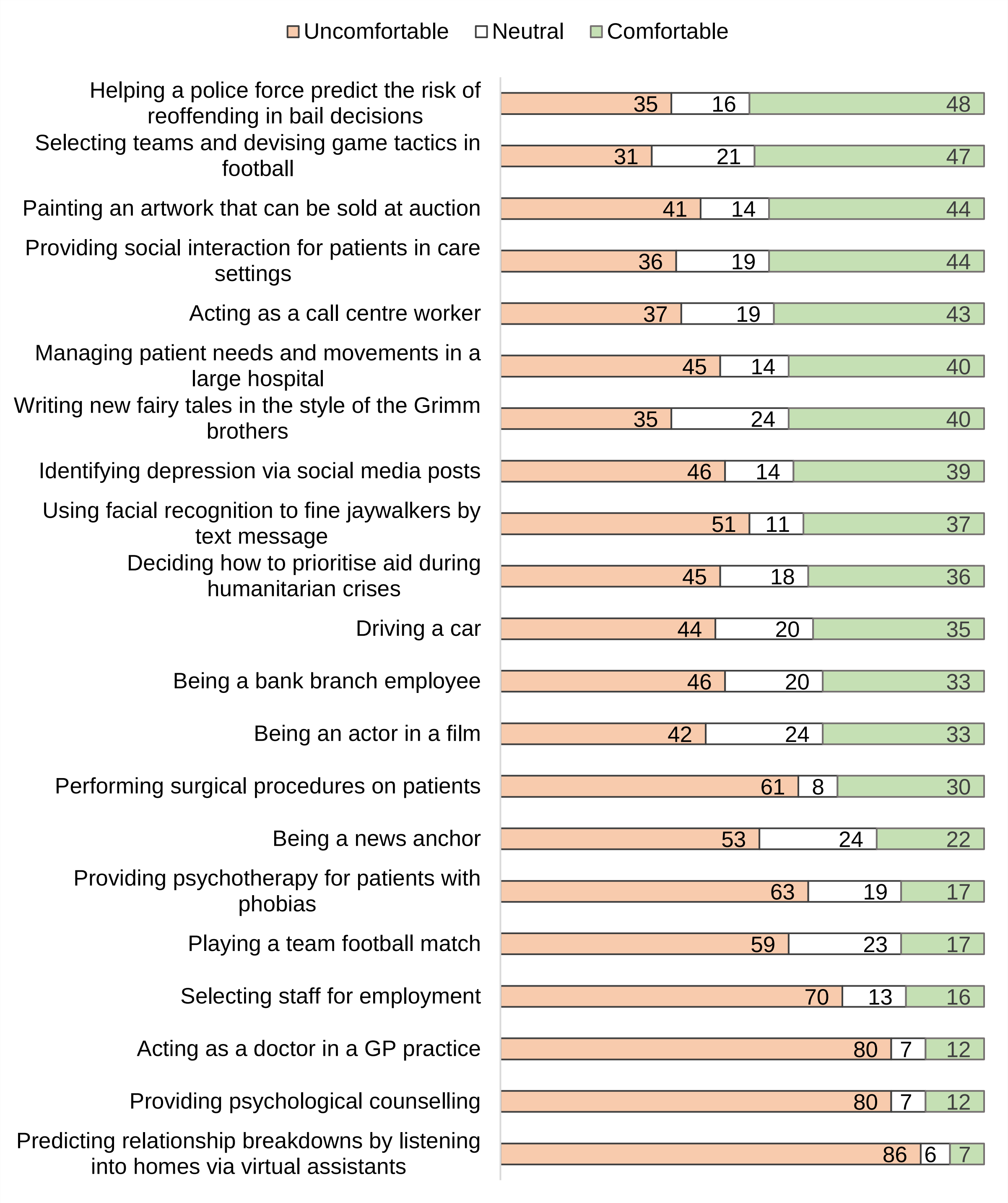
We again combined the “strongly” and “somewhat” categories to aid visual interpretation, and present frequency data for comfortableness in Figures 3A and 3B and perceived capability of AI compared to humans in Figures 4A and 4B. Participants were least comfortable with applications that may involve expert and complex social understanding (e.g. psychological counselling, acting as a doctor in general practice), while they were more comfortable with AI performing more scientific, less personal tasks (helping detect life on other planets, using smells in human breath to detect illness). The application with which people felt least comfortable was one that listened in on people’s conversations to predict relationship breakdowns. This is likely to have been thought to be a serious intrusion into people’s privacy, likely at odds with commonly accepted moral and ethical standards.

With regard to perceived capability, the applications for which AI was most frequently rated as more capable than humans all involved tasks that humans may find challenging due to a variety of limitations. These include cognitive and computational limitations (help detect life on other planets; detecting anomalies in data to aid cybersecurity; checking large volumes of documents for legal evidence), limitations in sensory capacities (detecting illness via smells in human breath), and knowledge limitations (translating speech in real time). AI applications that were most frequently rated as less capable than humans mostly involved elements of human compassion, judgement and social skills (e.g. psychological counselling, doctor in general practice, bank branch employee, selector of staff), or artistry, finesse and skill in performance (actor, news anchor, fiction writer, painter, football player).

--- Insert Figures 3A, 3B, 4A and 4B about here ---

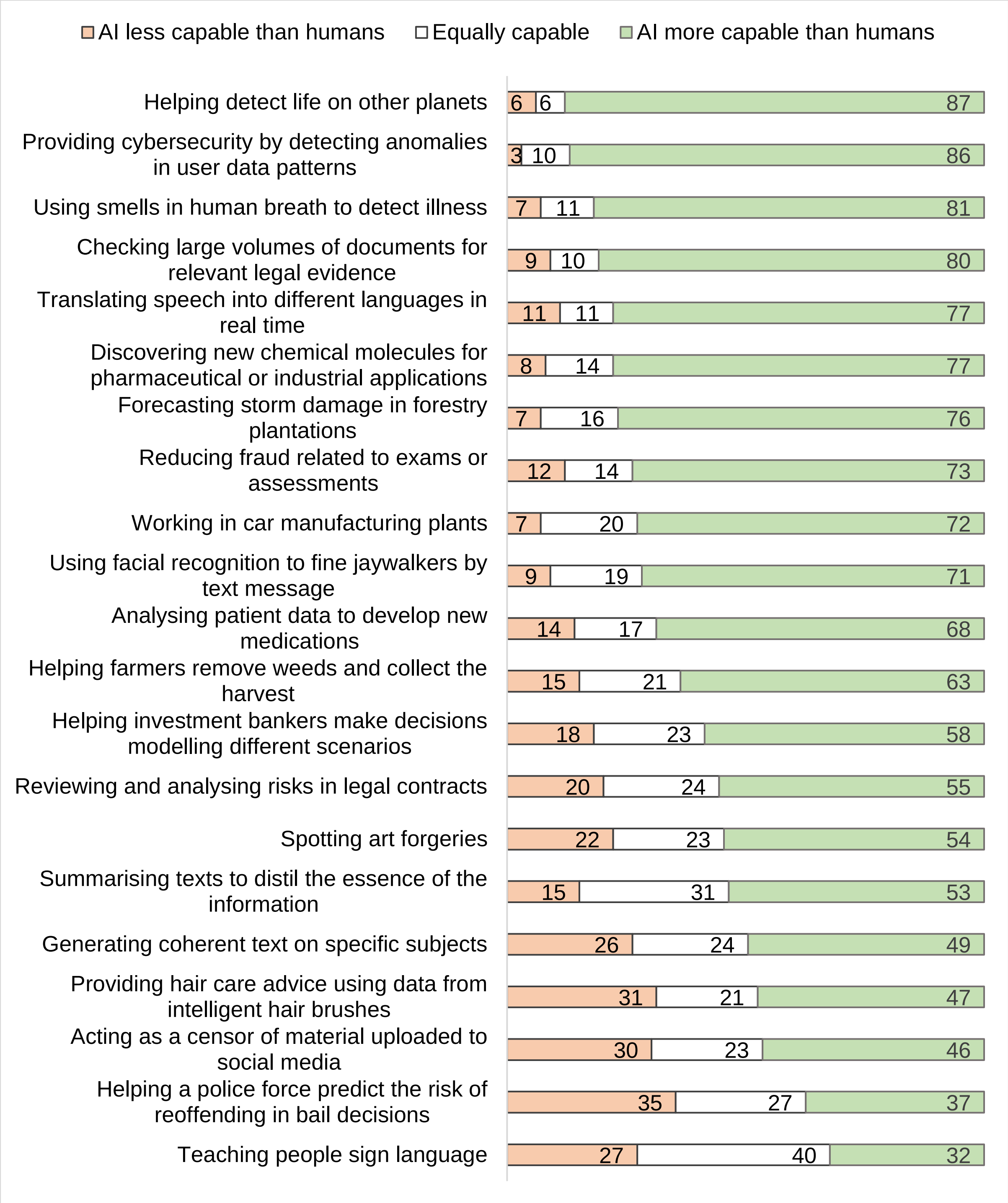
Figures 3A and 3B: Comfortableness ratings given to specific Artificial Intelligence Applications


Uncomfortable
Neutral
Comfortable
Translating speech into different languages in real time
6 4 89
Teaching people sign language
3 7 89
Helping detect life on other planets
1 11 87
Working in car manufacturing plants
7 7 85
Forecasting storm damage in forestry plantations
5 9 85
Using smells in human breath to detect illness
8 7 84
Helping farmers remove weeds and collect the harvest
7 8 84
Discovering new chemical molecules for pharmaceutical or industrial applications
7 11 81
Checking large volumes of documents for relevant legal evidence
11 8 80
Reducing fraud related to exams or assessments
10 11 78
Analysing patient data to develop new medications
8 14 77
Providing cybersecurity by detecting anomalies in user data patterns
7 15 77
Making arrangements by phone
13 13 73
Spotting art forgeries
13 15 71
Composing music
17 13 69
Reviewing and analysing risks in legal contracts
21 13 65
Summarising texts to distil the essence of the information
14 22 63
Generating coherent text on specific subjects
18 19 62
Helping investment bankers make decisions modelling different scenarios
14 24 61
Providing hair care advice using data from intelligent hair brushes
21 18 60
Acting as a censor of material uploaded to social media
24 18 57



Figures 3A and 3B Note: Figure 3A lists the applications rated as highest in comfortableness, Figure 3B the lowest. Data are collapsed over “somewhat” and “strongly”, while retaining neutral. N = 99 and raw frequencies are presented. The “uncomfortable” category is presented in orange on the left of the bars, neutral in white, centrally, and “comfortable” in green as the rightmost part of the bars.

Figures 4A and 4B: Perceived capability of specific AI applications in comparisons to humans



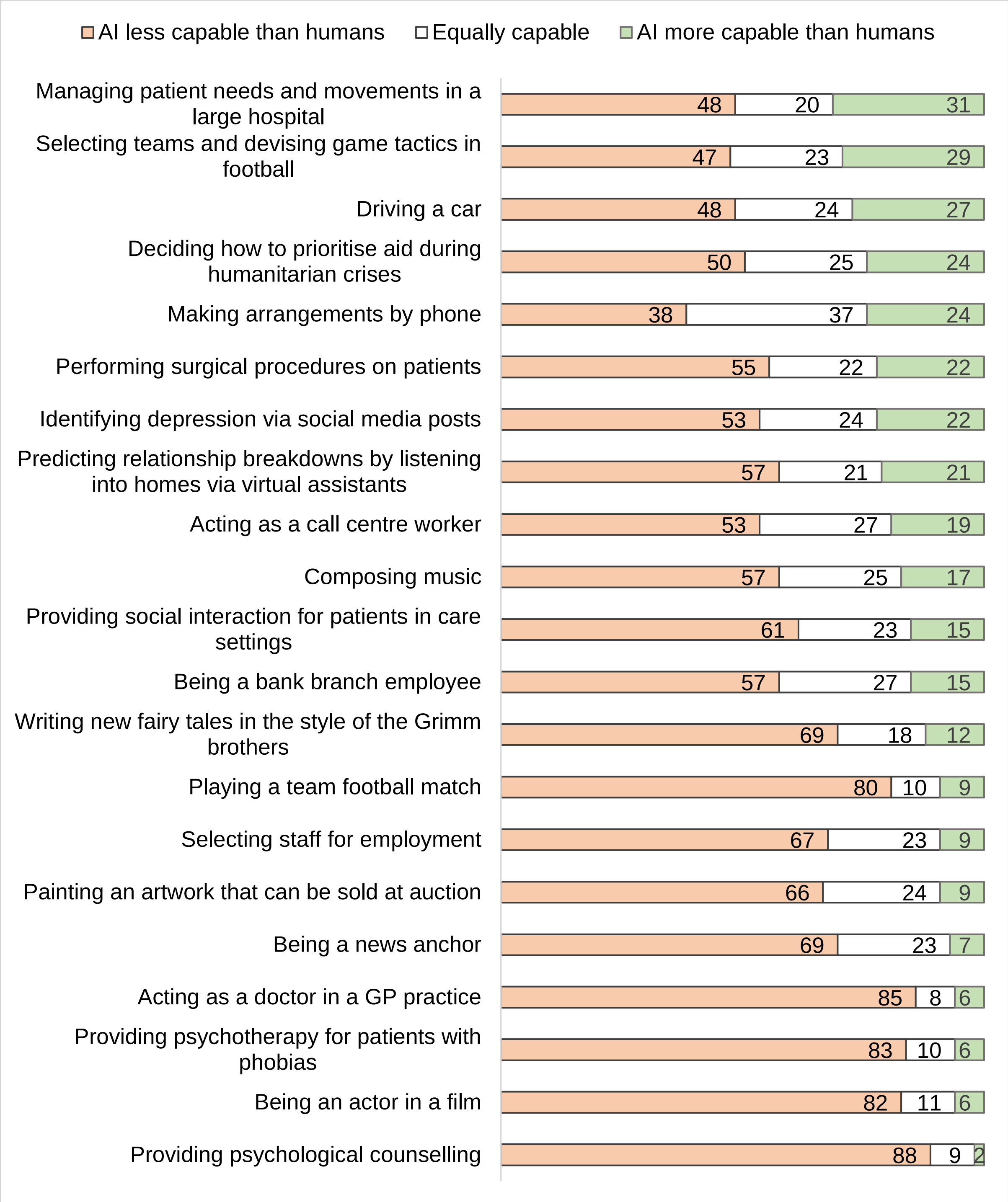


Figure 4A and 4B Note: The data are collapsed over “somewhat less / more ” and “much less / more”, while retaining neutral. N = 99 and raw frequencies are presented. The “AI less capable than humans” category is presented in orange on the left of the bars, neutral in white, centrally, and “AI more capable than humans” in green as the rightmost part of the bars. Figure 4A lists the AI applications rated as highest in capability, Figure 4B the lowest.

3.6.2. By-items correlations for comfortableness and perceived capability

Impressionistically, capability ratings showed some overlap in rankings with the comfortableness ratings. However, there were also differences in relative rankings. To explore the extent to which rated comfortableness could be captured as a function of perceived capability of AI in comparison with humans, a correlation was run on the average rating for each item on both these measures (see Supplementary Materials for the processed data). Shapiro-Wilks tests detected no significant deviation from a normal distribution for either measure. Therefore, a Pearson's correlation was run, giving $r = .83$, $N = 42$, $p < .001$, $r^2 = .69$. This was a relatively high association between the two variables, but with 31% of residual variance.

To explore which items may play a particularly strong role in the residual variance we calculated the standardised residuals (ZRes) for each pair of data when predicting comfortableness from perceived capability in a linear regression. We inspected the items with values that were more than 1.96 z-score removed from zero in either direction. At one end of the spectrum, these were "Using facial recognition to fine jaywalkers by text message" (ZRes = -3.02), and "Predicting relationship breakdowns by listening into homes via virtual assistants" (ZRes = -2.83) where comfortableness was rated much lower than could be expected from the capability rating. The reasons for this are most probably because both applications were intrusive, yet AI may be perceived as highly capable of the tasks. At the other end of the spectrum, people showed higher levels of comfortableness than could be expected from their capability ratings for applications described as "Composing music" (ZRes = 1.99) and "Teaching people sign language" (ZRes = 2.27).

3.6.3. Exploratory Factor Analysis: Comfortableness

We ran Exploratory Factor Analyses with the same parameters as before for attitudes (Section 3.2.2.2), but without any reverse scoring. There were no a priori expectations for factors. We eliminated 13 items involved in multiple very low correlations ($r < .1$, a slightly more stringent cut-off than before because of larger number of items). Initial Exploratory Factor Analysis (Minimum Residuals, promax) on the remaining items identified two factors based on parallel analysis, in which 6 items did not load onto either factor, and these were also removed. A final Exploratory Factor Analysis was run with the remaining 23 items. Bartlett's Test of Sphericity showed, $\chi^2 = 1090$, $df = 253$, $p < .001$. KMO MSA overall was .86. The final analysis identified two factors accounting for 23.8% and 18.8% of the variance, respectively, total 42.5%. The factors were correlated with $r = .64$. The RMSEA was .075, 90% CI [.045, .080], TLI .88, and the model test showed $\chi^2 = 291$, $df = 208$, $p < .001$, suggesting a reasonable fit to support dimension reduction and naming latent factors. Factor loadings for comfortableness are presented in Table 6. Factor 1 primarily captured items with a high mean, indicating high levels of comfortableness. In turn, many items loading on this factor appeared to feature readily automatable tasks, often based on big data. Factor 2 primarily captured items with a low mean. In turn, many of these items described task that required a human judgement. Two measures were created, based on the mean across the relevant items. The first was a factor which we named "Comfortableness with AI applications for big data and automation" (Factor 1, $\alpha = .90$). The second was "Comfortableness with AI applications for Human judgement tasks" (Factor 2, $\alpha = .86$). Unidimensionality assessment was irrelevant and is therefore not reported.

--- insert Table 6 about here ---

Table 6: Factor loadings from the Exploratory Factor Analysis of Comfortableness with Specific Applications of Artificial Intelligence

	F1	F2	U	IRC	Mean	SD
Reducing fraud related to exams or assessments	0.86		0.31	0.70	4.10	1.06
Using smells in human breath to detect illness	0.75		0.54	0.53	4.21	1.02
Discovering new chemical molecules for pharmaceutical or industrial applications	0.73		0.44	0.65	4.33	1.00
Translating speech into different languages in real time	0.72		0.62	0.42	4.54	0.91
Helping farmers remove weeds and collect the harvest	0.66		0.59	0.54	4.33	1.00
Reviewing and analysing risks in legal contracts	0.64		0.48	0.65	3.62	1.28
Forecasting storm damage in forestry plantations	0.63		0.59	0.56	4.30	0.91
Spotting art forgeries	0.59		0.66	0.49	4.04	1.20
Working in car manufacturing plants	0.59		0.50	0.66	4.35	0.99
Providing hair care advice using data from intelligent hair brushes	0.56		0.63	0.54	3.57	1.30
Checking large volumes of documents for relevant legal evidence	0.54		0.66	0.52	4.11	1.03
Helping investment bankers make decisions modelling different scenarios	0.48		0.48	0.69	3.70	1.15
Acting as a censor of material uploaded to social media	0.41		0.82	0.37	3.42	1.38
Selecting staff for employment		0.85	0.47	0.48	2.13	1.21
Being a bank branch employee		0.79	0.44	0.59	2.77	1.34
Acting as a doctor in a GP practice		0.72	0.53	0.56	1.77	1.11
Managing patient needs and movements in a large hospital		0.67	0.55	0.57	2.96	1.32
Acting as a call centre worker		0.65	0.53	0.60	3.08	1.36
Providing social interaction for patients in care settings		0.56	0.71	0.45	3.09	1.35
Driving a car		0.52	0.68	0.49	2.79	1.43

Writing new fairy tales in the style of the Grimm brothers	0.50	0.70	0.49	3.05	1.41
Deciding how to prioritise aid during humanitarian crises	0.50	0.59	0.61	2.78	1.34
Selecting teams and devising game tactics in football	0.44	0.75	0.46	3.26	1.31

Table 6 Note: Factor loadings onto Factor 1 (F1, Comfortableness with AI applications for big data and automation) and Factor 2 (F2, Comfortableness with AI applications for Human judgement tasks), with Uniqueness (U), item-rest correction (IRC), item mean and standard deviation (SD) for the 23 items retained in the Exploratory Factor Analysis of comfortableness ratings.

3.6.4. Exploratory Factor Analysis: Perceived capability

The same Exploratory Factor Analysis process as was run on the comfortableness data was performed on the capability data, again without a priori structural expectations. Multiple low correlations ($r < .1$) were detected in 14 items, and these were eliminated, as were four items that showed low loadings on either of the two factors extracted in the initial Exploratory Factor Analysis. Further iterations revealed further low loading or cross-loading items, which were removed in turn. The final analysis was on 21 items. Bartlett's Test of Sphericity in this analysis was significant, $\chi^2 = 1122$, df = 210, $p < .001$, while KMO MSA was .87. Two factors were extracted based on parallel analysis accounting for 24.5% and 22.9% of the variance, respectively, total 47.4%. The correlation between the two factors was $r = .57$. RMSEA was .081, 90% CI [.053, .089], TLI .88, and the model test showed $\chi^2 = 254$, df = 169, $p < .001$, suggesting a reasonable fit, which would support dimension reduction and the naming of latent factors. Factor loadings for perceived capability are presented in Table 7. Factor 1 contained many items which involve human judgement or skilled finesse, and we named this "Perceived capability of AI for tasks involving human judgement", creating a factor mean based on the items loading onto this factor (α =.89). Factor 2 seemed to contain items that all involve algorithmic processing of "big data" and we named this factor "Perceived capability of AI for tasks involving big data" (α = .90).

--- insert Table 7 about here ---

Table 7: Factor loadings from the Exploratory Factor Analysis of Perceived capability of specific applications of Artificial Intelligence

	Factor 1	Factor 2	U	IRC	Mean	SD
Providing psychotherapy for patients with phobias	0.81		0.47	0.54	1.79	0.96
Acting as a doctor in a GP practice	0.77		0.52	0.53	1.64	0.94
Selecting staff for employment	0.73		0.53	0.56	2.14	1.02
Performing surgical procedures on patients	0.71		0.52	0.59	2.44	1.21
Being a bank branch employee	0.71		0.55	0.55	2.39	1.13
Driving a car	0.60		0.52	0.65	2.70	1.25
Deciding how to prioritise aid during humanitarian crises	0.58		0.43	0.71	2.63	1.23
Playing a team football match	0.56		0.75	0.37	1.81	1.13
Managing patient needs and movements in a large hospital	0.54		0.46	0.70	2.80	1.31
Identifying depression via social media posts	0.49		0.64	0.57	2.57	1.14
Making arrangements by phone	0.48		0.68	0.53	2.85	1.06
Acting as a call centre worker	0.47		0.70	0.51	2.57	1.17
Painting an artwork that can be sold at auction	0.42		0.82	0.37	2.10	1.03
Helping detect life on other planets		0.93	0.34	0.48	4.54	0.90
Discovering new chemical molecules for pharmaceutical or industrial applications		0.90	0.34	0.54	4.17	1.02

Checking large volumes of documents for relevant legal evidence	0.88	0.33	0.57	4.25	0.97
Reducing fraud related to exams or assessments	0.85	0.30	0.65	3.91	1.08
Reviewing and analysing risks in legal contracts	0.64	0.43	0.67	3.54	1.13
Spotting art forgeries	0.64	0.55	0.56	3.59	1.26
Helping investment bankers make decisions modelling different scenarios	0.57	0.45	0.69	3.59	1.16
Summarising texts to distil the essence of the information	0.47	0.71	0.47	3.58	1.03

Table 7 Note: Factor loadings onto Factor 1 (F1, Perceived capability of AI for tasks involving Human Judgement) and Factor 2 (F2, Perceived capability of AI for tasks involving Big Data), with Uniqueness (U), item-rest correction (IRC), item mean and standard deviation (SD) for the 21 items retained in the Exploratory Factor Analysis of perceived capability ratings.

3.7. Means and Standard Deviations for comfortableness and perceived capability composite measures

We computed means and standard deviations for the factors of comfortableness and perceived capability (see Table 8). Participants showed positive views of the use of AI for tasks involving big data or automation, but negative views of AI being used in tasks involving human judgement, rating their perceived capabilities particularly low.

--- insert Table 8 about here ---

Table 8: Means and Standard Deviations for the composite measures of Comfortableness and Perceived capability

	Mean	SD
Comfortableness		
Comfortableness with AI for tasks involving big data / automation	4.05	0.74
Comfortableness with AI for tasks involving human judgement	2.77	0.88
Perceived capability		
Perceived capability of AI for tasks involving big data	3.89	0.82
Perceived capability of AI for tasks involving human judgement	2.34	0.74

Table 8 Note: Means and SDs for composite measures. For all scales, 3 was the neutral centre. Scores below that point reflect negative views, above reflect positive views. Minimum possible score was 1, maximum possible score was 5.

3.8. Cross-validation general and specific views: Correlation and regression analyses

To explore to what extent individuals' attitudes towards AI in general were associated with their comfortableness with specific applications, and their perception of the capability of AI, we again ran correlation analyses on a descriptive exploratory basis, with p-values reported for reference, but not to test hypotheses (see Table 9). We double-checked the key patterns using the more stringent ANOVA factor contributions via linear multiple regression models. We predicted the positive and then the negative subscale of General Attitudes towards AI from the four factors related to specific applications (four-predictor model), reporting F and p for each of the coefficients in Table 9. The strong prediction of the General Attitudes from comfortableness with specific applications provides cross-validation of the General

Attitudes towards AI subscales. The four-predictor model suggested that rated capabilities of specific applications of AI were less strongly predictive of general attitudes, suggesting these were more independent. To explore the pattern in more detail, we checked whether the capability ratings predicted the positive subscale if the comfortableness ratings were eliminated from the model (a two-predictor model), and their coefficients were significant ($p < .001$ for Big Data, $p = .002$ for Human Judgement). However, perceived capability did not significantly predict negative general attitudes in an equivalent two-predictor model ($p = .09$ for Big Data, $p = .56$ for Human Judgement). Overall, the pattern provides cross-validation between the general and specific views.

--- insert Table 9 about here ---

Table 9: Correlations and multiple regression coefficients associating subscales of General Attitudes towards Artificial and Comfortableness with and Perceived capability of specific applications of Artificial Intelligence

		Comfortableness with AI for…		Perceived capability of AI for…	
		big data / automation	human judgement	big data	human judgement
Positive General Attitudes towards AI	*r*	0.65	0.68	0.57	0.52
	p	< .001	< .001	< .001	< .001
	F	10.14	16.29	0.24	0.13
	p	.002	< .001	.63	.71
Negative General Attitudes towards AI	*r*	0.46	0.36	0.24	0.18
	p	< .001	< .001	.018	.081
	F	15.80	4.25	3.62	0.95
	p	< .001	.04	.06	.33

Table 9 Note: Correlations (r, p), and ANOVA tests (F, p). General Attitudes towards Artificial Intelligence subscales are listed in the leftmost column, and cross-validation factor composites capturing attitudes towards specific applications of Artificial Intelligence are listed on the top row, N = 99. The p-values for the correlations are based on two-tailed tests with alpha at .05. F and p are from the multiple regression's ANOVA for the factors, calculated with type 3 Sums of Squares, with dfs 1, 94. Please be reminded that all negative items on both scales were reverse-scored, so the higher a score the more positive the attitude.

4. Discussion

The Discussion contains a consideration of the psychometrics and validity of the GAAIS, followed by an evaluation of more global conceptual findings of this study, and an evaluation of the limitations, future research that is needed to build on the work presented here, finishing with a conclusion.

4.1. Scale psychometrics and validity

The study yielded an initially validated General Attitudes towards Artificial Intelligence Scale (GAAIS) with positive and negative subscales, which had good psychometric properties. A unidimensionality assessment showed that the subscales should not be merged into an overall composite scale score. Subscales of the Technology Readiness Index that related to societal use of technology predicted our General Attitudes towards AI subscales as hypothesised. These regression patterns provided convergent validity for our new subscales. The associations were not maximal, and did not involve subscales of the Technology Readiness Index that related to individual user experiences of technology. This provided discriminant validity, which is evidence of the novelty and distinctiveness of our new scale. Our rationale for our new AI scale was that older Technology Acceptance Scales such as the TAM (Davis, 1989) reflect users' individual choices to use technology, but AI often involves decisions by others. Our results support the need for measurement tools that capture these key aspects of AI, and our new scale addresses this gap.

The subscale averages provided valuable information on attitudes towards Artificial Intelligence. Overall, participants held slightly positive views on the positive subscale, which consisted of items expressing enthusiasm and perceived utility of AI. The sample mean was just below neutral for the negative subscale. This balance of both positive and negative views in the same sample concurs well with the findings from recent surveys discussed in Section 1.1.

Cross-validation of general attitudes using specific applications was successful, adding further validity to our new scale. It was useful that these insights emerged “bottom-up” from a list of AI innovations, and that clustering was likewise identified “bottom-up” via the statistical analysis, providing independent cross-validation. However, comfortableness was a better predictor of General Attitudes towards AI than overall perceived capability. This is probably because people may hold very positive attitudes towards the potential benefits of AI, but may nevertheless make a separate assessment about current limitations of specific AI applications. This would seem a rational position to hold given the current limitations of AI, especially given the novelty of the specific applications in our items. In contrast, people may have rated comfortableness more hypothetically, assuming that a system was fully capable of the task described. Furthermore, comfortableness is more closely related psychologically to general attitudinal constructs than capability assessments are. The latter were probably based on rational assessments, because we asked participants to judge AI vs. humans on each task. In contrast, comfortableness is likely to be more emotionally based. Overall, comfortableness with specific applications formed good cross-validation for the new General Attitudes towards Artificial Intelligence Scale.

4.2. Conceptual insights

The study yielded important conceptual insights. One important source of insight is an inspection of items that were retained following exploratory factor analysis, how these items clustered, and which items had the strongest item-rest correlations. For the general attitudes, items that loaded onto the positive factor expressed societal or personal benefits of AI, or a preference of AI over humans in some contexts (e.g. in routine transactions), with some items capturing emotional matters (AI being exciting, impressive, making people happier, enhancing their wellbeing). Items involving personal use of AI were also present (use in own job, interest in using AI). In all, the balance in the positive items was towards utility, both in the number of items, and in the items with the highest item-rest correlations (see Table 3). In the negative subscale, more items were eliminated from the initial pool, and those that were retained were dominated by emotions (sinister, dangerous, discomfort, suffering for “people like me”), and dystopian views of the uses of AI (unethical, error prone-ness, taking control, spying). Here, the more emotional items tended to have higher item-rest correlations, suggesting that the retained negative items may reflected more affective attitudes. Some negative items were not retained in factor analysis, because they did not correlate strongly with the other item set. Two such eliminated negative items “The rise of Artificial Intelligence poses a threat to people's job security” and “I am concerned about Artificial Intelligence applications mining my personal data” showed high levels of participant concern in the survey data. However, they did not load onto the negative factor. Overall, the positive items were dominated by utility, and negative items by negative emotions and dystopian concerns.

Insights could also be gained from the clustering of the data on specific applications of AI. When asked about their comfortableness with these specific applications as well as their perceived capability in comparison with humans, two clusters emerged via the data analysis. In one cluster, there were applications that featured big data or other readily automatable tasks, and participants held positive views about these, feeling comfortable with them, and attributing high capabilities to such applications. Underlying this may be the common feature that these applications aided humans in their endeavours (e.g. molecule screening, aiding bankers, detecting fraud), but where humans are not replaced by AI, and AI did not gain autonomy or control. In the other cluster there were applications involving some aspect of human judgement, empathy, skill, or social understanding, and participants felt negatively towards AI performing these functions. Discomfort and low capability were, for example, associated with AI performing staff selection, decisions on the allocation of aid, and driving a car. This is an important finding, which suggests that people may make clear distinctions in the classes of tasks for which they will currently accept AI. Another important source of conceptual insights regarding specific applications comes from the survey data, particularly via an inspection of applications that attracted ratings near the extremes. It is interesting to note that among the lowest rated applications of AI, both for comfortableness and capability, were applications related to individual health interactions, e.g. acting as a doctor. This raises issues in the context of ongoing work developing medical AI applications (see e.g. Fenech et al. 2018, Kadvany, 2019). Our data also showed very low ratings for applications involving psychotherapy. This is despite evidence of people's tendency to anthropomorphise and form emotional connections with extremely basic classic psychotherapy systems such as ELIZA (Weizenbaum, 1976). This is another important finding, as our data suggest that there may be initial resistance to using such applications, and their developers may need to overcome this if they want their applications to be effective.

Further conceptual insights were gained by correlating perceived capability and comfortableness. While these correlated strongly across applications, we argued that an ethical dimension led to a partial decoupling between comfortableness and perceived capability, which was pronounced in some items. For example, while some applications may be perceived as capable (e.g. fining people for offences based on automatic facial recognition), participants reported levels of discomfort that were out of line with the perceived capability of such applications. This may be related to the intrusiveness of these types of applications (see also House of Commons, 2019, p. 14 on automatic live facial recognition). Notwithstanding this, live facial recognition has now been introduced in London (Metropolitan Police, 2020), with the important aim of fighting crime. More recently, facial recognition has also been deployed in Moscow for surveillance of compliance with coronavirus / Covid-19 quarantine regulations (Rainsford, 2020). Our findings suggest the general public may not feel entirely comfortable with these types of applications in all contexts, at least not in the UK. It would be interesting and useful to examine to what extent the public may perceive the end as justifying the means in such types of applications of AI. This is likely to vary across cultures and contexts.

4.3. Evaluation of limitations and future research

It is important to evaluate the limitations as well as the strengths of our research. First, our sample size was relatively small, for resource-related reasons. A reason why a small sample could be problematic is that Exploratory Factor Analysis needs a reasonable sample size to be valid. However, the KMO MSAs for all Exploratory Factor Analyses showed good sampling adequacy. KMO MSA is an empirical measure of sampling adequacy that supersedes sample-size heuristics. These heuristics often work on a worst-case scenario, and can therefore overestimate the sample sizes needed (see e.g. Taherdoost, Sahibuddin, & Jalaliyoon, 2014, for a recent discussion). Our data also showed good internal consistency indices. Thus, we argue that our sample size was sufficient for the analyses reported. A further potential weakness is that the population from which the sample was drawn may not be sufficiently informed to express valid views on AI. Similarly, both the newspaper articles and our summaries may have oversimplified the complexities of the AI applications (Wilks, 2019). However, it was our intention to survey ordinary people's reactions to the type of information that may reach them via general media outlets. News channels often simplify matters, while headlines condense and simplify matters even more. This condensed information was likely to reflect many people's exposure to AI developments. Finally, our scale went through initial validation using Exploratory Factor Analysis, but would benefit from further validation via a Confirmatory Factor Analysis with a new and larger sample. It would also be beneficial to run studies that link the new measure to other samples, demographics, and other social factors. This is planned as future research.

4.4. Conclusion

In summary and conclusion, our research produced a usable two-factor General Attitudes towards Artificial Intelligence Scale (GAAIS) with good psychometric properties, convergent and discriminant validity, and good cross-validation patterns. It will be helpful to further validate this tool in future research with a new, larger sample. Attitudes towards AI need to be gauged regularly, given the rapid development in these technologies and their profound impact on society. Data on acceptance of AI by the public can inform legislators and organisations developing AI applications on ways in which their introduction may need to be managed if these applications are to be accepted by the end users. Useful measurement tools are therefore important. Our new initially validated General Attitudes towards AI Scale is a useful tool to help accomplish these aims. We include it ready for use in Appendix B.

References

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

Barnes, M., Elliott, L.R., Wright, J., Scharine, A., and Chen, J. (2019). *Human–Robot Interaction Design Research: From Teleoperations to Human–Agent Teaming* (Technical Report ARL-TR-8770). Army Research Laboratory Aberdeen Proving Ground, MD 21005 United States. https://apps.dtic.mil/dtic/tr/fulltext/u2/1079789.pdf

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

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

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

Chui, M., Manyika, J., & Miremadi, M. (2016). Where machines could replace humans—and where they can't (yet). *McKinsey Quarterly*, *30*(2), 1-9. Via http://www.oregon4biz.com/assets/e-lib/Workforce/MachReplaceHumans.pdf

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319-340. https://doi.org/10.2307/249008

Duquemin, H., Rabaiotti, G., Tomlinson, I. & Stephens, M. (2019). *Services sector, UK: 2008 to 2018*. Office for National Statistics, 2 April 2019. https://www.ons.gov.uk/economy/economicoutputandproductivity/output/articles/servicessectoruk/2008to2018

Edelman (2019). Artificial Intelligence (AI) Survey. https://www.edelman.com/sites/g/files/aatuss191/files/2019-03/2019_Edelman_AI_Survey_Whitepaper.pdf

Jamovi project (2019). *jamovi*. (Version 0.9) [Computer Software]. https://www.jamovi.org/

Fast, E., & Horvitz, E. (2017). Long-term trends in the public perception of artificial intelligence. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17),* p 963-969. https://www.microsoft.com/en-us/research/wp-content/uploads/2016/11/long_term_AI_trends.pdf

Fenech, M., Strukelj, N., & Buston, O. (2018). Ethical, social and political challenges of artificial intelligence in health. http://futureadvocacy.com/wp-content/uploads/2018/04/1804_26_FA_ETHICS_08-DIGITAL.pdf

Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological forecasting and social change*, *114*, 254-280. https://doi.org/10.1016/j.techfore.2016.08.019

Frontier Economics (2018) The impact of artificial intelligence on work: an evidence review prepared for the Royal Society and the British Academy. London: Frontier Economics. https://royalsociety.org/-/media/policy/projects/ai-and-work/frontier-review-the-impact-of-AI-on-work.pdf

Granulo, A., Fuchs, C., & Puntoni, S. (2019). Psychological reactions to human versus robotic job replacement. *Nature human behaviour*, *3*(10), 1062-1069. https://doi.org/10.1038/s41562-019-0670-y

House of Commons Science and Technology Committee (2019). The work of the Biometrics Commissioner and the Forensic Science Regulator. Nineteenth Report of Session 2017–19. https://publications.parliament.uk/pa/cm201719/cmselect/cmsctech/1970/1970.pdf

Kadvany, E. (2019). 'The glue between therapist and patient': can Silicon Valley fix mental health care?. *The Guardian*, 30 July 2019.

https://www.theguardian.com/society/2019/jul/29/mental-health-care-therapy-silicon-valley

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

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

Lorenzo-Seva, U. & Ferrando, P. J. (2019). *Factor*. Departament de Psicologia Universitat Rovira i Virgili, Tarragona (Spain), 2006, Updated 2nd May, 2019. http://psico.fcep.urv.es/utilitats/factor/index.html

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

Metropolitan Policy (2020). Met begins operational use of Live Facial Recognition (LFR) technology. Press Release, 24 January 2020. http://news.met.police.uk/news/met-begins-operational-use-of-live-facial-recognition-lfr-technology-392451

Olhede SC, Wolfe PJ. 2018 The growing ubiquity of algorithms in society: implications, impacts and innovations. *Philosophical Transactions of the Royal Society. A* **376**: 20170364. http://dx.doi.org/10.1098/rsta.2017.0364

Parasuraman, A., & Colby, C. L. (2015). An Updated and Streamlined Technology Readiness Index: TRI 2.0. *Journal of Service Research*, *18*(1), 59–74. https://doi.org/10.1177/1094670514539730

R Core Team (2018). *R: A Language and envionment for statistical computing*. [Computer software]. https://cran.r-project.org/

Rainsford, S. (2020). Coronavirus: Russia uses facial recognition to tackle Covid-19. BBC News, 3 April 2020. https://www.bbc.co.uk/news/av/world-europe-52157131/coronavirus-russia-uses-facial-recognition-to-tackle-covid-19

Revelle, W. (2019). *psych: Procedures for Psychological, Psychometric, and Personality Research*. [R package]. https://cran.r-project.org/package=psych

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/

Schaefer, K. E., Chen, J. Y., Szalma, J. L., & Hancock, P. A. (2016). A meta-analysis of factors influencing the development of trust in automation: Implications for understanding autonomy in future systems. *Human factors*, *58*(3), 377-400. https://doi.org/10.1177%2F0018720816634228

Sheridan, T. B. (2019). Individual differences in attributes of trust in automation: measurement and application to system design. *Frontiers in psychology*, *10*, 1117. https://doi.org/10.3389/fpsyg.2019.01117

Sloman, S. A., & Lagnado, D. A. (2005). The Problem of Induction. In: Holyoak, K. & Morrison, R. (Eds.). *Cambridge handbook of thinking and reasoning*. Cambridge: Cambridge University Press.

Taherdoost, H., Sahibuddin, S. & Jalaliyoon, N. (2014). Exploratory Factor Analysis; Concepts and Theory. In Balicky, J. *Advances in Applied and Pure Mathematics*: *Proceedings of the 2nd International Conference on Mathematical, Computational and Statistical Sciences (MCSS '14).* Gdańsk, Poland: World Scientific and Engineering Academy and Society Press.

Vayena, E., Blasimme, A., & Cohen, I. G. (2018). Machine learning in medicine: Addressing ethical challenges. *PLoS medicine*, *15*(11), PLoS Med 15(11): e1002689. https://doi.org/10.1371/journal.pmed.1002689

Weizenbaum, J. (1976). *Computer power and human reason: from judgment to calculation*. New York: W. H. Freeman and Company.

White, S., Lacey, A., & Ardanaz-Badia, A. (2019). The probability of automation in England: 2011 and 2017. Office for National Statistics. https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/articles/theprobabilityofautomationinengland/2011and2017

Wilks, Y. (2019). *Artificial Intelligence: Modern Magic or Dangerous Future?* London: Icon Books Ltd.

Zhang, B. and 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

Appendices

Appendix A: News stories and their sources

Items derived from newspaper articles reporting new AI technologies, summarised by the authors to form items rated for comfortableness and perceived capability with the URLs for their sources.

Item	Summary, item phrasing	URL for newspaper article
1	Translating speech into different languages in real time	https://www.theguardian.com/technology/2019/feb/17/is-the-era-of-artificial-speech-translation-upon-us
2	Providing psychological counselling	https://www.theguardian.com/lifeandstyle/shortcuts/2019/jan/02/woebots-ai-counselling-future-therapy-mental-health

3	Painting an artwork that can be sold at auction	https://www.theguardian.com/technology/audio/2019/jan/18/can-a-computer-be-creative-chips-with-everything-podcast
4	Being a news anchor	https://www.theguardian.com/world/2018/nov/09/worlds-first-ai-news-anchor-unveiled-in-china
5	Composing music	https://www.theguardian.com/music/2018/oct/22/ai-artificial-intelligence-composing
6	Selecting staff for employment	https://www.theguardian.com/technology/2018/oct/10/amazon-hiring-ai-gender-bias-recruiting-engine
7	Being an actor in a film	https://www.theguardian.com/film/2018/aug/16/tony-kaye-sam-khoze-2nd-born-robot-actors
8	Spotting art forgeries	https://www.theguardian.com/us-news/2018/aug/06/the-new-tool-in-the-art-of-spotting-forgeries-artificial-intelligence
9	Providing psychotherapy for patients with phobias	https://www.theguardian.com/science/2018/jul/11/automated-virtual-reality-therapy-helps-people-overcome-phobia-of-heights
10	Providing social interaction for patients in care settings	https://www.theguardian.com/commentisfree/2018/jul/02/robo-carers-human-principles-technology-care-crisis
11	Playing a team football match	https://www.theguardian.com/science/2018/jun/26/the-world-cup-of-robot-football-no-need-for-humans-to-worry-yet
12	Making arrangements by phone	https://www.theguardian.com/technology/2018/may/08/google-duplex-assistant-phone-calls-robot-human
13	Acting as a call centre worker	https://www.theguardian.com/business/2018/may/12/robot-technology-threat-terminist-uk-call-centre-workforce
14	Being a bank branch employee	https://www.theguardian.com/cities/2018/may/14/shanghai-robot-bank-china-worlds-first-human-free-branch-construction
15	Generating coherent text on specific subjects	https://www.independent.co.uk/life-style/gadgets-and-tech/news/ai-text-generator-fake-news-articles-misuse-dangerous-open-source-a8780686.html
16	Selecting teams and devising game tactics in football	https://www.independent.co.uk/life-style/gadgets-and-tech/news/football-ai-coach-artificial-intelligence-wingate-finchley-fc-big-bang-fair-a8742466.html
17	Teaching people sign language	https://www.independent.co.uk/news/science/sign-language-ai-help-video-game-artificial-intelligence-computer-deaf-hearing-a8739141.html
18	Identifying depression via social media posts	https://www.independent.co.uk/news/health/facebook-ai-depression-mental-health-social-media-anxiety-machine-learning-a8585301.html
19	Discovering new chemical molecules for pharmaceutical or industrial applications	https://www.independent.co.uk/news/science/university-glasgow-chemistry-robot-machine-learning-lee-cronin-a8453851.html
20	Acting as a doctor in a GP practice	https://www.independent.co.uk/voices/loneliness-kills-artificial-intelligence-chatbot-doctors-health-risk-diagnoses-a8423321.html
21	Writing new fairy tales in the style of the Grimm brothers	https://www.independent.co.uk/news/long_reads/ai-robot-brothers-grimm-fairytale-write-story-the-princes-and-fox-a8393826.html
22	Using smells in human breath to detect illness	https://www.independent.co.uk/news/science/ai-artificial-intelligence-smell-detect-illness-science-technology-a8394706.html
23	Forecasting storm damage in forestry plantations	https://www.independent.co.uk/news/science/artificial-intelligence-help-repair-storm-damage-costs-billions-a8319411.html
24	Using facial recognition to fine jaywalkers by text message	https://www.independent.co.uk/news/world/asia/china-police-facial-recognition-technology-ai-jaywalkers-fines-text-wechat-weibo-cctv-a8279531.html

25	Helping detect life on other planets	https://www.independent.co.uk/news/science/archaeology/news/alien-life-ai-artifical-intelligence-space-predict-ufo-planets-a8288326.html
26	Providing hair care advice using data from intelligent hair brushes	https://www.independent.co.uk/news/business/news/loreal-modiface-takeover-ai-makeup-beauty-digital-firm-virtual-augmented-reality-a8259301.html
27	Deciding how to prioritise aid during humanitarian crises	https://www.independent.co.uk/news/science/artifical-intelligence-disaster-response-humanitarian-crisis-ai-help-a8319361.html
28	Predicting relationship breakdowns by listening into homes via virtual assistants	https://www.independent.co.uk/life-style/gadgets-and-tech/alexa-relationship-dating-google-home-advice-imperial-college-research-a8658976.html
29	Helping a police force predict the risk of reoffending in bail decisions	https://www.ft.com/content/9559efbe-2958-11e9-a5ab-ff8ef2b976c7
30	Summarising texts to distil the essence of the information	https://www.ft.com/content/a3943548-e9cb-11e8-94da-a6478f64c783
31	Analysing patient data to develop new medications	https://www.ft.com/content/e450a688-ddfb-11e8-b173-ebef6ab1374a
32	Performing surgical procedures on patients	https://www.ft.com/content/5230b9c4-fd3a-11e8-b03f-bc62050f3c4e
33	Checking large volumes of documents for relevant legal evidence	https://www.ft.com/content/ad042a78-052f-11e9-9d01-cd4d49afbbe3
34	Managing patient needs and movements in a large hospital	https://www.ft.com/content/ed665ed8-cdfd-11e8-9fe5-24ad351828ab
35	Reviewing and analysing risks in legal contracts	https://www.ft.com/content/50b0eba4-d063-11e8-9a3c-5d5eac8f1ab4
36	Helping farmers remove weeds and collect the harvest	https://www.ft.com/content/5854088a-ddda-11e8-b173-ebef6ab1374a
37	Helping investment bankers make decisions modelling different scenarios	https://www.ft.com/content/3ab7cbf4-8281-11e8-96dd-fa565ec55929
38	Providing cybersecurity by detecting anomalies in user data patterns	https://www.ft.com/content/d8e073d2-869e-11e8-9199-c2a4754b5a0e
39	Working in car manufacturing plants	https://www.ft.com/content/3a453bb8-c000-11e8-84cd-9e601db069b8
40	Driving a car	https://www.ft.com/content/8c94ab24-c77b-11e8-ba8f-ee390057b8c9
41	Reducing fraud related to exams or assessments	https://www.ft.com/content/540e77fa-9fe2-11e8-85da-eeb7a9ce36e4
42	Acting as a censor of material uploaded to social media	https://www.ft.com/content/9728b178-59b4-11e8-bdb7-f6677d2e1ce8

Appendix B: The General Attitudes towards Artificial Intelligence Scale (GAAIS)

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, 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 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, Disgree = 2, Neutral = 3, Agree = 4, and Strongly agree = 5. Score the items marked "Negative" in reverse so that Strongly disagree = 5, Disgree = 4, Neutral = 3, 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.