WHAT MAKES A (RO)BOT SMART? EXAMINING THE ANTECEDENTS OF PERCEIVED INTELLIGENCE IN THE CONTEXT OF USING PHYSICAL ROBOTS, SOFTWARE ROBOTS, AND CHATBOTS AT WORK

Research full-length paper

Makkonen, Markus, University of Jyvaskyla, Jyvaskyla, Finland, markus.v.makkonen@jyu.fi

Salo, Markus, University of Jyvaskyla, Jyvaskyla, Finland, markus.t.salo@jyu.fi

Pirkkalainen, Henri, Tampere University, Tampere, Finland, henri.pirkkalainen@tuni.fi

Abstract

In recent years, the acceptance and use of intelligent robots and other kinds of intelligent systems have begun to gain more and more attention also in information systems research. Here, many studies have found the perceived intelligence of robots to act as one critical antecedent for their acceptance and use, but few studies have focused on the antecedents of perceived intelligence itself. In this study, we aimed to address this gap in prior research by examining the effects of individual intelligence dimensions on the overall intelligence perception of robots in the work context. In addition, we also examined the potential differences in these effects as well as in the individual intelligence dimensions and overall intelligence perception themselves between three common types of robots: physical robots, software robots, and chatbots. These examinations were based on online survey data from 1,080 present or prior users of robots at work. In summary, we found that adaptability, personality, autonomy, and multifunctionality act as the most influential antecedents of perceived intelligence in the case of all three types of robots. In addition, we also found that software robots and chatbots perform better than physical robots in most individual intelligence dimensions and in overall intelligence perception.

Keywords: Perceived Intelligence, Intelligence Dimensions, Physical Robots, Software Robots, Chatbots, Work Context.

1 Introduction

During the next few decades, the rapid diffusion of intelligent robots and other kinds of intelligent systems will revolutionise our lives not only at home but also at work (Servoz, 2019). On one hand, this diffusion is likely to result in substantial job replacement and retraining, as it has been predicted that about 14% of current jobs may be completely automated and about 32% of them are likely to change considerably in the future (OECD, 2022). On the other hand, the diffusion is likely to promote productivity (Servoz, 2019) and may also provide more human-centred benefits, such as more meaningful work (Smids et al., 2020). However, in order for such revolutions to take place, these technologies must first be adopted by their potential users, which is why it is not surprising that the acceptance and use of intelligent robots and other kinds of intelligent systems have begun to gain attention also in information systems (IS) research. Here, most prior studies have focused on the antecedents of their acceptance and use, highlighting especially the role of perceived intelligence as a critical factor. For example, perceived intelligence has been found to positively affect factors like the perceived usefulness and perceived ease of use of these technologies (Moussawi et al., 2021, 2022) as well as their trustworthiness and the trust toward them (W. Kim et al., 2020; H. Kim et al., 2022; Moussawi et al., 2021; Moussawi and Benbunan-Fich, 2021), thus also having a positive effect on the intention to adopt (Tussyadiah and Park, 2018; Pillai and Sivathanu, 2020; Moussawi et al., 2021), use (Blut et al., 2021; Chuah, 2021; Moussawi and Benbunan-Fich, 2021; Cai et al., 2022; H. Kim et al., 2022), and continue using them (Balakrishnan et al., 2022; Maroufkhani et al., 2022; Moussawi et al., 2022).

However, whereas many prior studies have focused on perceived intelligence as an antecedent of acceptance and use, few of them have focused on the antecedents of perceived intelligence itself. This is a serious shortcoming because an understanding of such antecedents is considered a critical prerequisite for the systematic promotion of perceived intelligence in new product development and marketing (Rijsdijk et al., 2007; Rijsdijk and Hultink, 2009; Rokonuzzaman et al., 2022). For example, no prior study that we are aware of has examined how individual intelligence dimensions, such as the ability to act autonomously or learn, affect the overall intelligence perception of robots. In this study, we aimed to address this gap in prior research by examining the effects of individual intelligence dimensions on the overall intelligence perception of robots in the work context. Moreover, because these effects may vary between different kinds of robots, we also examined the potential differences in the effects as well as in the individual intelligence dimensions and overall intelligence perception themselves between three common types of robots: physical robots, software robots, and chatbots. These examinations were done by utilising data from 1,080 present or prior users of robots at work that was collected with an online survey and analysed with structural equation modelling (SEM). As a result, we contribute to IS literature with a better theoretical understanding of the antecedents and definition of perceived intelligence in the context of using robots at work as well as multiple implications for practice.

This paper consists of six sections. After this introductory section, we briefly present our research model in Section 2. The methodology and results of the paper are reported in Sections 3 and 4, and the results are discussed in more detail in Section 5. Finally, the paper concludes with a brief discussion of the limitations of the study and some potential paths for future research in Section 6.

2 Research Background and Research Model

When reviewing prior literature on intelligent robots and other kinds of intelligent systems in both IS and other relevant disciplines, we found several studies that highlighted especially the role of perceived intelligence as a critical antecedent for their use and acceptance. For example, in the case of physical and non-physical service robots, perceived intelligence has been found to positively affect the trustworthiness of and the trust toward the robots (W. Kim et al., 2020; H. Kim et al., 2022), rapport-building between humans and robots (Qiu et al., 2020), the intention to adopt the robots (Tussyadiah and Park, 2018), and the intention to use the robots (Chuah, 2021; H. Kim et al., 2022). Of these, the last-mentioned effect has also been confirmed in the meta-analysis by Blut et al. (2021). In turn, in the case of chatbots, perceived intelligence has been found to positively affect the intention to adopt the bots (Pillai and Sivathanu, 2020), the intention to use the bots (Cai et al., 2022), as well as the attitude toward and the intention to continue using the bots (Balakrishnan et al., 2022). Finally, in the case of artificial intelligence based virtual assistants, perceived intelligence has been found to positively affect their perceived usefulness (Moussawi et al., 2021, 2022), their perceived ease of use (Moussawi et al., 2021), their perceived enjoyment (Moussawi et al., 2021, 2022), their perceived anthropomorphism (Moussawi et al., 2021, 2022; Moussawi and Benbunan-Fich, 2021), the trust toward them (Moussawi et al., 2021; Moussawi and Benbunan-Fich, 2021), the evaluations of them (Guha et al., 2022), the satisfaction with them and with their use (Marikyan et al., 2022; Moussawi et al., 2022), the positive disconfirmation of expectations (Moussawi et al., 2022), consumer brand engagement (McLean et al., 2021), smart-shopping perception (Aw et al., 2022), the attitude toward them and the purchase intention through them (Balakrishnan and Dwivedi, 2021), the intention to adopt them (Moussawi et al., 2021), the intention to use them (Moussawi and Benbunan-Fich, 2021), and the intention to continue using them (Maroufkhani et al., 2022; Moussawi et al., 2022).

However, in all the aforementioned studies, perceived intelligence has been operationalised as a simple unidimensional construct that has typically been measured by using the perceived intelligence dimension of the Godspeed scale (a commonly used scale for measuring user perceptions in human–robot interaction research) by Bartneck et al. (2009). In contrast, few prior studies have employed more complex multidimensional operationalisations. One exception to this is the study by Lee and Shin (2018) that examined how individual intelligence dimensions affected overall intelligence perception in the context of smartphone use by consumers. However, due to contextual differences, the generalisability of its findings to the context of using robots at work can be seen as questionable. For identifying the individual intelligence dimensions, the study by Lee and Shin (2018) used the taxonomy by Rijsdijk and Hultink (2003, 2009; Rijsdijk et al., 2007) that was developed for the context of intelligent or smart products. It defines seven individual intelligence dimensions: (1) *autonomy* (i.e., the extent to which a product is able to operate in an independent and goal-directed way without the interference of the user), (2) *adaptability* (i.e., the ability of a product to improve the match between its functioning and its environment), (3) *reactivity* (i.e., the ability of a product to react to changes in its environment), (4) *multifunctionality* (i.e., the ability of a single product to fulfil multiple functions), (5) *ability to cooperate* (i.e., the ability of a product to collaborate with other products in order to achieve a common goal), (6) *humanlike interaction* (i.e., the degree to which a product communicates and interacts with the user in a natural, human way), and (7) *personality* (i.e., the ability of a product to show the properties of a credible character). Together, these dimensions form the overall intelligence perception of a product, meaning that the better the product performs in each dimension, the more intelligent it is perceived to be. In theory, all the dimensions are assumed to be distinct from each other, meaning that performing well in one dimension does not necessarily correlate with performing well in the other dimensions. However, in practice, as noted by Rijsdijk et al. (2007), such correlations may still occur because the information technology (IT) that is used to promote the intelligence of a product in one dimension can often be utilised to promote its intelligence also in the other dimensions.

Over the years, also other similar taxonomies have been proposed. One example of these is the taxonomy by Rokonuzzaman et al. (2022), which is an application of the original taxonomy by Rijsdijk and Hultink (2003, 2009; Rijsdijk et al., 2007) to the context of the Internet of Things (IoT) for measuring the Smartness of a Thing (SoT). It defines ten individual intelligence dimensions: (1) *ability to cooperate*, (2) *autonomy*, (3) *environmental agility*, (4) *learning*, (5) *novelty*, (6) *personality*, (7) *real-time information processing*, (8) *two-way communication*, (9) *upgradability*, and (10) *visual appeal*. As can be seen, many of these dimensions are identical to the dimensions of the original taxonomy, whereas others are novel and intended to cover the somewhat broader concept of smartness by Rokonuzzaman et al. (2022) in comparison to the narrower concept of intelligence in the original taxonomy. Although not so widely used, other examples of similar taxonomies are the ones by Maass and Varshney (2008), Porter and Heppelmann (2014), Dawid et al. (2017), as well as Novak and Hoffman (2019).

In this study, we based our research model on the original taxonomy by Rijsdijk and Hultink (2003, 2009; Rijsdijk et al., 2007), which we used to identify the individual intelligence dimensions that affect the overall perceived intelligence of robots, as illustrated in Figure 1. The reason for selecting this taxonomy over the others was based on the generic nature of its dimensions in comparison to the more context-specific nature of the dimensions of other similar taxonomies, which promotes its applicability from the context of intelligent products also to the context of intelligent robots. In fact, Rijsdijk and Hultink (2003, 2009) as well as Rijsdijk et al. (2007) themselves demonstrated this applicability in their studies by using robotic vacuum cleaners, robotic lawnmowers, as well as humanoid and animal robots as examples of intelligent products. So, in a sense, many intelligent robots, although not all, can also be seen as intelligent products. More specifically, we included in our research model six individual intelligence dimensions that we define as follows: (1) *autonomy* (i.e., the extent to which a robot is able to operate in an independent and goal-directed way without the interference of the user), (2) *adaptability* (i.e., the ability of a robot to improve the match between its functioning and its environment), (3) *reactivity* (i.e., the ability of a robot to react to changes in its environment), (4) *multifunctionality* (i.e., the ability of a single robot to fulfil multiple functions), (5) *ability to cooperate* (i.e., the ability of a robot to collaborate with other systems in order to achieve a common goal), and (6) *personality* (i.e., the ability of a robot to show the properties of a credible character). Of these, the first five dimensions are the same ones that were also included in the study by Lee and Shin (2018) and can be seen to be highly applicable not only to the context of intelligent products but also to the context of intelligent robots. For example, it makes sense to assess practically all types of robots in each of these five dimensions, and robots that perform better in these dimensions can be assumed to be perceived as more intelligent by their users. In addition to them, we also included the personality dimension due to the observed strong correlations between perceived intelligence and perceived anthropomorphism in several prior studies on the acceptance and use of intelligent systems (e.g., Moussawi et al., 2021, 2022; Moussawi and Benbunan-Fich, 2021). However, we excluded the humanlike interaction dimension for three reasons. First, it can be seen to overlap with the personality dimension because a robot that is perceived as better in humanlike interaction is also likely to be perceived as having a more humanlike personality. Second, it can be seen as somewhat less generic in comparison to the other dimensions because, whereas it makes sense to assess practically all types of robots in the autonomy, adaptability, reactivity, multifunctionality, ability to cooperate, and personality dimensions, it only makes sense to assess a robot in the humanlike interaction dimension if it actually interacts with humans in some substantial way. This is not the case with all types of robots, such as many industrial robots that are working in factories. Third, humanlike interaction was also assessed as the least important individual intelligence dimension in terms of overall intelligence perception in the study by Rijsdijk et al. (2007) based on interviews with multiple intelligent product experts.

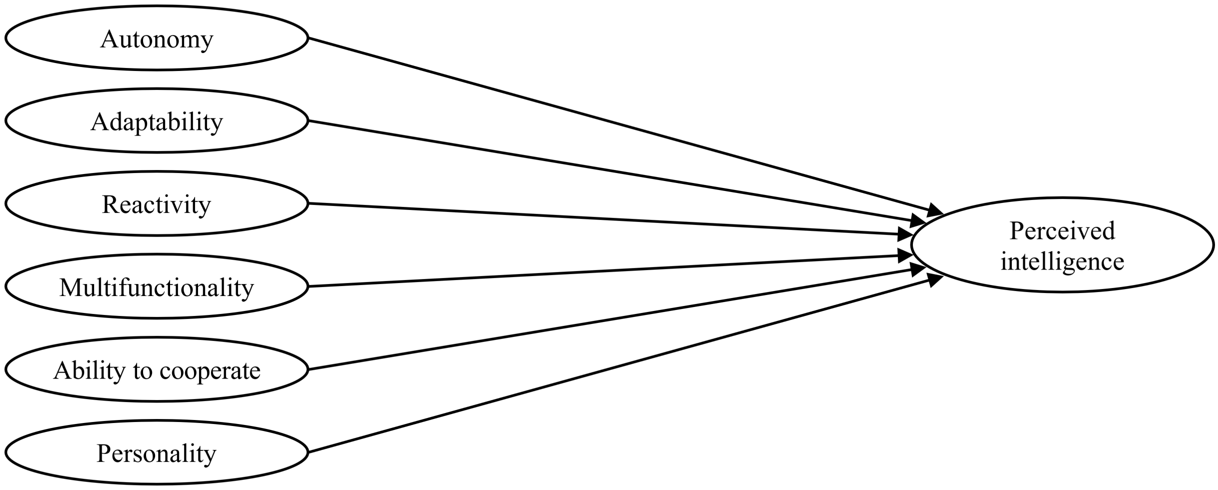

Figure 1. Research model of the study.

3 Methodology

The data for testing the research model was collected with an online survey that was conducted by using the LimeSurvey service. The respondents for the survey were recruited by using an online crowdsourcing service, which have been deemed a reliable and valid method of collecting data also in IS research (Lowry et al., 2016). More specifically, we used the Prolific service, which has been found to provide better or at least equal data quality and a more heterogeneous population of participants than its alternatives, such as the Amazon Mechanical Turk (MTurk) service (Peer et al., 2017, 2021). Because we were interested in using robots at work, we recruited only respondents who were employed either full-time (≥ 30 h / week) or part-time (< 30 h / week) and resided in the UK, the US, or Canada, which are all countries that have been found to have high usage rates of robots at work (World Economic Forum, 2020) and which also constitute a homogeneous Anglospheric cultural domain. In order to promote data quality, as recommended by Peer et al. (2014) and the guidelines by Jia et al. (2017) for using online crowdsourcing services in IS research, we also recruited only respondents who had a minimum approval rate of 98% for their submissions as well as had a minimum of 20 submissions and a maximum of 10,000 submissions. All the respondents were paid a monetary reward for their participation that was above the US federal minimum wage of USD 7.25 per hour.

The survey consisted of two parts. The first part was a short one-minute screening survey, in which the respondents were shown a list of different types of intelligent robots or systems and asked to select the ones that they have used at work. The second part was the actual survey, which was targeted at only those respondents who had indicated that they had used some type of an intelligent robot or system at work. In both surveys, we distinguished between four different types of intelligent robots or systems that were defined to the respondents as follows:

- **Physical robots:** A physical robot refers to robot technology with a physical embodiment (in contrast to software robots, chatbots, etc.). Typically, a physical robot is a programmable machine that has a movable physical structure and is capable of executing specific tasks with a varying degree of autonomy (e.g., industrial robots, service robots, social robots, and care robots).
- **Software robots:** Software robots or software robotics refers to the use of bot programs (except chatbots) to automate computer tasks usually performed by people. The term software robotics is often used synonymously with the term robotic process automation (RPA).
- **Chatbots:** A chatbot is a piece of software that is used to simulate an online chat conversation via text or text-to-speech instead of providing direct contact with a human (e.g., a customer service person).
- **Virtual assistants (omitted in this study, cf. Section 4):** A virtual assistant is a piece of software that can perform tasks, which are often based on verbal or written commands, for a user (e.g., Apple Siri, Amazon Alexa, Microsoft Cortana, and Google Home / Nest).

In the actual survey, in order to avoid response fatigue, the respondents were inquired in detail about the use of only one of the aforementioned types of technologies: (1) physical robots, if they had used them, (2) software robots, if they had used them and had not used physical robots, (3) chatbots, if they had used them and had not used physical or software robots, and (4) virtual assistants, if they had used them and had not used physical robots, software robots, or chatbots. In addition, all the respondents were asked questions about their demographics, job, as well as hopes and fears related to using robots at work. The six individual intelligence dimensions and overall perceived intelligence were measured reflectively by four items each. The items for measuring the six individual intelligence dimensions were adapted from the studies by Rijsdijk and Hultink (2003, 2009), Rijsdijk et al. (2007), as well as Rokonuzzaman et al. (2022), whereas the items for measuring the overall perceived intelligence were adapted from the study by Lee and Shin (2018). The items were shown to the respondents in an order that was randomised individually for each respondent. The exact wordings of the items are reported in

Appendix A. The measurement scale was a standard five-point Likert scale (1 = strongly disagree … 5 = strongly agree). In order to avoid forced responses, the respondents also had the option not to respond to a specific item, which resulted in a missing value.

The collected data was analysed with covariance-based structural equation modelling (CB-SEM) by using the Mplus version 8.8 software (Muthén and Muthén, 2022) and following the guidelines by Gefen et al. (2011) for SEM in administrative and social science research. As the model estimator, we used the MLR option of Mplus, which stands for maximum likelihood estimator robust to non-normal data. The potential missing values were handled by using the FIML option of Mplus, which stands for full information maximum likelihood and uses all the available data in model estimation. The potential differences between different types of robots were examined with multiple group analysis (MGA) by following the testing procedure proposed by Steenkamp and Baumgartner (1998) for establishing measurement invariance. In it, increasingly strict constraints on parameter equality are added across the groups and the fit of the resulting constrained model is compared to the fit of the unconstrained model. If the constraints result in no statistically significant deterioration in model fit, then the hypothesis on the specific type of measurement invariance is supported. Configural invariance is tested by estimating the model separately in each group while constraining only the simple model structure as equal across the groups, whereas metric and scalar invariance are tested by additionally constraining the indicator loadings and indicator intercepts as equal across the groups. After this, the differences in the model constructs can be tested by examining their estimated mean scores in each group. Of the groups, one typically acts as a reference group, in which the construct mean scores are fixed to zero and against which the construct mean scores of the other groups are compared. In addition, the differences in the effects between the model constructs can be tested by constraining the estimated effect sizes as equal across the groups. As a statistical test for examining the potential deteriorations in model fit, we used the χ^2 test of difference, in which the value of the test statistic was corrected with the Satorra-Bentler (2001) scaling correction factor (SCF) due to the use of MLR as the model estimator. However, because the χ^2 test of difference is known to suffer from a similar sensitivity to sample size as the χ^2 test of model fit, we also considered the potential changes in the model fit indices, as suggested by Steenkamp and Baumgartner (1998).

4 Results

In total, we received 1,436 responses to the actual survey. However, of them, we had to omit 228 responses due to the respondents not having used any of the four types of technologies, despite having indicated so in the screening survey. In addition, we omitted 54 responses due to an invalid response to an attention check item and 12 responses due to a missing value in all the items that measured the six individual intelligence dimensions and overall perceived intelligence. Due to the too small number of responses from a statistical perspective, we also omitted 62 responses from those who had responded to the measurement items in the case of virtual assistants. This resulted in a sample size of 1,080 responses to be used in this study. Of the respondents, 375 (34.7%) had responded to the measurement items in the case of physical robots, 474 (43.9%) had responded to the measurement items in the case of software robots, and 231 (21.4%) had responded to the measurement items in the case of chatbots. The demographics of the sample in terms of the gender, age, country of residence, educational attainment, employment status, and total work experience of the respondents are reported in Table 1, showing the sample to have a good representativeness of the working-age population. Note that the employment status could also be a combination of two or more options (e.g., employed and student). The age of the respondents ranged from 18 to 75 years, with a mean of 35.9 years and a standard deviation of 10.5 years, and their mean response time to the survey was about 15 minutes.

In the following three subsections, we first assess the reliability and validity of the estimated model at the both indicator and construct levels, then report the model fit and model estimates, and finally examine the potential differences between different types of robots.

	N	%		N	%
Gender			Undergraduate	557	51.6
Man	529	49.0	Graduate or postgraduate	285	26.4
Woman	541	50.1	Other or unknown	4	0.4
Other or unknown	10	0.9	**Employment status**		
Age			Employed full-time	857	79.4
18–29 years	337	31.2	Employed part-time	147	13.6
30–39 years	406	37.6	Self-employed	79	7.3
40–49 years	192	17.8	Unable to work	1	0.1
50 years or over	145	13.4	Student	59	5.5
Country of residence			**Total work experience**		
UK	721	66.8	Under a year	12	1.1
US	283	26.2	1–2 years	41	3.8
Canada	70	6.5	3–5 years	148	13.7
Other or unknown	6	0.6	6–10 years	227	21.0
Educational attainment			11–20 years	328	30.4
Secondary or high school	95	8.8	Over 20 years	323	29.9
Some post-secondary studies	139	12.9	Unknown	1	0.1

Table 1. Demographics of the sample (N = 1,080).

4.1 Model reliability and validity

Construct reliabilities were assessed by using the composite reliabilities (CR) of the constructs (Fornell and Larcker, 1981), which are commonly expected to be greater than or equal to 0.7 (Nunally and Bernstein, 1994). The CR of each construct is reported in the first column of Table 2, showing that all the constructs met this criterion. In turn, construct validities were assessed by examining the convergent and discriminant validities of the constructs by using the two criteria proposed by Fornell and Larcker (1981). They are both based on the average variance extracted (AVE) of the constructs, which refers to the average proportion of variance that a construct explains in its indicators. In order to have acceptable convergent validity, the first criterion expects each construct to have an AVE of at least 0.5. This means that, on average, each construct should explain at least half of the variance in its indicators. The AVE of each construct is reported in the second column of Table 2, showing that all the constructs met this criterion. In turn, in order to have acceptable discriminant validity, the second criterion expects each construct to have a square root of AVE greater than or equal to its absolute correlations with the other model constructs. This means that, on average, each construct should share at least an equal proportion of variance with its indicators to what it shares with these other model constructs. The square root of AVE of each construct (the on-diagonal cells) and the construct intercorrelations (the off-diagonal cells) are reported in the remaining columns of Table 2, showing that this criterion was also met by all the constructs.

Finally, indicator reliabilities and validities were assessed by using the standardised loadings of the indicators, which are reported in Appendix A together with the means and standard deviations (SD) of the indicator scores as well as the percentages of missing values in the indicators. In the typical case of each indicator loading on only one construct, the standardised loading of each indicator is commonly expected to be statistically significant and greater than or equal to 0.707 (Fornell and Larcker, 1981). This is equivalent to the standardised residual of each indicator being less than or equal to 0.5, meaning that at least half of the variance in each indicator is explained by the construct on which it loads. As can be seen, all the indicators met this criterion. In addition, the percentages of missing values were very modest, with 990 out of the 1,080 respondents (91.7%) actually having no missing values at all.

Construct	CR	AVE	AUT	ADA	REA	MF	ATC	PER	PI
Autonomy (AUT)	0.900	0.692	0.832						
Adaptability (ADA)	0.903	0.701	0.719	0.837					
Reactivity (REA)	0.898	0.687	0.665	0.712	0.829				
Multifunctionality (MF)	0.898	0.687	0.382	0.499	0.443	0.829			
Ability to cooperate (ATC)	0.892	0.674	0.333	0.444	0.437	0.646	0.821		
Personality (PER)	0.915	0.730	0.623	0.571	0.562	0.366	0.248	0.854	
Perceived intelligence (PI)	0.867	0.620	0.683	0.736	0.604	0.542	0.429	0.633	0.787

Table 2. Composite reliabilities, average variances extracted, and construct intercorrelations.

4.2 Model fit and model estimates

The results of model estimation in terms of model fit, the standardised effect sizes and their statistical significance, as well as the proportion of explained variance (R^2) in overall perceived intelligence are reported in the first column of Table 3. Model fit was assessed by using the χ^2 test of model fit and four model fit indices recommended in recent methodological literature (Hu and Bentler, 1999): the comparative fit index (CFI), the Tucker-Lewis index (TLI), the root mean square error of approximation (RMSEA), and the standardised root mean square residual (SRMR). Of them, the χ^2 test of model fit rejected the null hypothesis of the model fitting the data, which is common in the case of large samples (Bentler and Bonett, 1980). In contrast, the values of the four model fit indices all clearly met the cut-off criteria recommended by Hu and Bentler (1999): CFI $\geq$ 0.95, TLI $\geq$ 0.95, RMSEA $\leq$ 0.06, and SRMR $\leq$ 0.08. Thus, we consider the overall fit of the model acceptable. We also found no serious signs of multicollinearity or common method bias. For example, the variance inflation factor (VIF) scores calculated from the factor scores were all clearly less than four (Hair et al., 2018), and the Harman's single factor test (Podsakoff et al., 2003) suggested a very bad model fit ($\chi^2(350) = 7{,}128.916$, $p < 0.001$, CFI = 0.600, TLI = 0.568, RMSEA = 0.134, SRMR = 0.113).

	All robots (N = 1,080)	Physical robots (N = 375)	Software robots (N = 474)	Chatbots (N = 231)
Model fit				
χ^2	683.113	541.381	549.505	418.835
df	329	329	329	329
p	< 0.001	< 0.001	< 0.001	0.001
CFI	0.979	0.968	0.969	0.978
TLI	0.976	0.963	0.964	0.974
RMSEA	0.032	0.041	0.038	0.034
SRMR	0.031	0.035	0.037	0.038
Effects				
AUT → PI	0.212***	0.161*	0.231**	0.281***
ADA → PI	0.358***	0.335***	0.411***	0.213*
REA → PI	-0.017	0.063	-0.138*	0.053
MF → PI	0.187***	0.166**	0.185**	0.261**
ATC → PI	0.029	0.012	0.036	0.013
PER → PI	0.230***	0.268***	0.245***	0.228**
R^2 in PI	65.7%	69.1%	60.1%	72.6%

*Table 3. Model fit and model estimates (*** = p < 0.001, ** = p < 0.01, * = p < 0.05).*

Of the six individual intelligence dimensions, adaptability, personality, autonomy, and multifunctionality were each found to have a positive and statistically significant effect on overall perceived intelligence, whereas the effects of the ability to cooperate and reactivity were found to be statistically not significant. Together, the six individual intelligence dimensions were found to explain 65.7% of the variance in overall perceived intelligence.

4.3 Differences between different types of robots

In order to examine the potential differences between different types of robots, we first divided the sample into three mutually exclusive and collectively exhaustive groups based on whether a respondent had responded to the measurement items in the case of physical robots (N = 375), software robots (N = 474), or chatbots (N = 231). We then estimated the research model separately for each of these groups. The results of these estimations are reported in the remaining three columns of Table 3. As can be seen, the model fit of these three models remained approximately as good as in the case of the model estimated by using the whole sample without the group separation, but there were some differences in the model estimates. However, before examining these differences closer, we first tested the measurement invariance across the groups. The results of these tests are reported in Table 4. As can be seen, the tests supported the hypothesis on both configural and full metric invariance but only partial scalar invariance. The indicator intercepts that were not found to be invariant across the groups were those of REA4 and REA3 in the case of physical robots, which were both found to be slightly lower than in the case of software robots and chatbots. This means that the respondents tended to score these two items slightly lower in the case of physical robots than in the case of software robots and chatbots, regardless of the score of the reactivity construct that they were measuring. However, this partial scalar invariance cannot be considered to compromise the mean score comparisons across the groups in the case of any of the constructs because all the constructs were still measured by at least one indicator that had both an invariant loading and an invariant intercept across all the compared groups (cf. Steenkamp and Baumgartner, 1998).

Invariance	χ^2	df	SCF	CFI	TLI	RMSEA	SRMR	$\Delta\chi^2$	Δdf	p
Configural	1,511.593	987	1.1697	0.970	0.966	0.038	0.037	–	–	–
Full metric	1,559.758	1,029	1.1623	0.970	0.967	0.038	0.039	45.322	42	0.335
Full scalar	1,641.548	1,071	1.1562	0.968	0.966	0.038	0.040	84.481	42	< 0.001
Partial scalar (REA4)	1,621.137	1,070	1.1564	0.969	0.967	0.038	0.040	61.266	41	0.022
Partial scalar (REA3)	1,610.344	1,069	1.1564	0.970	0.968	0.038	0.040	49.068	40	0.154
Full path	1,619.192	1,081	1.1571	0.970	0.968	0.037	0.041	9.320	12	0.675

Table 4. Results of measurement invariance tests (non-invariant indicators in parenthesis).

Finally, we moved on to examine the potential differences in the construct scores and the effects between the constructs. First, in terms of the differences in the construct scores, the results of the construct mean score comparisons across the groups are reported in a tabular form in Table 5 and a graphical form in Figure 2. In terms of autonomy and adaptability, software robots and chatbots were found to have higher scores in comparison to physical robots, but no statistically significant difference was found between software robots and chatbots. In contrast, in terms of reactivity, physical robots were found to have higher scores in comparison to software robots and chatbots, but, once again, no statistically significant difference was found between software robots and chatbots. In terms of multifunctionality, software robots were found to have higher scores in comparison to physical robots and chatbots, but no statistically significant difference was found between physical robots and chatbots. In terms of the ability to cooperate, chatbots were found to have higher scores in comparison to software robots, but no statistically significant differences were found between physical and software robots or between physical robots and chatbots. In terms of personality, chatbots were found to have higher scores in comparison to physical and software robots, and also software robots were found to have higher scores in comparison to physical robots. Finally, in terms of overall perceived intelligence, software robots and chatbots were found to have higher scores in comparison to physical robots, but no statistically significant difference was found between software robots and chatbots. Second, in terms of the differences in the effects between the constructs, the last row in Table 4 reports the results of the full path invariance test. It suggested that, overall, there were no statistically significant differences across the groups in the effects between the constructs. Thus, the small differences in the model estimates that can be seen in Table 3 are only due to random variation.

Construct	**Software robots vs. physical robots**	**Chatbots vs. physical robots**	**Chatbots vs. software robots**
Autonomy (AUT)	0.387***	0.470***	0.083
Adaptability (ADA)	0.467***	0.530***	0.063
Reactivity (REA)	-0.230**	-0.216*	0.014
Multifunctionality (MF)	0.164*	-0.017	-0.181*
Ability to cooperate (ATC)	0.123	-0.084	-0.207**
Personality (PER)	0.134*	0.539***	0.405***
Perceived intelligence (PI)	0.380***	0.304**	-0.075

*Table 5. Differences in construct mean scores (*** = $p < 0.001$, ** = $p < 0.01$, * = $p < 0.05$).*

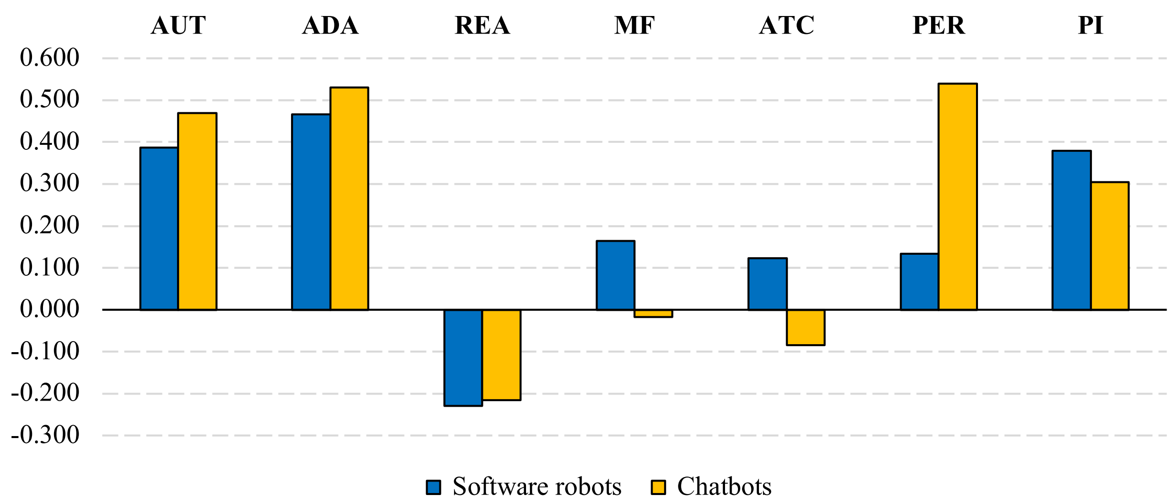

Figure 2. Differences in construct mean scores (physical robots act as the reference group).

5 Discussion and Conclusion

In this study, we examined how individual intelligence dimensions affect the overall intelligence perception of robots in the work context. Moreover, we also examined the potential differences in these effects as well as in the individual intelligence dimensions and overall intelligence perception themselves between three common types of robots: physical robots, software robots, and chatbots. All in all, we made three main findings.

First, we found that the overall perceived intelligence of robots in the work context is affected most strongly by their adaptability, followed by their personality, autonomy, and multifunctionality. Each of these four individual intelligence dimensions was found to have a positive and statistically significant effect on the overall perceived intelligence of robots. In contrast, the effects of the ability to cooperate and reactivity were found to be statistically not significant. These findings are partly in accordance but partly in conflict with the prior findings by Lee and Shin (2018) in the context of smartphone use by consumers, which suggested that only multifunctionality and adaptability have a positive and statistically significant effect on the overall perceived intelligence of smartphones, whereas the effects of reactivity, the ability to cooperate, and autonomy are statistically not significant. Thus, there seem to be some substantial differences between this context and the context of using robots at work in terms of how the individual intelligence dimensions affect the overall intelligence perception. Of the findings, especially the positive and statistically significant effects of adaptability and autonomy on the overall perceived intelligence of robots cannot be considered surprising because these were assessed as the most important individual intelligence dimensions in terms of overall intelligence perception also in the study by Rijsdijk et al. (2007) based on interviews with multiple intelligent product experts, many of whom considered these as "the ultimate intelligence". In addition, they are both central themes in artificial intelligence (AI), on which also the intelligence of intelligent robots is based (Russell and Norvig, 2022). Similarly, we do not consider the positive and statistically significant effects of multifunctionality and personality on the overall perceived intelligence of robots to be surprising. On one hand, being competent in not only one but multiple things is likely to promote the perception of general instead of only domain-specific intelligence. On the other hand, perceived intelligence has been found to correlate strongly with perceived anthropomorphism in several prior studies on the acceptance and use of intelligent systems (e.g., Moussawi et al., 2021, 2022; Moussawi and Benbunan-Fich, 2021). In contrast, the statistically not significant effects of the ability to cooperate and reactivity on the overall perceived intelligence of robots can be considered somewhat more surprising. They may both be explained by the ever-higher expectations that users place on the intelligence of robots. For example, due to the widespread diffusion of IoT technologies, the ability of a robot to connect to as well as communicate and cooperate with other systems may already be seen as a standard feature of many modern robots rather than any special sign of intelligence. Similarly, many users may expect a robot to be not only reactive but also adaptive in order for it to be perceived as intelligent. As explained by Rijsdijk et al. (2007), reactivity refers to very straightforward and reflexive responses to environmental stimuli that remain constant over time. Instead, adaptability involves more sophisticated gathering, storing, and processing of information in order to build an internal model of the environment that is then used for learning and adapting the responses to the external stimuli over time.

Second, we found no differences in the aforementioned effects between different types of robots, meaning that users seem to form their overall intelligence perception based on the individual intelligence dimensions similarly in the case of physical robots, software robots, and chatbots alike. Thus, the differences in the aforementioned effects can be characterised more as between-context differences (i.e., differences between the context of using robots at work and other contexts) than within-context differences (i.e., differences within the context of using robots at work). In contrast, third, we found differences in the individual intelligence dimensions and overall intelligence perception themselves between different types of robots. In general, both software robots and chatbots seemed to be perceived as at least equally or more intelligent than physical robots in the case of all the individual intelligence dimensions except for reactivity, in the case of which physical robots were perceived as more intelligent than both software robots and chatbots. The most substantial differences were found in autonomy, adaptability, and personality. As a result, also the overall perceived intelligence of both software robots and chatbots was found to be higher in comparison to physical robots. However, between software robots and chatbots, the differences were found to be less substantial, which is likely explained by their commonalities (e.g., being both non-physical instead of physical) in comparison to physical robots. Here, the only differences were that chatbots were perceived to be slightly less multifunctional and less able to cooperate, but to have a more humanlike personality. Of these, especially the difference in personality is by no means surprising because chatbots typically communicate with users in a very humanlike manner and may be represented by humanlike avatars, which obviously offers them more opportunities for building humanlike personalities. This is also reflected by the interest that several prior studies have shown in chatbot personalities (e.g., Shumanov and Johnson, 2021).

From a theoretical perspective, the aforementioned findings promote our understanding of the perceived intelligence of robots by suggesting how individual intelligence dimensions affect the formation of overall intelligence perception in the work context. On one hand, this understanding can be seen to offer us new insights into the antecedents of perceived intelligence. On the other hand, it can also be seen to offer us building blocks for a more precise definition of the concept of perceived intelligence itself. As exemplified by Legg and Hutter (2007), intelligence is a very ambiguous concept for which dozens of different definitions have been suggested. Many of these definitions are based on a list of abilities associated with intelligence, but this list varies widely from one definition to another. Our study suggests that in the context of intelligent robots, and more specifically in the context of using intelligent robots at work, the most precise definition for perceived intelligence would seem to be the ability of a robot to be adaptable, autonomous, and multifunctional as well as to have a humanlike personality. This more precise definition enables future studies to move on from the simple unidimensional operationalisations of perceived intelligence that have been used in most prior studies (cf. Section 2), to more complex multidimensional operationalisations that may further promote our understanding of the phenomenon in question.

In turn, from a practical perspective, the findings of the study offer the organisations that have already taken or are thinking about taking robots into use valuable insights into promoting their perceived intelligence. Such promotional actions can be considered important from at least two perspectives. On one hand, from an organisation-centred perspective, the higher perceived intelligence of robots is likely to accelerate their diffusion in the organisations, as illustrated by the positive effects of perceived intelligence on adoption, use, and use continuance intentions (cf. Section 2), thus also resulting in productivity improvements. On the other hand, from a human-centred perspective, the higher perceived intelligence of robots is likely to improve human–robot interaction among the employees of the organisations, as illustrated by the positive effects of perceived intelligence on factors like perceived usefulness, perceived ease of use, and perceived enjoyment (cf. Section 2), thus augmenting the benefits (e.g., more meaningful work – cf. Smids et al., 2020) and mitigating the hindrances (e.g., technostress or robostress – cf. Tarafdar et al., 2007; Ragu-Nathan et al., 2008; Vänni et al., 2019) that may result from human–robot collaboration. What the promotional actions should concretely be obviously depends a lot on the robot or context in question, such as whether the robot in question is an industrial robot that is working in a factory or a service robot that is working in close cooperation with humans. However, our findings suggest that, in the case of all robots and contexts, the main focus of the promotional actions should be on adaptability, personality, autonomy, and multifunctionality because these individual intelligence dimensions were found as the most influential in affecting the overall intelligence perception of robots. Thus, promoting employee perceptions on these four dimensions should be paid special attention during the whole lifespan of the robots, not only when designing and developing the robots but also when introducing them to their end-users because the acceptance and use of all technological innovations are ultimately determined not only by their internal technical characteristics but also by external social cues, such as the communication concerning them (Rogers, 2003). In addition, our findings suggest that the promotional actions are needed most urgently in the case of physical robots, which were found to lag behind software robots and chatbots in most of the individual intelligence dimensions and in the overall intelligence perception. However, in the case of all types of robots, there seems to be plenty of room for improvement, as it is suggested by the mean values of the measurement items reported in Appendix A. According to these, the respondents found robots to perform worse than moderately in all the individual intelligence dimensions except for the ability to cooperate and multifunctionality as well as particularly badly in personality.

6 Limitations and Future Research

This study can be considered to have three main limitations. First, the study targeted only US, UK, and Canadian residents, which limits the generalisability of its findings to other countries and cultures. Second, the study focused on the use of robots only in the work context. Although this context has acted as one of the main drivers for the diffusion of intelligent robots, intelligent robots are obviously also being used in non-work contexts, making it important to expand the examination to also these other contexts. Third, although the six individual intelligence dimensions in our research model were able to explain about two-thirds of the variance in the overall perceived intelligence of robots, one-third of this variance remained unexplained. Thus, in order to further promote the proportion of explained variance, it is important to expand the set of the examined individual intelligence dimensions, for example, by considering the use of also other taxonomies than the taxonomy by Rijsdijk and Hultink (2003, 2009; Rijsdijk et al., 2007) that we used in this study.

We see that one important path for future research is to address the aforementioned limitations by replicating the present study in other countries and cultures as well as in contexts other than the work context, while also potentially identifying additional individual intelligence dimensions that may affect the overall intelligence perception of robots. In addition, future research may benefit from less quantitative and more qualitative studies that provide more in-depth information on the specific use contexts of robots. Such studies may be particularly helpful in more thoroughly explaining some of the findings of this study, such as the differences found in the individual intelligence dimensions and overall intelligence perception between different types of robots. They may also offer further insights on concrete actions that could be taken to promote the individual intelligence dimensions that were found as the most influential in affecting the overall intelligence perception of robots in this study.

Acknowledgements

This study was funded by the Academy of Finland (341359).

References

Aw, E. C.-X., Tan, G. W.-H., Cham, T.-H., Raman, R., and Ooi, K.-B. (2022). "Alexa, what's on my shopping list? Transforming customer experience with digital voice assistants." *Technological Forecasting and Social Change* 180, 121711.

Balakrishnan, J., Abed, S. S., and Jones, P. (2022). "The role of meta-UTAUT factors, perceived anthropomorphism, perceived intelligence, and social self-efficacy in chatbot-based services?" *Technological Forecasting and Social Change* 180, 121692.

Balakrishnan, J. and Dwivedi, Y. K. (2021). "Conversational commerce: Entering the next stage of AI-powered digital assistants." *Annals of Operations Research*.

Bartneck, C., Kulić, D., Croft, E., and Zoghbi, S. (2009). "Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots." *International Journal of Social Robotics* 1 (1), 71–81.

Bentler, P. M. and Bonett, D. G. (1980). "Significance tests and goodness of fit in the analysis of covariance structures." *Psychological Bulletin* 88 (3), 588–606.

Blut, M., Wang, C., Wünderlich, N. V., and Brock, C. (2021). "Understanding anthropomorphism in service provision: A meta-analysis of physical robots, chatbots, and other AI." *Journal of the Academy of Marketing Science* 49 (4), 632–658.

Cai, D., Li, H., and Law, R. (2022). "Anthropomorphism and OTA chatbot adoption: A mixed methods study." *Journal of Travel & Tourism Marketing* 39 (2), 228–255.

Chuah, S. H.-W., Aw, E. C.-X., and Yee, D. (2021). "Unveiling the complexity of consumers' intention to use service robots: An fsQCA approach." *Computers in Human Behavior* 123, 106870.

Dawid, H., Decker, R., Hermann, T., Jahnke, H., Klat, W., König, R., and Stummer, C. (2017). "Management science in the era of smart consumer products: Challenges and research perspectives." *Central European Journal of Operations Research* 25 (1), 203–230.

Fornell, C. and Larcker, D. F. (1981). "Evaluating structural equation models with unobservable variables and measurement error." *Journal of Marketing Research* 18 (1), 39–50.

Gefen, D., Rigdon, E. E., and Straub, D. (2011). “Editor’s comments: An update and extension to SEM guidelines for administrative and social science research.” *MIS Quarterly* 35 (2), iii–xiv.

Guha, A., Bressgott, T., Grewal, D., Mahr, D., Wetzels, M., and Schweiger, E. (2022). “How artificiality and intelligence affect voice assistant evaluations.” *Journal of the Academy of Marketing Science*.

Hair, J. F., Jr., Black, W. C., Babin, B. J., and Anderson, R. E. (2018). *Multivariate Data Analysis*. 8th Edition. Andover, United Kingdom: Cengage.

Hu, L.-t. and Bentler, P. M. (1999). “Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives.” *Structural Equation Modeling* 6 (1), 1–55.

Jia, R., Steelman, Z. R., and Reich, B. H. (2017). “Using Mechanical Turk data in IS research: Risks, rewards, and recommendations.” *Communications of the Association for Information Systems* 41, 14, 301–318.

Kim, H., So, K. K. F., and Wirtz, J. (2022). “Service robots: Applying social exchange theory to better understand human–robot interactions.” *Tourism Management* 92, 104537.

Kim, W., Kim, N., Lyons, J. B., and Nam, C. S. (2020). “Factors affecting trust in high-vulnerability human-robot interaction contexts: A structural equation modelling approach.” *Applied Ergonomics* 85, 103056.

Lee, W.-j. and Shin, S. (2018). “Effects of product smartness on satisfaction: Focused on the perceived characteristics of smartphones.” *Journal of Theoretical and Applied Electronic Commerce Research* 13 (2), 1–14.

Legg, S. and Hutter, M. (2007). “A collection of definitions of intelligence.” In: *Advances in Artificial General Intelligence: Concepts, Architectures and Algorithms*. Ed. by B. Goertzel and P. Wang. Frontiers in Artificial Intelligence and Applications 157. Amsterdam, Netherlands: IOS Press, pp. 17–24.

Lowry, P. B., D’Arcy, J., Hammer, B., and Moody, G. D. (2016). ““Cargo Cult” science in traditional organization and information systems survey research: A case for using nontraditional methods of data collection, including Mechanical Turk and online panels.” *Journal of Strategic Information Systems* 25 (3), 232–240.

Maass, W. and Varshney, U. (2008). “Preface to the focus theme section: ‘Smart products’.” *Electronic Markets* 18 (3), 211–215.

Marikyan, D., Papagiannidis, S., Rana, O. F., Ranjan, R., and Morgan, G. (2022). ““Alexa, let’s talk about my productivity”: The impact of digital assistants on work productivity.” *Journal of Business Research* 142, 572–584.

Maroufkhani, P., Asadi, S., Ghobakhloo, M., Jannesari, M. T., and Ismail, W. K. W. (2022). ”How do interactive voice assistants build brands’ loyalty?” *Technological Forecasting and Social Change* 183, 121870.

McLean, G., Osei-Frimpong, K., and Barhorst, J. (2021). “Alexa, do voice assistants influence consumer brand engagement? – Examining the role of AI powered voice assistants in influencing consumer brand engagement.” *Journal of Business Research* 124, 312–328.

Moussawi, S. and Benbunan-Fich, R. (2021). “The effect of voice and humour on users’ perceptions of personal intelligent agents.” *Behaviour & Information Technology* 40 (15), 1603–1626.

Moussawi, S., Koufaris, M., and Benbunan-Fich, R. (2021). “How perceptions of intelligence and anthropomorphism affect adoption of personal intelligent agents.” *Electronic Markets* 31 (2), 343–364.

Moussawi, S., Koufaris, M., and Benbunan-Fich, R. (2022). “The role of user perceptions of intelligence, anthropomorphism, and self-extension on continuance of use of personal intelligent agents.” *European Journal of Information Systems*.

Muthén, L. K. and Muthén, B. O. (2022). *Mplus Home Page*. URL: https://www.statmodel.com (visited on 08/30/2022).

Novak, T. P. and Hoffman, D. L. (2019). “Relationship journeys in the internet of things: A new framework for understanding interactions between consumers and smart objects.” *Journal of the Academy of Marketing Science* 47 (2), 216–237.

Nunnally, J. C. and Bernstein, I. H. (1994). *Psychometric Theory*. 3rd Edition. New York, NY: McGraw-Hill.

OECD (2022). *Data on the Future of Work*. URL: https://www.oecd.org/future-of-work/reports-and-data/data-infographics.htm (visited on 08/30/2022).

Peer, E., Brandimarte, L., Samat, S., and Acquisti, A. (2017). "Beyond the Turk: Alternative platforms for crowdsourcing behavioral research." *Journal of Experimental Social Psychology* 70, 153–163.

Peer, E., Rothschild, D., Gordon, A., Evernden, Z., and Damer, E. (2021). "Data quality of platforms and panels for online behavioral research." *Behavior Research Methods* 54 (4), 1643–1662.

Peer, E., Vosgerau, J., and Acquisti, A. (2014). "Reputation as a sufficient condition for data quality on Amazon Mechanical Turk." *Behavior Research Methods* 46 (4), 1023–1031.

Pillai, R. and Sivathanu, B. (2020). "Adoption of AI-based chatbots for hospitality and tourism." *International Journal of Contemporary Hospitality Management* 32 (10), 3199–3226.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., and Podsakoff, N. P. (2003). "Common method biases in behavioral research: A critical review of the literature and recommended remedies." *Journal of Applied Psychology* 88 (5), 879–903.

Porter, M. E. and Heppelmann, J. E. (2014). "How smart, connected products are transforming competition." *Harvard Business Review* 92 (11), 64–88.

Qiu, H., Li, M., Shu, B., and Bai, B. (2020). "Enhancing hospitality experience with service robots: The mediating role of rapport building." *Journal of Hospitality Marketing & Management* 29 (3), 247–268.

Ragu-Nathan, T. S., Tarafdar, M., Ragu-Nathan, B. S., and Tu, Q. (2008). "The consequences of technostress for end users in organizations: Conceptual development and empirical validation." *Information Systems Research* 19 (4), 417–433.

Rijsdijk, S. A. and Hultink, E. J. (2003). ""Honey, have you seen our hamster?" Consumer evaluations of autonomous domestic products." *Journal of Product Innovation Management* 20 (3), 204–216.

Rijsdijk, S. A. and Hultink, E. J. (2009). "How today's consumers perceive tomorrow's smart products." *Journal of Product Innovation Management* 26 (1), 24–42.

Rijsdijk, S. A., Hultink, E. J., and Diamantopoulos, A. (2007). "Product intelligence: Its conceptualization, measurement and impact on consumer satisfaction." *Journal of the Academy of Marketing Science* 35 (3), 340–356.

Rogers, E. M. (2003). *Diffusion of Innovations*. 5th Edition. New York, NY: Free Press.

Rokonuzzaman, M., Kim, K. K., Dugar, K. K., and Fox, J. (2022). "What makes an object smart? Conceptualization, development, and validation of a scale to measure the Smartness of a Thing (SoT)." *Journal of Business Research* 141, 337–354.

Russell, S. J. and Norvig, P. (2022). *Artificial Intelligence: A Modern Approach*. 4th Global Edition. Harlow, England: Pearson.

Satorra, A. and Bentler, P. M. (2001). "A scaled difference chi-square test statistic for moment structure analysis." *Psychometrika* 66 (4), 507–514.

Servoz, M. (2019). *AI: The Future of Work? Work of the Future! On How Artificial Intelligence, Robotics and Automation Are Transforming Jobs and the Economy in Europe*. URL: https://data.europa.eu/doi/10.2872/913422 (visited on 08/30/2022).

Shumanov, M. and Johnson, L. (2021). "Making conversations with chatbots more personalized." *Computers in Human Behavior* 117, 106627.

Smids, J., Nyholm, S., and Berkers, H. (2020). "Robots in the workplace: A threat to – or opportunity for – meaningful work?" *Philosophy & Technology* 33 (3), 503–522.

Steenkamp, J.-B. E. M. and Baumgartner, H. (1998). "Assessing measurement invariance in cross-national consumer research." *Journal of Consumer Research* 25 (1), 78–90.

Tarafdar, M., Tu, Q., Ragu-Nathan, B. S., and Ragu-Nathan, T. S. (2007). "The impact of technostress on role stress and productivity." *Journal of Management Information Systems* 24 (1), 301–328.

Tussyadiah, I. P. and Park, S. (2018). "Consumer evaluation of hotel service robots." In: *Information and Communication Technologies in Tourism 2018*. Ed. by B. Stangl and J. Pesonen. Cham, Switzerland: Springer, pp. 308–320.

Vänni, K. J., Salin, S. E., Cabibihan, J.-J., and Kanda, T. (2019). "Robostress, a new approach to understanding robot usage, technology, and stress." In: *Proceedings of the 11th International Confer-*

ence on Social Robotics. Ed. by M. A. Salichs, S. S. Ge, E. I. Barakova, J.-J. Cabibihan, A. R. Wagner, Á. Castro-González, and H. He. Lecture Notes in Computer Science 11876. Cham, Switzerland: Springer, pp. 515–524.
World Economic Forum (2020). *The Future of Jobs Report 2020*. URL: https://www.weforum.org/reports/the-future-of-jobs-report-2020 (visited on 08/30/2022).

Appendix A

When thinking about the [physical robots / software robots / chatbots / virtual assistants] that you [use / have used] at work, to what extent do you agree or disagree with the following statements?

Item	Wording	Mean	SD	Missing	Loading
AUT1	They make decisions by themselves.	2.315	1.269	1.8%	0.826***
AUT2	They determine by themselves how to do things.	2.288	1.233	1.4%	0.834***
AUT3	They determine by themselves what to do.	2.296	1.249	1.3%	0.828***
AUT4	They decide things independently.	2.231	1.217	2.0%	0.839***
ADA1	They have an ability to learn.	2.746	1.328	1.9%	0.864***
ADA2	They learn from experience.	2.587	1.329	1.9%	0.860***
ADA3	They improve themselves over time.	2.682	1.327	1.8%	0.890***
ADA4	They deliver better and better performance over time.	3.008	1.268	0.7%	0.724***
REA1	They observe their environment.	2.457	1.300	2.6%	0.847***
REA2	They keep an eye on their environment.	2.431	1.299	2.8%	0.820***
REA3	They react to changes in the environment.	2.555	1.292	2.1%	0.837***
REA4	They adapt their behaviour to the environment.	2.517	1.284	2.1%	0.811***
MF1	They have multiple functions.	3.613	1.191	0.1%	0.841***
MF2	They perform multiple functionalities.	3.543	1.222	0.4%	0.856***
MF3	They fulfil multiple functional needs.	3.511	1.166	0.7%	0.789***
MF4	They can do a lot of different things.	3.407	1.241	0.6%	0.828***
ATC1	They can cooperate with other systems.	3.715	1.148	0.7%	0.847***
ATC2	They can work in cooperation with other systems.	3.715	1.141	0.7%	0.817***
ATC3	They can connect or be connected with other systems.	3.804	1.139	1.1%	0.796***
ATC4	They can communicate with other systems.	3.708	1.152	0.7%	0.824***
PER1	They have humanlike personalities.	1.766	1.049	2.0%	0.876***
PER2	They have human properties.	1.901	1.101	1.5%	0.808***
PER3	They behave like human beings.	1.841	1.078	1.9%	0.867***
PER4	They are like people.	1.702	0.978	1.9%	0.864***
PI1	They are intelligent.	2.781	1.396	1.3%	0.889***
PI2	They are smart.	2.891	1.369	1.4%	0.874***
PI3	They are clever.	2.734	1.380	1.5%	0.851***
PI4	They are savvy.	2.437	1.310	2.2%	0.760***

*** = $p < 0.001$