Computers in Human Behavior 89 (2018) 279–288

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Full length article

Why trust an algorithm? Performance, cognition, and neurophysiology

Veronika Alexander, Collin Blinder, Paul J. Zak*

Center for Neuroeconomics Studies, Claremont Graduate University, USA

ARTICLE INFO

Keywords:
Automation
Computers
Decisions
Neurophysiology

ABSTRACT

OBJECTIVE: We measured neurophysiologic responses and task performance while participants solved mazes after choosing whether to adopt an imperfect helper algorithm.
BACKGROUND: Every day we must decide whether to trust or distrust algorithms. Will an algorithm improve our performance on a task? What if we trust it too much?
METHOD: Participants had to pay to use the algorithm and were aware that it offered imperfect help. We varied the information about the algorithm to assess the factors that affected adoption while measuring participants' peripheral neurophysiology.
RESULTS: We found that information about previous adoption by others had a larger effect on adoption and resulted in lower cognitive load than did information about algorithm accuracy. The neurophysiologic measurement showed that algorithm adoption without any information resulted in low cognitive engagement during the task and impaired task performance. Conversely, algorithm use after information about others' use improved engagement and performance.
CONCLUSION: By objectively measuring cognitive load and task performance, we identified how to increase algorithm adoption while sustaining high performance by human operators.
APPLICATION: Algorithm adoption can be increased by sharing previous use information and performance improved by providing a reason to monitor the algorithm.
Precis: We collected neurophysiologic data while varying information about an algorithm that assisted participants in solving a timed and incentivized maze and found that information about prior use by others more effectively influenced adoption, reduced cognitive load, and improved performance compared to algorithm accuracy information.

1. Introduction

Why do people trust algorithms? Algorithms are everywhere: from self-driving cars and self-flying airplanes, to search engines and online stores recommending what you should buy next. This paper presents results of the first study, to our knowledge, measuring neurophysiologic responses when people choose to trust algorithms. We took this approach to uncover how information about algorithms affects an algorithm's adoption, neurophysiological activity and task performance. Neurophysiologic measurement allowed us to gauge participant engagement during the task when advised by an algorithm and without algorithm help. The recent fatal crash of a Tesla automobile using a driver-assist algorithm may have occurred in part because of over-reliance on the software by the driver.

Humans are naturally prone to error, and errors can be useful when it comes to creativity and innovation (Schultz, 2011). At the same time, human error contributes to 70% of aviation accidents (Shappell & Wiegmann, 2012); to 75% of roadway crashes (Stanton & Salmon, 2009); and is a major cause of preventable death in hospitals (Makary & Daniel, 2016). Algorithms outperform human judgement in various domains, including some aspects of healthcare (e.g., diagnosing medical conditions) (Gladwell, 2007; Grove, Zald, Lebow, Snitz, & Nelson, 2000), finance (e.g., automated financial forecasts) (Dietvorst, Simmons, & Massey, 2015; Silver, 2012), the military (e.g., automated change detection in war zones) (Parasuraman, Cosenzo, & De Visser, 2009), and education (e.g., academic success of students) (Dietvorst et al., 2015; Xing, Guo, Petakovic, & Goggins, 2015). Yet even when algorithms are evidence-based and improve outcomes, many people shun their use (Dietvorst et al., 2015). People can be forgiven for making the occasional mistake, but automated systems are expected to function perfectly every time (Alvarado-Valencia & Barrero, 2014). As a result, when people see an algorithm falter, they are less likely to keep using it, even when the machine outperforms humans on average (Dietvorst et al., 2015).

* Corresponding author. Center for Neuroeconomics Studies, Claremont Graduate University, 160 E. 10th St., Claremont, CA 91711. USA.
E-mail address: paul.zak@cgu.edu (P.J. Zak).

https://doi.org/10.1016/j.chb.2018.07.026
Received 17 November 2017; Received in revised form 18 July 2018; Accepted 21 July 2018
Available online 25 July 2018
0747-5632/

The increasing dependence on algorithms has many sources. These include the high cognitive load of complex work (Alvarado-Valencia & Barrero, 2014; Goddard, Roudsari, & Wyatt, 2012), and a desire for system reliability (Alvarado-Valencia & Barrero, 2014; Pak, Rovira, McLaughlin, & Baldwin, 2016). For example, algorithmic financial forecasts are less frequently adjusted when they came from a well-known source and were based on sound explanations and assumptions (Gönül, Önkal, & Goodwin, 2009). At the same time, people tend to assign more credibility to human experts rather than computerized advice (Önkal, Goodwin, Thomson, Gönül, & Pollock, 2009; Alvarado bibValencia & Barrero"2009; Alvarado-Valencia & Barrero, 2014; Lerch, Prietula, & Kulik, 1997; Waern & Ramberg, 1996). One reason for the conflict between using human advice versus algorithmic help is that people are thought to integrate more information from more sources than do automated systems, including social information, but the quality of social information may be poor (Sundar, Xu, & Oeldorf-Hirsch, 2009).

An additional concern when using automated systems is inattention. Automation complacency could be catastrophic if one fails to properly evaluate and monitor system performance (Parasuraman & Manzey, 2010). Yet, monitoring automated systems often demands high cognitive load–something that automation seeks to alleviate (Mosier & Skitka, 1996; Parasuraman & Manzey, 2010). Other reasons for absent or haphazard algorithm monitoring include a simplistic belief that an automated solution must work perfectly or nearly perfectly (Alvarado-Valencia & Barrero, 2014).

We focus on "trust" in an algorithm rather than "use" because we are interested in not only the choice to adopt an algorithm, but how people respond while using it. Our study design informs participants that algorithms are imperfect. There are multiple points at which participants can choose to follow the advice of the algorithm ("trust" it) or ignore it. Our point of departure is the research we have done on neurocorrelates of trust between unacquainted human beings (Zak, 1994, 2004; Kosfeld, Heinrichs, Zak, Fischbacher, & Fehr, 2005; Zak, 2008; Zak, 2017); we hypothesized that related neurophysiologic mechanisms may mediate trust in algorithms. In the present paper, trust is operationalized as a behavior, rather than an intention, and we associate neurophysiologic responses with whether advice from the algorithm was acted upon as trustworthy or not.

Trust in algorithms may depend on a person's dispositional trust. Since dispositional trust varies with culture, age, gender, and other personality traits (Hoff & Bashir, 2015), these factors may influence algorithm adoption. Cultural factors such as politeness have been shown to affect how people interact with automation and technology (Nass, Moon, & Carney, 1999) and if they view automated solutions as being as trustworthy as a human partner (DeSteno et al., 2012). For example, trust in an online social interaction activate similar neurophysiological mechanisms when trusting someone in person (Riedl, Hubert, & Kenning, 2010). Several groups have reported that older adults trust automation more than younger ones (Hoff & Bashir, 2015; McBride, Rogers, & Fisk, 2011) and may be less vigilant when assessing a system's properties (Ho, Kiff, Plocher, & Haigh, 2005). Many studies find that women are less trusting than men, yet once trusted are more trustworthy compared to males (Buchan, Croson, & Solnick, 2008; Zak, 2012). Yet, gender has mixed effects on trust in technology (Hoff & Bashir, 2015). Intriguingly, extroverts may use more automation than introverts (Merritt & Ilgen, 2008). Many factors affect the decision to use an automated system but the studies to date have mostly focused on traits that promote or inhibit use rather than the neurophysiologic drivers of these decisions.

The present study measured neurophysiologic responses during an algorithm adoption decision as the information about the algorithm varied. The algorithm adoption decision was designed to be difficult for two reasons. The first was that participants were aware that the algorithm we wrote offered advice that was imperfect and as a result they could follow or ignore advice throughout the task they were assigned to complete. Second, participants had to pay for the algorithm to help them earn money during the tasks in the experiment. This design sought to make the algorithm choice nontrivial. We also monitored cardiac activity during the algorithm use to assess attention. In addition, the experiment included a behavioral task that allowed us to quantify performance. We sought to connect algorithm adoption to attention and task performance in order to test if participants exhibited over-reliance or under-reliance on the algorithm. If algorithm use was not incentivized, or the algorithm provided perfect assistance, this would not capture the key issue we sought to understand, whether and why one trusts an algorithm.

We hypothesized that higher algorithm accuracy and greater prior use by others would result in higher algorithm adoption. We also expected that social proof would influence adoption more than algorithm accuracy information. Social proof (using information about actions of others) serves as an important signal of trust (MacCoun, 2012) and has been connected to purchasing behaviors, especially when buying a new product (Amblee & Bui, 2011; Hajli, 2015; Hajli, Lin, Featherman, & Wang, 2014, pp. 673–689).

From a neurophysiologic perspective, we anticipated that those with high basal cognitive load would also be more likely to adopt algorithms to conserve neural resources. Finally, we expected participants who enjoy computers and those who are less risk-averse would be more likely to adopt algorithms.

2. Methods

2.1. Participants

A total of 46 participants (52% women, M_{age} = 23.7 years, SD_{age} = 6.4) from the Claremont Colleges and surrounding community volunteered for this study. Caucasians made up 41% of the participants, with the remaining being 33% Asian, 9% Hispanic, 9% African-American, and 8% other. The Institutional Review Board of Claremont Graduate University approved the study, and written informed consent was obtained from all participants prior to inclusion.

2.2. Research design

Participants were instructed in, and then asked to solve, 12 standard two-dimensional mazes on a laptop computer (i.e., mazes that do not loop back on themselves, see Fig. 1). This produced 552 (46*12) observations that produces a power of tests of 0.99 using G*power (Faul, Erdfelder, Lang, & Buchner, 2007) and the average size effect and standard errors in Barraza, Alexander, Beavin, Terris, and Zak (2015). Participants had 60 s to solve each maze using the arrow keys to navigate. Two blocks of six mazes were presented in a counterbalanced design. Mazes were initially covered in grey "fog" so participants could not see the entire maze. All participants completed two practice mazes during a training session before data collection began. After trying to

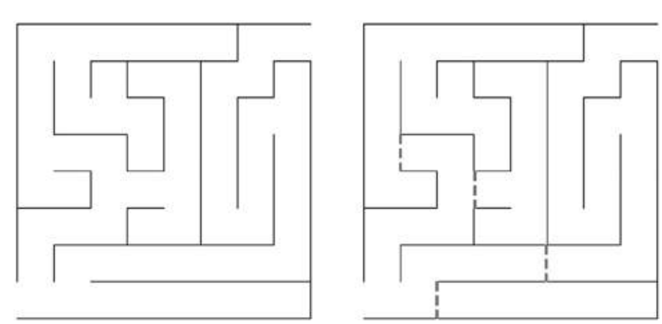

Fig. 1. An example of the maze without (left) and with (right) the algorithm: note it has four dashed "walls" added in, with one blocking the right path (which was considered as a "mistake" of the algorithm).

solve each maze, there was a 1-min quiet period to allow neurophysiologic measures to return to baseline. After the first six-maze block, participants had a 3-min rest period to reduce cognitive load before a new treatment began.

Participants were informed that they would earn \$5 for each maze they solved in 60 s or less. Participants earned \$2 for mazes they could not complete in the allotted time so they were fairly compensated, though this was not disclosed to them in advance. Before starting each maze, participants were given an opportunity to purchase a "helper" algorithm for \$2 that could make solving the maze easier. They were told that the algorithm was imperfect because each maze was generated differently, but that it would offer advice that could be followed or ignored. There were four different algorithm information treatments (see below) that remained constant within a 6-maze block.

Fig. 1 shows an example of a maze, with and without the "helper" algorithm suggestions. When the helper algorithm was purchased, mazes had several extra blue "walls" added that recommended the correct path to the maze's end. Participants were informed they could cross the walls (i.e. ignore the advice) if they wished. Most of the blue walls covered incorrect paths. We designed imperfection into the algorithms by, in one-half of the mazes, adding four blue walls with one of the four blocking the correct path. In the other half of the mazes, three blue walls were added in which one wall failed to cover a wrong path.

2.3. Experimental conditions

Participants were randomly assigned to be in two of four conditions that varied information about the algorithm. After providing information about the algorithm but before each maze was shown, the software prompted participants with "Would you like to buy this algorithm for \$2?" The first condition was the "No information" condition (NI) in which participants were offered the chance to purchase a helper algorithm but no information was given about it. The instructions read, "We have no information about this algorithm." The second condition provided participants with accuracy information (AI) about the algorithm, stating that "This algorithm is 75 percent accurate." The third condition provided low social proof (LS) regarding previous use, stating that "54 percent of people have used this algorithm." The final condition provided high social proof (HS), saying "70 percent of people have used this algorithm." Each condition remained constant during the six-maze blocks.

Block 2 always presented a different algorithm to purchase than in Block 1. We used the following four orderings to test our hypotheses (Fig. 3): (1) *NI* in Block 1 and *AI* in Block 2; (2) *AI* in Block 1 and *HS* in Block 2; (3) *NI* in Block 1 and *LS* in Block 2; and (4) *LS* in Block 1 and *NI* in Block 2.

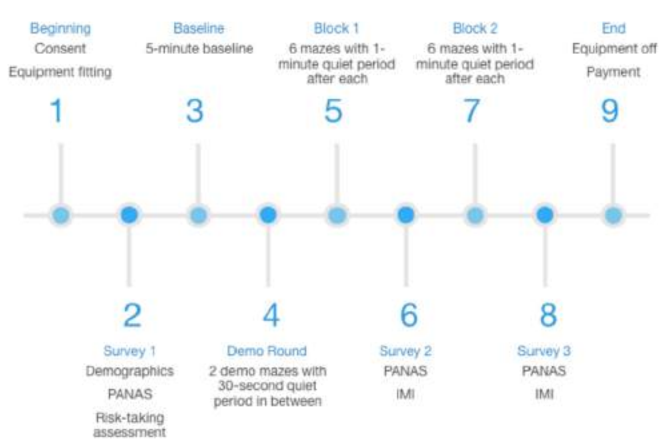

Fig. 2. The timeline of the experiment. PANAS = Positive and Negative Affect Schedule; IRI = Interpersonal Reactivity Inventory; IMI = Intrinsic Motivation Inventory.

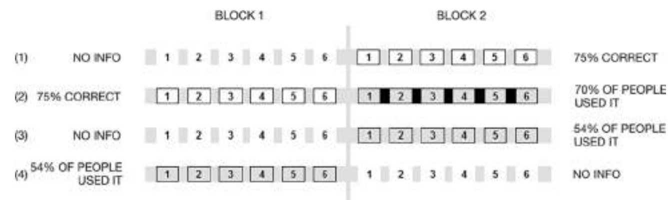

Fig. 3. Summary of experimental conditions and their orderings in each block of 6 mazes. No info = *No Information* condition (NI); 75% correct = *Accuracy Information* condition (AI); 54% of people used it = *Low Social Proof* condition (LS); 70% of people used it = *High Social Proof* condition (HS).

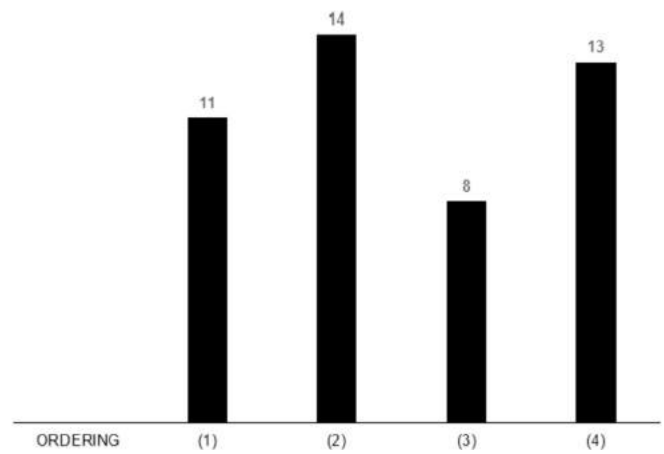

Fig. 4. Number of people in each of the orderings. (1) *No Information (NI)* in Block 1 and *Accuracy Information (AI)* in Block 2; (2) *Accuracy Information (AI)* in Block 1 and *High Social Proof (HS)* in Block 2; (3) *No Information (NI* in Block 1 and *Low Social Proof (LS)* in Block 2; and (4) *Low Social Proof (LS)* in Block 1 and *No Information (NI)* in Block 2.

Out of the total of 46 participants, 11 were assigned to the first ordering, 14 to the second, 8 to the third, and 13 to the fourth (Fig. 4). In total, 32 people were in the *NI*, 23 in the *AI*, 20 in *LS*, and 14 in *HS* conditions (Fig. 5). Note that 3 participants did not complete Block 2; for these participants only their Block 1 data was included for analyses. Important for our statistical tests, the mazes themselves and extra walls due to algorithm purchase were unchanged across conditions. This allowed us to isolate how the information about the algorithms affected the purchase decisions.

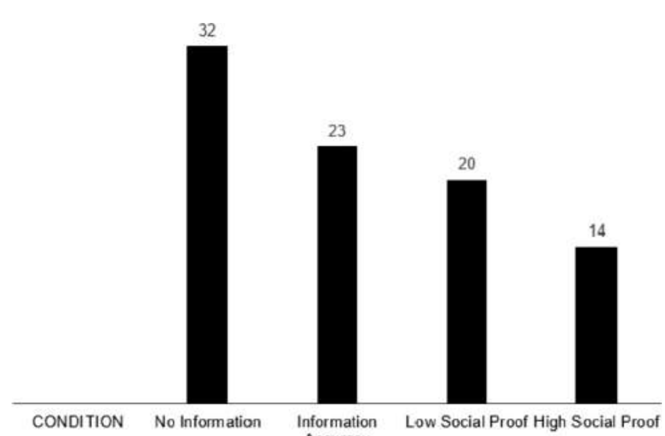

Fig. 5. Number of people in each condition.

2.4. Neurophysiologic measures

Cognitive load (the amount of mental effort needed to solve a problem; Paas & Van Merriënboer, 1993; Sweller, 1988) can be measured using electrocardiography by increased heart rate and decreased heart rate variability (Brookhuis & de Waard, 2010). In the present study, we used changes in heart rate (HR) from baseline to measure cognitive load/arousal during the task. An increase in HR from baseline is associated with enhanced prefrontal activity and, over a range of non-anxious responses, is associated with improved cognitive performance (Fredericks, Choi, Hart, Butt, & Mital, 2005; Lahtinen, Koskelo, Laitinen, & Leino, 2007; Thayer, Hansen, Saus-Rose, & Johnsen, 2009). We also used the standard deviation of HR to approximate high-frequency heart rate variability (HF-HRV). Heart rate variability is defined as the variation in the time intervals between adjacent heartbeats (Sztajzel, 2004). HF-HRV is correlated with vagal tone and indicates increased parasympathetic activity.

2.5. Procedure

Fig. 2 shows the timeline of the study. After consent, participants were fitted with electrocardiogram (ECG) electrodes on the left and right costal arch at the level of the 10th rib, with reference below the right collarbone. These wirelessly sent cardiac data to a BIOPAC® MP150 data acquisition system (BIOPAC Systems Inc., Goleta, California). ECG data were obtained continuously during the study.

Participants were then asked to complete a set of surveys, including demographics, mood (Positive and Negative Affect Schedule, PANAS; Watson, Clark, & Tellegen, 1988), dispositional empathy (Interpersonal Reactivity Index, IRI; Davis, 1983), and risk preferences (Risk Taking Preference Questionnaire, adapted from Holt & Laury, 2002). After completing the surveys, participants were instructed to sit quietly for 5 min in order to obtain baseline ECG measures. Immediately afterwards, participants were given two practice mazes to complete in 60 s, with a 30-s break between them.

After the practice session, participants were given a 3-min break after which data collection began as participants completed two sets of 6 mazes. After the first block of mazes, participants completed surveys assessing mood (PANAS) and intrinsic motivation (Intrinsic Motivation Inventory, IMI; adapted from McAuley, Duncan, & Tammen, 1989). Then participants were given another set of six 60-s mazes (Block 2). The setup was identical to that of Block 1, except the information describing the "helper" algorithm was different (see conditions below). Participants then completed a final survey set (PANAS and IMI). A research assistant then helped participants remove the sensors and paid them in private. Participants completed the experiment in about an hour.

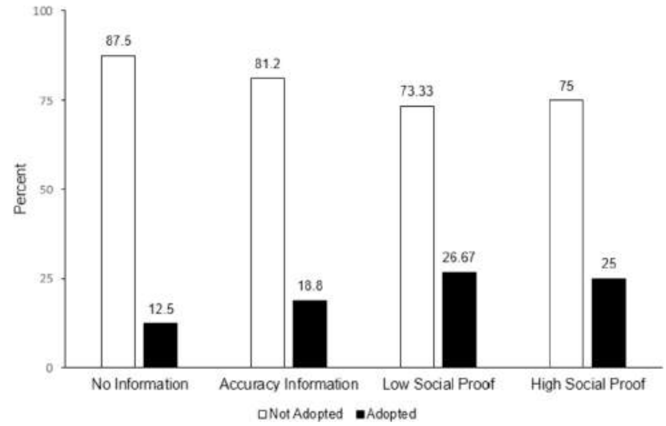

Fig. 6. Algorithm adoption by condition.

Table 1
Mixed-effects multilevel logistic regression results on decision to adopt an algorithm.

Variables	Odds Ratio	*p*-value	95% Confidence Interval	
			Lower	Upper
Fixed effects				
Maze	1.07 (0.09)	.40	0.91	1.25
Block #2	0.16** (0.10)	< .01	0.05	0.54
HR %	0.91** (0.03)	< .01	0.86	0.96
Accuracy Information Condition	1.46 (0.76)	.46	0.53	4.04
Social Condition	4.50** (1.99)	< .01	1.90	10.69
Gender (female)	3.53* (2.26)	.05	1.00	12.39
Finished the maze	1.20 (0.36)	.55	0.67	2.16
HF-HRV	1.16 (0.09)	.07	0.99	1.36
Constant	0.02** (0.02)	< .01	0.00	0.10
Random effects				
	Estimate	*p*-value	95% Confidence Interval Lower	Upper
Participant's ID (σ of intercept)	1.72** (0.36)	< .01	1.14	2.58
Observations	523			

Note. Standard errors are in parentheses. HR = Heart Rate; HF-HRV = High Frequency Heart Rate Variability.
*$p < .05$, **$p < .01$.

3. Results

3.1. Behavior

Sixty-one percent of participants chose the helper algorithm at least one time. In the NI condition, the algorithm was used 12.5% of the time, followed by 18.8% of the time in the AI condition, and 26.67% and 25% in LS and HS conditions (Fig. 6). There is a significant effect of algorithm condition on algorithm adoption ($F(3, 525) = 3.95, p < .01$, $\eta_p^2 = 0.022$). The adoption rate for the NI condition is not significantly different from that of the AI condition ($t(323) = -1.56, p > .1$), but is significantly less than both social proof conditions (NI vs. LS: $t(310) = -3.21$, $p < .01$, $d = 0.37$; NI vs. HS: $t(274) = -2.61$, $p < .01$, $d = 0.34$). The adoption rate was not statistically significant for the LS and HS conditions ($t(202) = 0.27$, $p > .1$), so we will combine them into a single condition we will call S.

We estimated a mixed-effects multilevel logistic regression (Table 1) and found that the maze order did not affect decisions to purchase an algorithm (Odds Ratio (OR) was not significantly different from 1 (95% Confidence Interval (CI) = [0.91, 1.25]), $p > .1$). As a result, each choice was considered an independent observation. We found that algorithm purchase was 42% lower in Block 2 than in Block 1 (95% CI for OR = [0.05, 0.54], $p < .01$). Having accuracy information (AI), compared to having no information (NI), did not affect the decision to use an algorithm (95% CI for OR = [0.53, 4.04], $p > .1$). At the same time, being in either of the social conditions (LS or HS) resulted in a greater use of an algorithm (95% CI for OR = [1.90, 10.69], $p < .01$) consistent with our finding above.

The proportion of mazes completed was unrelated to algorithm use (95% CI for OR = [0.67, 2.16], $p > .1$) as was expected from an experimental design that used imperfect algorithms. By comparing NI and S conditions, however, we found differing effects on performance depending on algorithm type. In the NI condition, the completion rate for those who did not purchase the algorithm was marginally *higher* than that of algorithm-adopters (54% of non-users finished the maze in time, compared to 33% for algorithm-users, $t(190) = 1.92$, $p = .06$, $d = 0.42$). Conversely, in the combined S condition, the performance of non-adopters was marginally *lower* than that of adopters (51% finished the maze in time without an algorithm, compared to 66% with an algorithm, $t(202) = -1.90$, $p = .06$, $d = 0.30$). There was no difference in performance for the AI condition for algorithm-users compared to non-users ($t(131) = 0.81$, $p > .10$).

3.2. Neurophysiological responses

Algorithm use was associated with a decrease in heart rate, $t(521) = 3.06$, $p < .01$, $d = 0.34$. For non-users, heart rate (HR) increased 3.9% on average from baseline to the maze-solving period ($t(423) = 8.29$, $p < .01$, $d = 0.40$), while users' HRs did not significantly change ($t(99) = 0.69$, $p > .10$). Higher baseline HR did not influence algorithm adoption ($r = 0.23$, $p > .10$). For NI condition, percentage change in HR was significantly higher for non-users compared to that of users ($t(41.47) = 4.07$, $p < .01$, $d = 0.61$): HR increased by 3.11% for non-users ($t(167) = 4.62$, $p < .01$, $d = 0.36$), but did not significantly change for users ($t(22) = -1.91$, $p = .07$)). For the AI condition, HR increased by 3.60% for non-users ($t(103) = 3.62$, $p < .01$, $d = 0.35$), but did not significantly change for users ($t(23) = 0.74$, $p > .1$). Finally, for S condition, percentage change in HR was significantly higher for non-users, compared to that of users ($t(202) = 2.19$, $p = .03$, $d = 0.35$): HR increased by 4.85% for non-users ($t(150) = 5.95$, $p < .01$, $d = 0.48$), but did not significantly change for users ($t(52) = 1.02$, $p > .1$).

We examined second-by-second cardiac activity during the maze completion by combining the two information conditions NI and AI that have identical adoption rates ($t(323) = -1.56$, $p > .10$) into one condition we call O. We then compared this to the combined social condition (S). The adoption rate in S was 72% higher than that in O ($t(367.61) = -2.98$, $p < .01$, $d = 0.28$). Fig. 7 depicts the percentage change in heart rate from baseline for the S and O conditions over the 65 s participants spent solving the maze. The average change in HR is higher for S than for O while participants are completing the maze for seconds 20 to 60. We also found that HF-HRV in S was 49.6% higher compared to that in O ($F(64, 64) = 0.85$, $p < .01$).

We next examined the HR time series for completed mazes using helper algorithms compared to completed mazes without help. Fig. 8 shows that those who purchased the algorithm in the NI condition had heart rates that were, on average, 2% below baseline throughout the 60 s maze period. Standard statistical tests require that the time series is stationary (i.e., not a random walk) and we confirmed this (Dickey-Fuller test $p < .01$). Those who solved the maze without the helper algorithm experienced greater arousal, with heart rate increasing linearly during the task ($p < .01$). There was no significant difference in HF-HRV between algorithm-users and non-users in the NI condition ($p > .10$). Time series for both AI and S conditions are shown in Figs. 9 and 10. We also found that those who used the algorithms had significantly higher HF-HRV in the AI and S conditions, compared to non-adopters ($F(64, 64) = 0.35$, $p < .01$ for AI and $F(64, 64) = 0.59$, $p = .04$ for S).

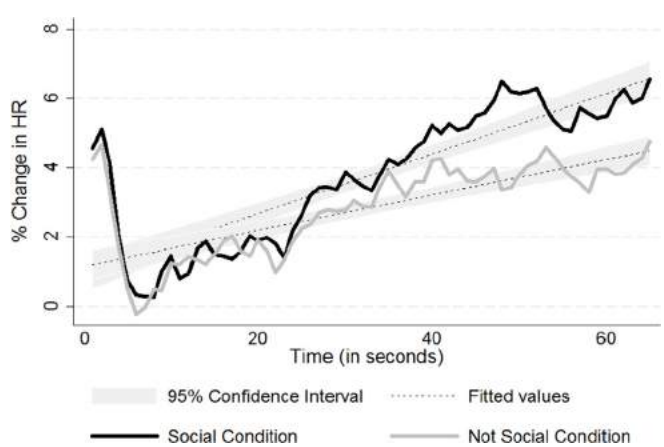

Fig. 7. % Change in heart rate (HR) for *Social* vs. *Not-Social* conditions.

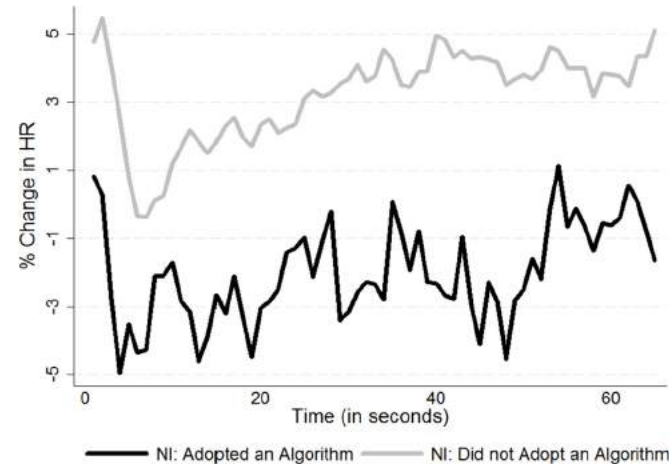

Fig. 8. % Change in heart rate (HR) for *No Information (NI)* condition: Adopters vs. Non-adopters.

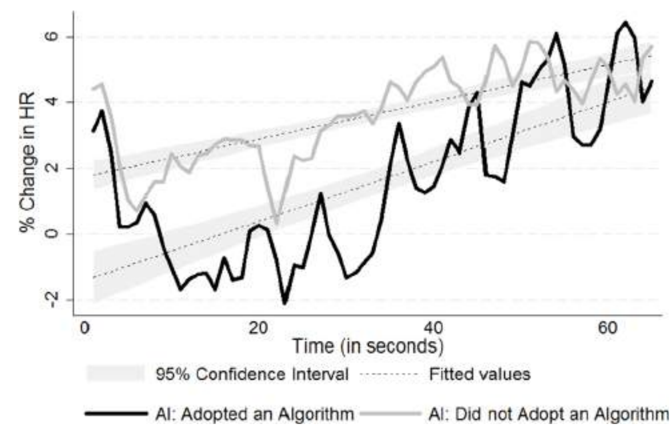

Fig. 9. % Change in heart rate (HR) for *Accuracy Information (AI)* condition: Adopters vs. Non-adopters.

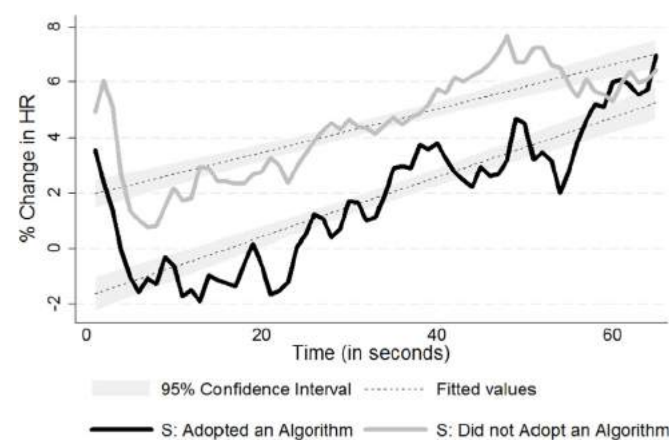

Fig. 10. % Change in heart rate (HR) for *Social* condition: Adopters vs. Non-adopters.

Table 2
Pairwise correlations of the major study variables (p-values in parentheses).

	Variable	(1)	(2)	(3)	(4)	(5)	(6)
(1)	% of the time algorithm was used in Block 1	1.0000					
		–					
(2)	% of the time algorithm was used in Block 2	**.4499****	1.0000				
		(.0025)	–				
(3)	Age	−.2115	.0459	1.0000			
		(.1582)	(.7703)	–			
(4)	Gender (0 = Male; 1 = Female)	**.3160***	02312	−.1321	1.0000		
		(.0324)	(.1358)	(.3813)	–		
(5)	Do you like computer science?	.2513	.0927	.0799	−.2152	1.0000	
		(.0921)	(.5546)	(.5976)	(.1509)	–	
(6)	Most people can be trusted	**.3522***	.1431	−.0453	.1260	.1023	1.0000
		(.0164)	(.3600)	(.7650)	(.4042)	(.4988)	–
(7)	Degree of risk-aversion	.2401	−.0023	−.0295	.1151	.0844	.0270
		(.1080)	(.9884)	(.8456)	(.4462)	(.5769)	(.8584)
(8)	% of mazes completed in Block 1	−**.3111***	.0544	.0162	−.0722	.0422	−.0875
		(.0354)	(.7290)	(.9150)	(.6335)	(.7804)	(.5633)
(9)	% of mazes completed in Block 2	−.1143	−.0044	−.1507	−**.4188****	**.3411***	−.0517
		(.4655)	(.9777)	(.3346)	(.0052)	(.0252)	(.7419)
(10)	Is your major math related (computer science, economics, math, engineering)?	−.0241	−.0016	−.0481	−.2179	.1510	−.1823
		(.8738)	(.9917)	(.7507)	(.1457)	(.3166)	(.2252)
(11)	Interpersonal Reactivity Inventory: Empathic Concern	.1903	.1050	.2015	.2898	**.3264***	−.1399
		(.2052)	(.5029)	(.1793)	(.0507)	(.0269)	(.3538)
(12)	Interpersonal Reactivity Inventory: Personal Distress	.0533	.1458	−.0335	**.3488***	−.1892	−.1299
		(.7252)	(.3509)	(.8249)	(.0175)	(.2080)	(.3895)
(13)	Baseline positive mood	.2115	.1054	.2522	.0718	**.4841****	.0046
		(.1583)	(.5011)	(.0909)	(.6355)	(.0007)	(.9757)
(14)	Baseline negative mood	−.0273	−.0757	−.2289	.1147	−.2415	−.2571
		(.8569)	(.6293)	(.1260)	(.4477)	(.1059)	(.0845)
(15)	Intrinsic Motivation Inventory (IMI) after Block 1: Interest/Enjoyment	−.0763	.1134	.2723	.0273	.2185	−.1515
		(.6142)	(.4691)	(.0671)	(.8569)	(.1446)	(.3147)
(16)	IMI after Block 1: Competence	−.0998	.0703	.1682	−.1144	**.3968****	.0295
		(.5093)	(.6544)	(.2638)	(.4492)	(.0063)	(.8457)
(17)	IMI after Block 1: Effort	−.0094	.0470	.0714	.0303	−.1290	.1796
		(.9508)	(.7646)	(.6374)	(.8414)	(.3928)	(.2325)
(18)	IMI after Block 1: Feeling tense or pressured	.0699	−.1270	−.2022	.2578	−.1513	.0960
		(.6444)	(.4172)	(.1778)	(.0837)	(.3156)	(.5256)
(19)	Positive mood after Block 1	−.0792	.0941	**.3476***	−.0477	**.4067****	−.0760
		(.6010)	(.5482)	(.0179)	(.7531)	(.0050)	(.6157)
(20)	Negative mood after Block 1	.1229	−.0812	−.2231	.1601	−.1418	.1243
		(.4158)	(.6046)	(.1362)	(.2878)	(.3474)	(.4106)
(21)	IMI after Block 2: Interest/Enjoyment	.1115	.0913	.1838	−.0895	**.3166***	−.0796
		(.4822)	(.5653)	(.2439)	(.5728)	(.0411)	(.6162)
(22)	Intrinsic Motivation Inventory (IMI) after Block 2: Competence	.0604	.1223	.0055	−**.3055***	**.4536****	−.0307
		(.7039)	(.4404)	(.9723)	(.0492)	(.0026)	(.8469)
(23)	IMI after Block 2: Effort	.2734	.1447	−.0720	.0328	−.0242	.1728
		(.0797)	(.3606)	(.6503)	(.8368)	(.8793)	(.2738)
(24)	IMI after Block 2: Feeling tense or pressured	.1113	−.1727	−.2877	.2461	−.2002	.2303
		(.4830)	(.2742)	(.0647)	(.1161)	(.2036)	(.1424)
(25)	Positive mood after Block 2	.1725	.1233	.2053	−.0857	**.3828***	−.0367
		(.2746)	(.4368)	(.1922)	(.5895)	(.0123)	(.8177)
(26)	Negative mood after Block 2	.1365	.0468	−.1898	.2058	−.0221	−.0092
		(.3886)	(.7685)	(.2286)	(.1910)	(.8896)	(.9541)
(27)	Baseline Heart Rate (HR)	.1559	.2653	.1377	.2696	−.2245	.0143
		(.3008)	(.0855)	(.3613)	(.0700)	(.1335)	(.9248)
(28)	% change in HR from baseline to Block 1	−.1638	−.1558	−.0078	−.1826	.1182	−.0724
		(.2768)	(.3185)	(.9591)	(.2245)	(.4339)	(.6323)
(29)	% change in HR from baseline to Block 2	−.0367	−.1325	−.1606	−.1250	.2141	−.0078
		(.8154)	(.4027)	(.3037)	(.4245)	(.1681)	(.9605)
(30)	High Frequency Heart Rate Variability (HF-HRV): Baseline	.0604	.2124	−.1892	−.0116	−.1143	.1009
		(.6901)	(.1714)	(.2079)	(.9390)	(.4493)	(.5046)
(31)	HF-HRV: Block 1	−.0311	.2089	−**.3426***	−.1152	−.1878	−.0375
		(.8375)	(.1788)	(.0198)	(.4457)	(.2113)	(.8046)
(32)	HF-HRV: Block 2	−.0799	.2335	−**.3151***	−.1915	−.1878	−.0203
		(.6063)	(.1318)	(.0372)	(.2131)	(.2221)	(.8960)

	Variable	(7)	(8)	(9)	(10)	(11)	(12)
(7)	Degree of risk-aversion	1					
		–					
(8)	% of mazes completed in Block 1	−.0458	1.0000				
		(.7627)	–				
(9)	% of mazes completed in Block 2	.1004	.0870	1.0000			
		(.5218)	(.5791)	–			

(*continued on next page*)

Table 2 (*continued*)

(10)	Is your major math related (computer science, economics, math, engineering)?	−.1522	.1756	−.0531	1.0000		
		(.3125)	(.2432)	(.7351)	–		
(11)	Interpersonal Reactivity Inventory: Empathic Concern	.1929	.0962	−.0145	−.1661	1.0000	
		(.1989)	(.5250)	(.9265)	(.2698)	–	
(12)	Interpersonal Reactivity Inventory: Personal Distress	.2538	.0563	−.2794	.0139	.2822	1.0000
		(.0887)	(.7102)	(.0696)	(.9271)	(.0575)	–
(13)	Baseline positive mood	.1921	−.0884	.2234	−.0276	**.4021****	.1100
		(.2009)	(.5592)	(.1499)	(.8554)	(.0056)	(.4666)
(14)	Baseline negative mood	.1109	.0747	−.0085	−.2567	.2103	**.4128****
		(.4630)	(.6216)	(.9570)	(.0850)	(.1607)	(.0044)
(15)	Intrinsic Motivation Inventory (IMI) after Block 1: Interest/Enjoyment	.2369	.1677	.0536	−.2415	.2250	−.0672
		(.1130)	(.2652)	(.7326)	(.1059)	(.1327)	(.6572)
(16)	IMI after Block 1: Competence	.2832	**.5403****	**.4023****	−.0921	**.3022***	−.1519
		(.0565)	(.0001)	(.0075)	(.5427)	(.0412)	(.3136)
(17)	IMI after Block 1: Effort	−.1549	−.0945	−.2545	.1296	−.0900	.1434
		(.3041)	(.5321)	(.0995)	(.3905)	(.5519)	(.3416)
(18)	Intrinsic Motivation Inventory (IMI) after Block 1: Feeling tense or pressured	−.0515	−.2083	−.2943	.0219	.0728	.2541
		(.7339)	(.1647)	(.0554)	(.8854)	(.6306)	(.0884)
(19)	Positive mood after Block 1	.0867	.1904	.1850	.0897	.2798	.0503
		(.5667)	(.2050)	(.2350)	(.5532)	(.0597)	(.7398)
(20)	Negative mood after Block 1	.2280	.0428	−.1059	−.2201	.2877	**.5082****
		(.1275)	(.7777)	(.4991)	(.1415)	(.0525)	(.0003)
(21)	IMI after Block 2: Interest/Enjoyment	.2118	−.2261	.2146	−.2537	.1564	−.1280
		(.1780)	(.1499)	(.1723)	(.1050)	(.3227)	(.4190)
(22)	IMI after Block 2: Competence	**.3924***	.1036	**.7315****	−.1753	.2689	−.0856
		(.0102)	(.5139)	(.0000)	(.2668)	(.0851)	(.5897)
(23)	IMI after Block 2: Effort	−.0762	**−.3067***	−.1589	.2036	−.0336	.2258
		(.6316)	(.0482)	(.3148)	(.1959)	(.8325)	(.1505)
(24)	IMI after Block 2: Feeling tense or pressured	−.0003	−.0573	**−.3851***	−.1427	.1714	**.3510***
		(.9983)	(.7187)	(.0118)	(.3675)	(.2779)	(.0226)
(25)	Positive mood after Block 2	.1100	−.0911	.2731	.0659	.1241	−.0487
		(.4882)	(.5660)	(.0802)	(.6782)	(.4337)	(.7592)
(26)	Negative mood after Block 2	.0800	.1163	−.1737	−.1006	**.4188****	**.3444***
		(.6145)	(.4634)	(.2713)	(.5261)	(.0058)	(.0255)
(27)	Baseline Heart Rate (HR)	−.1352	.0779	−.2514	−.0095	.2397	.0825
		(.3703)	(.6068)	(.1040)	(.9501)	(.1086)	(.5855)
(28)	% change in HR from baseline to Block 1	.1905	.0980	−.1086	.1559	−.0017	.1799
		(.2046)	(.5169)	(.4883)	(.3008)	(.9910)	(.2315)
(29)	% change in Heart Rate from baseline to Block 2	.2329	.0441	−.0583	.2743	.0490	.0966
		(.1328)	(.7788)	(.7137)	(.0750)	(.7549)	(.5376)
(30)	High Frequency Heart Rate Variability (HF-HRV): Baseline	−.1628	**.3304***	.0336	**.3458***	−.1912	.0280
		(.2797)	(.0249)	(.8308)	(.0186)	(.2031)	(.8533)
(31)	HF-HRV: Block 1	−.1484	**.3707***	−.0200	**.3005***	−.2269	.2392
		(.3251)	(.0112)	(.8987)	(.0424)	(.1295)	(.1093)
(32)	HF-HRV: Block 2	−.2098	**.3620***	.1852	.1759	−.2226	.1399
		(.1716)	(.0158)	(.2344)	(.2535)	(.1464)	(.3650)

	Variable	(13)	(14)	(15)	(16)	(17)	(18)
(13)	Baseline positive mood	1					
		–					
(14)	Baseline negative mood	.0505	1.0000				
		(.7387)	–				
(15)	Intrinsic Motivation Inventory (IMI) after Block 1: Interest/Enjoyment	**.3710***	.0124	1.0000			
		(.0111)	(.9349)	–			
(16)	IMI after Block 1: Competence	.2499	−.1043	**.3348***	1.0000		
		(.0939)	(.4901)	(.0229)	–		
(17)	IMI after Block 1: Effort	.0587	−.0713	**.2985***	−.2023	1.0000	
		(.6985)	(.6376)	(.0439)	(.1775)	–	
(18)	IMI after Block 1: Feeling tense or pressured	−.2026	.0717	−.0173	**−.3406***	**.6313****	1.0000
		(.1770)	(.6360)	(.9092)	(.0205)	(.0000)	–
(19)	Positive mood after Block 1	**.6941****	.1047	**.6970****	**.3543***	.2222	−.1847
		(.0000)	(.4887)	(.0000)	(.0157)	(.1379)	(.2190)
(20)	Negative mood after Block 1	.0852	**.5916****	.0589	−.0617	.2073	**.3156***
		(.5733)	(.0000)	(.6977)	(.6836)	(.1669)	(.0326)
(21)	Intrinsic Motivation Inventory (IMI) after Block 2: Interest/Enjoyment	**.3389***	.0244	**.7525****	.1101	.2406	.1022
		(.0281)	(.8781)	(.0000)	(.4876)	(.1248)	(.5195)
(22)	IMI after Block 2: Competence	**.4500****	.1494	.2699	**.6024****	−.1243	−.2228
		(.0028)	(.3450)	(.0839)	(.0000)	(.4328)	(.1561)
(23)	IMI after Block 2: Effort	.1755	−.0600	.1789	−.2395	**.8368****	**.6019****
		(.2663)	(.7060)	(.2569)	(.1265)	(.0000)	(.0000)
(24)	IMI after Block 2: Feeling tense or pressured	−.2574	.1711	.0420	−.1824	**.5182****	**.8195****
		(.0998)	(.2786)	(.7915)	(.2475)	(.0004)	(.0000)
(25)	Positive mood after Block 2	**.6910****	.0794	**.5627****	.2169	.1878	−.1625
		(.0000)	(.6170)	(.0001)	(.1677)	(.2336)	(.3037)
(26)	Negative mood after Block 2	.0974	**.4631****	.1007	.0088	.1338	.3005
		(.5395)	(.0020)	(.5258)	(.9560)	(.3982)	(.0531)

(*continued on next page*)

Table 2 (*continued*)

(27)	Baseline Heart Rate (HR)	.0934	.1785	.0319	−.0713	.0623	.0208
		(.5372)	(.2352)	(.8336)	(.6376)	(.6807)	(.8908)
(28)	% change in HR from baseline to Block 1	−.1940	−.1180	−.0447	.0464	.1326	.1483
		(.1965)	(.4348)	(.7683)	(.7596)	(.3796)	(.3254)
(29)	% change in HR from baseline to Block 2	−.1243	−.1719	−.0865	.0544	.1202	.1644
		(.4271)	(.2705)	(.5814)	(.7290)	(.4425)	(.2921)
(30)	High Frequency Heart Rate Variability (HF-HRV): Baseline	−.2170	−.1268	−.2437	.0693	−.0927	−.1025
		(.1474)	(.4010)	(.1026)	(.6471)	(.5402)	(.4980)
(31)	HF-HRV: Block 1	−.2672	.1051	−.2336	.0088	−.0147	−.0922
		(.0726)	(.4871)	(.1182)	(.9538)	(.9229)	(.5422)
(32)	HF-HRV: Block 2	−.2367	.1285	−.2916	.0638	−.0762	−.1302
		(.1218)	(.4059)	(.0548)	(.6809)	(.6230)	(.3997)
	Variable	(19)	(20)	(21)	(22)	(23)	(24)
(19)	Positive mood after Block 1	1					
		–					
(20)	Negative mood after Block 1	.0923	1.0000				
		(.5419)	–				
(21)	Intrinsic Motivation Inventory (IMI) after Block 2: Interest/Enjoyment	**.5653****	−.0241	1.0000			
		(.0001)	(.8795)	–			
(22)	IMI after Block 2: Competence	**.3939****	.0381	**.4085****	1.0000		
		(.0098)	(.8105)	(.0072)	–		
(23)	IMI after Block 2: Effort	.2181	.2616	**.3798***	−.0095	1.0000	
		(.1653)	(.0942)	(.0131)	(.9523)	–	
(24)	IMI after Block 2: Feeling tense or pressured	−.1809	**.5264****	.0654	−.2134	**.4695****	1.0000
		(.2516)	(.0003)	(.6807)	(.1748)	(.0017)	–
(25)	Positive mood after Block 2	**.8857****	−.0001	**.7049****	**.4686****	**.3734***	−.2046
		(.0000)	(.9996)	(.0000)	(.0017)	(.0149)	(.1938)
(26)	Negative mood after Block 2	.1860	**.8117****	.0231	−.0241	.1603	**.4453****
		(.2383)	(.0000)	(.8844)	(.8795)	(.3106)	(.0031)
(27)	Baseline Heart Rate (HR)	.2132	.1517	−.0358	−.1931	.1279	.0888
		(.1548)	(.3143)	(.8221)	(.2205)	(.4197)	(.5762)
(28)	% change in HR from baseline to Block 1	−.1791	−.1364	−.1621	.0699	−.0827	.1614
		(.2337)	(.3662)	(.3051)	(.6599)	(.6024)	(.3073)
(29)	% change in HR from baseline to Block 2	−.1641	−.1370	−.0802	.0804	.0056	.1256
		(.2929)	(.3810)	(.6184)	(.6173)	(.9725)	(.4338)
(30)	High Frequency Heart Rate Variability (HF-HRV): Baseline	−.0695	−.0048	**-.3104***	−.0741	−.0380	−.0158
		(.6460)	(.9745)	(.0454)	(.6409)	(.8110)	(.9207)
(31)	HF-HRV: Block 1	−.1213	.1207	−**.3672***	−.0504	.0940	.0971
		(.4218)	(.4243)	(.0168)	(.7511)	(.5538)	(.5406)
(32)	HF-HRV: Block 2	−.1657	.1543	−**.4144****	.0619	−.0271	.0187
		(.2825)	(.3172)	(.0064)	(.6969)	(.8647)	(.9063)
	Variable	(25)	(26)	(27)	(28)	(29)	(30)
(25)	Positive mood after Block 2	1					
		–					
(26)	Negative mood after Block 2	.0913	1.0000				
		(.5653)	–				
(27)	Baseline Heart Rate (HR)	.1652	**.3585***	1.0000			
		(.2957)	(.0197)	–			
(28)	% change in HR from baseline to Block 1	−**.3090***	−.2275	−**.4534****	1.0000		
		(.0465)	(.1473)	(.0016)	–		
(29)	% change in HR from baseline to Block 2	−.1927	−.2591	−**.5038****	**.8721****	1.0000	
		(.2274)	(.1019)	(.0006)	(.0000)	–	
(30)	HF-HRV: Baseline	−.0902	.1325	.2431	−.2863	**-.3472***	1.0000
		(.5699)	(.4029)	(.1036)	(.0537)	(.0225)	–
(31)	High Frequency Heart Rate Variability (HF-HRV): Block 1	−.1415	.2416	.1507	.1248	.0385	**.6283****
		(.3714)	(.1231)	(.3175)	(.4085)	(.8064)	(.0000)
(32)	HF-HRV: Block 2	−.2083	.1656	.2292	−.0264	−.1012	**.6731****
		(.1857)	(.2945)	(.1345)	(.8651)	(.5185)	(.0000)
	Variable	(31)	(32)				
(31)	HF-HRV: Block 1	1					
		–					
(32)	HF-HRV: Block 2	**.8702****	1.0000				
		(.0000)	–				

$^{*}p < .05$, $^{**}p < .01$.

3.3. Surveys

Pairwise correlations of survey and neurophysiological measures are in Table 2. There was no correlation between algorithm use and trait measures of empathic concern or personal distress (*ps* > .1). People who believed that others are generally trustworthy were more than twice as likely to adopt an algorithm (26% vs. 12%, $t(525.35) = -3.97$, $p < .01$, $d = 0.33$). Females were about twice as likely to adopt an algorithm as males (26% for females vs. 12% for males, $t(496.99) = -4.22$, $p < .01$, $d = 0.36$), but there was no significant difference in neurophysiological responses across genders. Individual-specific measures including age, academic major, initial mood, risk aversion, and affinity for computers were unrelated to algorithm adoption (*ps* > .10).

We did find that algorithm use reduced positive mood from baseline after Block 1 ($r = -0.35$, $p = .02$), but not in Block 2, suggesting acclimation to algorithm use. None of the items from the IMI survey (enjoyment, effort, competency, and being tense/pressured during the task) differed between blocks. Yet, when people successfully completed more mazes in Block 1, they reported putting less effort in solving the mazes in Block 2 ($r = -0.31$, $p = .05$).

4. Discussion

We designed this study to explore the question of why people choose to trust algorithms and how each person's performance and neurophysiology affects, and is affected by, these choices. We gave participants a maze-solving task in four different conditions: with no information about the algorithm (NI), accuracy information (AI), and low and high social proof (LS, HS).

A key finding was that social proof is the most effective way we tested to persuade people to adopt algorithms. Knowing that other people had used the algorithm increased participants' chances of adopting it. Indeed, there was no difference in the effect of low versus high social proof on algorithm adoption, consistent with research showing the power of even a modicum of social proof on decisions (Amblee & Bui, 2011; Hajli, 2015; Hajli et al., 2014, pp. 673–689). From the survey results, people who thought that others were generally trustworthy were more than twice as likely to adopt an algorithm. These individuals may be more prone to be affected by social influence This effect was strongest in the nonsocial conditions (NI, AI) where the information about the algorithm had less influence on decisions, (Trustworthy: 21.5%, Untrustworthy: 6.9%, $t(323) = 3.72$, $p < .01$, $d = 0.42$), but was not statistically significant in the social conditions (LS, HS) where the information was more influential (Trustworthy: 31.7%, Untrustworthy: 20.0%, 2-tailed $p = .08$).

We also found that participants in the social conditions had higher HF-HRV, a state that is associated with social connectedness and positive emotions (e.g., Kok & Fredrickson, 2010). This is consistent with research showing that people enjoy using technology if it triggers empathy (Shin, 2018). A speculative explanation for this finding is that people sometimes anthropomorphize algorithms: there is evidence that people use similar neurophysiological mechanisms to trust algorithms as they would other people (DeSteno et al., 2012; Nass et al., 1999; Riedl et al., 2010). Companies appear to make use of this effect by personalizing decision aids by, for example, giving it a name like Apple's Siri or Amazon's Alexa.

Our results extend the literature showing that social proof in uncertain environments affects decisions. Information about what others have done increases trust in new technologies and drives purchasing decisions (Amblee & Bui, 2011; Hajli, 2015; Hajli et al., 2014, pp. 673–689). In a similar way, building user-centric software interfaces that build trust increase the use of computerized health services (Shin & Biocca, 2017; Shin, Lee, & Hwang, 2017). Social network services (SNS), like Facebook and Twitter, seek to demonstrate security and privacy in order to build the perception of trustworthiness that sustains participation and influences adoption by new users (Shin, 2010). Moreover, social information about related products ("people who bought this item also bought …") positively influence sales of digital products and likely other products (Amblee & Bui, 2011). As social creatures, we are acutely aware, and influenced by what others, especially those "like us," are doing. In some situations, social influence can lead to conformity or decision disengagement and so should be deployed carefully (Cialdini & Goldstein, 2004).

We did not find support for our hypothesis that higher initial cognitive load would increase algorithm adoption. This may have occurred because the maze task was not sufficiently challenging or the money that could be earned was not high enough. This could be due to our peripheral measurement of cognitive load that is confounded with attention. Additional studies using other measurement techniques, for example, EEG, could further test this hypothesis. We also did not find that algorithm adoption varied with initial mood, dispositional empathy, or reactivity to stress. However, we did find that using an algorithm reduced cognitive load in all the conditions we tested. This finding carries additional weight than the previous studies because we used neurophysiologic measures rather than self-report.

Participants also reported that they put less effort into solving Block 2 mazes if they had done well in Block 1. This could reflect learning, cognitive fatigue, or a wealth effect from their Block 1 earnings. Future studies should vary the task difficulty, in addition to directly measuring cognitive load, since the former has been shown to increase algorithm adoption (Alvarado-Valencia & Barrero, 2014; Goddard et al., 2012).

A surprising finding was that women adopted algorithms twice as often as men did. This finding is at odds with most of the existing literature on algorithm adoption and could be an artifact of testing highly-educated participants. Algorithm adoption has been shown to increase with education, and both the men and women in our study were highly educated (David, 2015; The Nation's Report Card. Undated) and women tend to get better grades than men in college (The Nation's Report Card. Undated). It is possible, though, that women differ in the types of automation they prefer compared to men. In the present study that asked participants to solve mazes, the generally better average spatial abilities of men may be driving this finding (Galea & Kimura, 1993; Lövdén et al., 2007; Moffat, Hampson, & Hatzipantelis, 1998).

A final cautionary finding is that our results indicate that adopting automation could be ill-advised in some settings. We found that after algorithm adoption in the No Information condition, participants' cognitive loads did not increase from baseline as they did in the other conditions. This suggests that participants did not monitor the algorithm when there was no information about it—precisely the situation in which monitoring is most warranted. As a result of this relative inattention, performance in this condition suffered compared to those who did not use the algorithm for advice. In contrast, in the social conditions, adopters' attention was less than that of non-adopters' but remained above baseline showing that people monitored the algorithm's recommendations when solving the maze. At the same time, HF-HRV increased in the S condition showing the influence of social proof. This combination of sympathetic and parasympathetic switching produces "focused relaxation" and has been found to improve performance (Barraza et al., 2015). We found the same effect wherein algorithm adopters performed 29% better than non-adopters. This is an important issue that should be explored further — for example, it may not be best to abandon steering wheels all together just yet in order to get people to monitor the algorithms driving their cars. Blind trust is foolish, but so is avoiding algorithms altogether.

References

Alvarado-Valencia, J. A., & Barrero, L. H. (2014). Reliance, trust and heuristics in judgmental forecasting. *Computers in Human Behavior, 36*, 102–113.

Amblee, N., & Bui, T. (2011). Harnessing the influence of social proof in online shopping: The effect of electronic word of mouth on sales of digital microproducts. *International*

Journal of Electronic Commerce, 16(2), 91–114.
Barraza, J. A., Alexander, V., Beavin, L. E., Terris, E. T., & Zak, P. J. (2015). The heart of the story: Peripheral physiology during narrative exposure predicts charitable giving. *Biological Psychology, 105*, 138–143.
Brookhuis, K. A., & de Waard, D. (2010). Monitoring drivers' mental workload in driving simulators using physiological measures. *Accident Analysis & Prevention, 42*(3), 898–903.
Buchan, N. R., Croson, R. T., & Solnick, S. (2008). Trust and gender: An examination of behavior and beliefs in the Investment Game. *Journal of Economic Behavior & Organization, 68*(3), 466–476.
Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology, 55*, 591–621.
David, H. (2015). Why are there still so many jobs? The history and future of workplace automation. *The Journal of Economic Perspectives, 29*(3), 3–30.
Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44*(1), 113–126.
DeSteno, D., Breazeal, C., Frank, R. H., Pizarro, D., Baumann, J., Dickens, L., et al. (2012). Detecting the trustworthiness of novel partners in economic exchange. *Psychological Science* 0956797612448793.
Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General, 144*(1), 114.
Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191.
Fredericks, T. K., Choi, S. D., Hart, J., Butt, S. E., & Mital, A. (2005). An investigation of myocardial aerobic capacity as a measure of both physical and cognitive workloads. *International Journal of Industrial Ergonomics, 35*(12), 1097–1107.
Galea, L. A., & Kimura, D. (1993). Sex differences in route-learning. *Personality and Individual Differences, 14*(1), 53–65.
Gladwell, M. (2007). *Blink: The power of thinking without thinking: Back bay books*.
Goddard, K., Roudsari, A., & Wyatt, J. C. (2012). Automation bias: A systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association, 19*(1), 121–127.
Gönül, S., Önkal, D., & Goodwin, P. (2009). Expectations, use and judgmental adjustment of external financial and economic forecasts: An empirical investigation. *Journal of Forecasting, 28*(1), 19–37.
Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment, 12*(1), 19.
Hajli, N. (2015). Social commerce constructs and consumer's intention to buy. *International Journal of Information Management, 35*(2), 183–191.
Hajli, N., Lin, X., Featherman, M. S., & Wang, Y. (2014). *Social word of mouth: How trust develops in the market.* Hajli, N., Lin, X., Featherman, MS, Wang, Y.
Hoff, K. A., & Bashir, M. (2015). Trust in automation integrating empirical evidence on factors that influence trust. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 57*(3), 407–434.
Ho, G., Kiff, L. M., Plocher, T., & Haigh, K. Z. (2005). A model of trust & reliance of automation technology for older users. *Paper presented at the AAAI-2005 fall symposium: "Caring machines: AI in eldercare*.
Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *The American Economic Review, 92*(5), 1644–1655.
Kok, B. E., & Fredrickson, B. L. (2010). Upward spirals of the heart: Autonomic flexibility, as indexed by vagal tone, reciprocally and prospectively predicts positive emotions and social connectedness. *Biological Psychology, 85*(3), 432–436.
Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., & Fehr, E. (2005). Oxytocin increases trust in humans. *Nature, 435*(2), 673–676.
Lahtinen, T. M., Koskelo, J. P., Laitinen, T., & Leino, T. K. (2007). Heart rate and performance during combat missions in a flight simulator. *Aviation Space & Environmental Medicine, 78*(4), 387–391.
Lerch, F. J., Prietula, M. J., & Kulik, C. T. (1997). The Turing effect: The nature of trust in expert systems advice. *Paper presented at the expertise in context*.
Lövdén, M., Herlitz, A., Schellenbach, M., Grossman-Hutter, B., Krüger, A., & Lindenberger, U. (2007). Quantitative and qualitative sex differences in spatial navigation. *Scandinavian Journal of Psychology, 48*(5), 353–358.
MacCoun, R. J. (2012). The burden of social proof: Shared thresholds and social influence. *Psychological Review, 119*(2), 345.
Makary, M. A., & Daniel, M. (2016). Medical error—the third leading cause of death in the US. *BMJ, 353*, i2139.
McAuley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric properties of the intrinsic motivation inventory in a competitive sport setting: A confirmatory factor analysis. *Research Quarterly for Exercise & Sport, 60*(1), 48–58.
McBride, S. E., Rogers, W. A., & Fisk, A. D. (2011). Understanding the effect of workload on automation use for younger and older adults. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 53*(6), 672–686.
Merritt, S. M., & Ilgen, D. R. (2008). Not all trust is created equal: Dispositional and history-based trust in human-automation interactions. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 50*(2), 194–210.
Moffat, S. D., Hampson, E., & Hatzipantelis, M. (1998). Navigation in a "virtual" maze: Sex differences and correlation with psychometric measures of spatial ability in humans. *Evolution and Human Behavior, 19*(2), 73–87.
Mosier, K. L., & Skitka, L. J. (1996). Human decision makers and automated decision aids: Made for each other. *Automation and human performance: Theory and applications*, 201–220.
Nass, C., Moon, Y., & Carney, P. (1999). Are people polite to Computers? Responses to computer-based interviewing Systems1. *Journal of Applied Social Psychology, 29*(5), 1093–1109.
Önkal, D., Goodwin, P., Thomson, M., Gönül, S., & Pollock, A. (2009). The relative influence of advice from human experts and statistical methods on forecast adjustments. *Journal of Behavioral Decision Making, 22*(4), 390–409.
Paas, F. G., & Van Merriënboer, J. J. (1993). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human Factors, 35*(4), 737–743.
Pak, R., Rovira, E., McLaughlin, A. C., & Baldwin, N. (2016). Does the domain of technology impact user trust? Investigating trust in automation across different consumer-oriented domains in young adults, military, and older adults. *Theoretical Issues in Ergonomics Science*, 1–22.
Parasuraman, R., Cosenzo, K. A., & De Visser, E. (2009). Adaptive automation for human supervision of multiple uninhabited vehicles: Effects on change detection, situation awareness, and mental workload. *Military Psychology, 21*(2), 270.
Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors, 52*(3), 381–410.
Riedl, R., Hubert, M., & Kenning, P. (2010). Are there neural gender differences in online trust? An fMRI study on the perceived trustworthiness of eBay offers. *MIS Quarterly, 34*(2), 397–428.
Schultz, K. (2011). *Being wrong: Adventures in the margin of error. Granta Books*.
Shappell, S. A., & Wiegmann, D. A. (2012). *A human error approach to aviation accident analysis: The human factors analysis and classification system.* Ashgate Publishing, Ltd.
Shin, D. (2010). The effects of trust, security and privacy in social networking: A security-based approach to understand the pattern of adoption. *Interacting with Computers, 22*(5), 428–438. https://doi.org/10.1016/j.intcom.2010.05.001.
Shin, D. (2018). Empathy and embodied experience in virtual environment: To what extent can virtual reality stimulate empathy and embodied experience? *Computers in Human Behavior, 78*, 64–73. https://doi.org/10.1016/j.chb.2017.09.012.
Shin, D., & Biocca, F. (2017). Health experience model of personal informatics: The case of a quantified self. *Computers in Human Behavior, 69*, 62–74. https://doi.org/10.1016/j.chb.2016.12.019.
Shin, D., Lee, S., & Hwang, Y. (2017). How do credibility and utility affect the user experience of health informatics services? *Computers in Human Behavior, 67*, 292–302. https://doi.org/10.1016/j.chb.2016.11.007.
Silver, N. (2012). *The signal and the noise: Why so many predictions fail-but some don't: Penguin*.
Stanton, N. A., & Salmon, P. M. (2009). Human error taxonomies applied to driving: A generic driver error taxonomy and its implications for intelligent transport systems. *Safety Science, 47*(2), 227–237.
Sundar, S. S., Xu, Q., & Oeldorf-Hirsch, A. (2009). Authority vs. peer: How interface cues influence users. *Paper presented at the CHI'09 extended abstracts on human factors in computing systems*.
Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257–285.
Sztajzel, J. (2004). Heart rate variability: A noninvasive electrocardiographic method to measure the autonomic nervous system. *Swiss Medical Weekly, 134*, 514–522.
Thayer, J. F., Hansen, A. L., Saus-Rose, E., & Johnsen, B. H. (2009). Heart rate variability, prefrontal neural function, and cognitive performance: The neurovisceral integration perspective on self-regulation, adaptation, and health. *Annals of Behavioral Medicine, 37*(2), 141–153.
The Nation's Report Card. (Undated). Gender: Grade Point Average. National Assessment of Educational Progress. https://www.nationsreportcard.gov/hsts_2009/gender_gpa.aspx Retrieved 2/22/18.
Waern, Y., & Ramberg, R. (1996). People's perception of human and computer advice. *Computers in Human Behavior, 12*(1), 17–27.
Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063–1070.
Xing, W., Guo, R., Petakovic, E., & Goggins, S. (2015). Participation-based student final performance prediction model through interpretable Genetic Programming: Integrating learning analytics, educational data mining and theory. *Computers in Human Behavior, 47*, 168–181.
Zak, P. J. (2008). The neurobiology of trust. *Scientific American* June, 88–95.
Zak, P. J. (2012). *The Moral Molecule: The Source of Love and Prosperity.* New York: Random House.
Zak, P. J. (2017). *Trust Factor: The Science of Creating High-Performance Companies.* New York: AMACOM Div American Mgmt Assn.
Zak, P. J., Kurzban, R., & Matzner, W. T. (2004). The neurobiology of trust. *Annals of the New York Academy of Sciences, 1032*, 224–227.

Biographies

Alexander, V. (2014). *Center for Neuroeconomics Studies, Claremont Graduate University.* PhD, Economics Claremont Graduate University.
Blinder, C. (2017). *Center for Neuroeconomics Studies.* BA, Psychology, Pitzer College Claremont Graduate University.
Zak, P. J. (1994). *Center for Neuroeconomics Studies, Claremont Gaduate University.* PhD, Economics University of Pennsylvania.