MIS Quarterly

RESEARCH NOTE

THE PATH OF THE RIGHTEOUS: USING TRACE DATA TO UNDERSTAND FRAUD DECISIONS IN REAL TIME[1]

Markus Weinmann
University of Cologne, Cologne, GERMANY,
and Erasmus University, Rotterdam, THE NETHERLANDS {weinmann@wiso.uni-koeln.de}

Joseph S. Valacich
University of Arizona,
Tucson, AZ, U.S.A. {valacich@arizona.edu}

Christoph Schneider
IESE Business School, University of Navarra,
Barcelona, SPAIN {cschneider@iese.edu}

Jeffrey L. Jenkins
Brigham Young University,
Provo, UT, U.S.A. {jeffrey_jenkins@byu.edu}

Martin Hibbeln
University of Duisburg-Essen,
Duisburg, GERMANY {martin.hibbeln@uni-due.de}

Trace data—users' digital records when interacting with technology—can reveal their cognitive dynamics when making decisions on websites in real time. Here, we present a trace-data method, analyzing movements captured via a computer mouse, to assess potential fraud when filling out an online form. In contrast to existing fraud-detection methods, which analyze information after submission, mouse-movement traces can capture the cognitive deliberations as possible indicators of fraud as it is happening. We report two controlled studies using different tasks, where participants could freely commit fraud to benefit themselves financially. As they performed the tasks, we captured mouse-cursor movement data and found that participants who entered fraudulent responses moved their mouse significantly more slowly and with greater deviation. We show that the extent of fraud matters such that more extensive fraud increases movement deviation and decreases movement speed. These results demonstrate the efficacy of analyzing mouse-movement traces to detect fraud during online transactions in real time, enabling organizations to confront fraud proactively as it is happening at internet scale. Our method of analyzing actual user behaviors in real time can complement other behavioral methods in the context of fraud and a variety of other contexts and settings.

Keywords: Trace data, mouse-cursor movements, fraud, cognitive dissonance, Bayesian analysis

[1] Ron Cenfetelli was the accepting senior editor for this paper. Shuk Ying Ho served as the associate editor.

Introduction

Fraud[2] is prevalent in industry, amounting to an estimated US$5 trillion in damages annually (Crowe Global, 2021). Insurance fraud, for example, is associated with estimated yearly damages of US$80 billion (Coalition Against Insurance Fraud, 2021) and affects organizations and insurance consumers alike, costing individuals up to US$700 annually in increased premiums (FBI, 2021). The prevalence of fraud has been exacerbated by businesses increasingly moving online (Morrow & Maynard, 2021), but detecting it is challenging, particularly online. Most existing fraud detection methods analyze the *output* of a submission process for fraud, such as a submitted claim, and flag suspicious outputs for further investigation by an auditor. To complement existing methods, we present a method that analyzes the *input* during the submission process by using trace data—users' digital records when interacting with technology (cf. Berente et al., 2019)—namely, mouse-cursor movements.[3]

Analyzing mouse-movement data can provide an understanding of online users' psychological states and intentions. For example, Hibbeln et al. (2017) found that the speed and distance of mouse movements can indicate users' frustration on websites. Jenkins et al. (2019) found that similar characteristics of mouse movements correlated with deception and heightened galvanized skin responses in polygraph examinations during a sanctioned deception experiment. Although research has suggested that mouse-movement data may be useful for predicting fraud in online commerce (Valacich et al., 2013), little research has empirically tested the efficacy of analyzing mouse movements to predict fraud in a more natural context or how the extent of fraud influences mouse-movement data.

Drawing on theories from psychology and neuroscience, we explain and empirically validate how formulating a fraudulent response and the extent of fraud can subconsciously influence users' mouse movements. To this end, we (1) examine how mouse movements can indicate the decision-making process involving fraud, (2) create a measure of mouse-movement deviation that is valid in natural interactions, and (3) explore the relationship between fraud, the extent of fraud,[4] and mouse movements in a nonsanctioned context. Thus, our research question is: *How do online fraud decisions influence mouse movements?*

To answer this, we conducted two online studies. Study 1 focused on internal validity using an established visual-perception task that incentivized participants to commit fraud. Study 2 focused on external validity using an online insurance claim task. The results show that both the presence of fraud and the extent of fraud change mouse-movement behavior. These studies demonstrate that analyzing mouse-movement traces can be extended to the presence and extent of fraud in natural online settings and nonsanctioned tasks, creating a foundation for practitioners to reveal users' intent and identify fraud at scale in real time.

Related Work

Online fraud is widespread, leading to costs for both consumers and organizations. Hence, organizations use techniques ranging from statistical approaches to AI-based approaches to increase fraud detection accuracy (Carneiro et al., 2017), thereby lowering ex post monitoring costs. However, given the challenges of internet scale, auditors can investigate only a portion of the most egregious claims. Many misrepresentations slip through the cracks, amounting to billions of dollars lost (Hunter, 2015). Hence, there is clearly a need to improve fraud detection systems.

Most fraud detection approaches in practice use ex post auditing instead of real-time monitoring. They analyze information *after* submission, so fraud can be detected only after it is committed. We propose a trace-data method that analyzes users' mouse movements to detect possible fraud in real time. Tracking mouse movements can provide insight into various cognitive processes (Freeman et al., 2011; Freeman & Ambady, 2011; Kieslich et al., 2020), some of which are related to deception, including decision conflict (McKinstry et al., 2008), cognitive competition (Dale et al., 2007), stress (Banholzer et al., 2021), and emotional reactions (Hibbeln et al., 2017).

We contribute to this body of research by extending existing work, drawing on two papers that are relevant to our research. In particular, we build on the work of Hibbeln et al. (2017), who explained how mouse movements can indicate user frustration in online contexts, and Jenkins et al. (2019), who examined mouse movements in concealed information tasks. (See Table 1 for a comparison of our work with Hibbeln et al. (2017) and Jenkins et al. (2019).)

[2] In this article, we focus on fraudulent responses on online forms. Online fraud not on online forms, such as gift card fraud, dating fraud, and mobile phone scams committed by cybercriminals, is beyond the scope of this article.

[3] Mouse-cursor movements reflect how users move their computer mouse when interacting with an interface; in the remainder of the paper, we will use the simpler term *mouse movements*.

[4] We thank Reviewer 3 for this suggestion.

Table 1. Summary of Major Contributions of Our Paper Compared to Hibbeln et al. (2017) and Jenkins et al. (2019)

	Compared to Hibbeln et al. (2017)	Compared to Jenkins et al. (2019)
Theoretical contribution	Hibbeln et al. used attentional control theory to explain how frustration causes changes in people's mouse movements. By *combining the RAM, CLT, and cognitive dissonance/conflict theory*, we explain how the fraud decision process also influences mouse movements.	Jenkins et al. used the RAM and CLT to explain why an orienting response will change people's mouse movements when providing a predetermined response—participants were told beforehand how to respond. We extend these theories and *add cognitive dissonance/conflict theory* to explain how mouse movements can illuminate the cognitive dynamics of the real-time decision process to be fraudulent. In addition, we draw on these theories to *explain how the extent of fraud influences mouse movements.*
Methodological contributions	Hibbeln et al. calculated total distance and speed, assuming tasks have the same length for comparison. We introduce a valid measure of deviation that allows us to *compare tasks that differ in length and structure* by controlling for the shortest distance of movement segments.	Jenkins et al. used the graded motor response analysis to measure trajectory deviation for interactions with the same beginning and end points. We introduce a valid measure of deviation that allows interactions to have *different beginning, intermediate, and end points*.
Ecological validity	Hibbeln et al. provided an ecologically valid test of frustration in an experimental setting. We provide an ecologically valid test of *unsanctioned fraud*.	Jenkins et al. show that mouse movements differ between guilty and innocent people in sanctioned deception scenarios. We show that mouse movements differ in *unsanctioned fraud* scenarios and in natural interaction contexts.

In this paper, we extend the response activation model (RAM) used by Hibbeln et al. (2017) and cognitive load theory (CLT) to understand people's mouse movements, revealing the cognitive dynamics of a fraud decision in real time. In contrast to Jenkins et al. (2019), we examine fraud in an unsanctioned and realistic context, addressing their study's weaknesses of artificial interactions, sanctioning, and time delays between the fraudulent act and mouse-movement data collection. Further, we introduce a novel deviation measure computed in natural online interactions, which allows for the analysis of potentially fraudulent responses in commonly used online forms. Finally, we extend Jenkins et al.'s paper by examining how not only the presence of fraud but also the extent of fraud influences mouse movements.

Theory and Hypotheses

In this section, we apply theories on cognitive dissonance/conflict, the response activation model (RAM), and cognitive load theory (CLT) to understand the real-time cognitive dynamics of a decision to commit fraud. To develop hypotheses that explain how the fraud decision-making process correlates with mouse-movement changes, we build on two axioms of deception: deciding to be fraudulent increases both cognitive dissonance/conflict and cognitive load.

According to interpersonal deception theory, people are likely to experience cognitive and moral conflict when engaging in deception (Buller & Burgoon, 1996; Nuñez et al., 2005). People must formulate their deception by creating a credible fraudulent response, often resulting in additional cognitive conflict as they reevaluate their responses to ensure a credible story. Further, it is common for people to experience moral conflict and cognitive dissonance. They hesitate and reconsider their responses because they evaluate the consequences of being caught. Likewise, they are likely to experience increased cognitive dissonance because the act of being fraudulent may conflict with their self-image of being honest (Derrick et al., 2013; Nuñez et al., 2005). In both of these cases, people may reconsider their decisions or details of their responses.

According to the RAM, competing cognitions and reevaluation of response details can influence fine motor control (Welsh & Elliott, 2004). The RAM posits that hand movements respond to all thoughts that have even a minor potential to result in movement, so-called actionable potential (Welsh & Elliott, 2004). When people knowingly provide misleading information online, they are likely to

double-check, reconsider, hesitate, or question actions. Thus, when moving the mouse to commit fraud, they may consider stopping the action for fear of being caught; likewise, they may consider responding differently to make the fraud more believable. Even if people don't execute the corresponding actions, their fraudulent thoughts have actionable potential that can—in contrast to nonfraudulent thoughts—result in different movements.

The RAM explains the relationship between hand movements and thoughts. When a thought with actionable potential enters the mind, the mind automatically and subconsciously programs a movement response to fulfill that intention (Welsh & Elliott, 2004). This includes transmitting nerve impulses to the muscles to move the hand toward the stimulus and realize the intention (Song & Nakayama, 2008). If a person has accordant cognitions, their mouse trajectory follows roughly a straight line toward the movement's target—in this case, the intended input field on the online form. Deviations from the straight line can result from fraudulent intentions because the mind programs movement responses toward other stimuli with actionable potential. Such deviations manifest in different mouse-movement characteristics. In sum, we hypothesize:

H1a: *When providing fraudulent responses on an online form, people perform mouse movements that exhibit greater deviation from the shortest path.*

Cognitive processing is likely to differ based on the fraud decision (Buller & Burgoon, 1996). In some instances, a fraud decision is binary (Should I steal this candy bar?); in other contexts, the decision is more complex and also involves extent (Should I misrepresent my insurance claim? By how much?). The fraud extent influences the amount and duration of the heightened cognitive deliberation (Thomas & Biros, 2011). Extensive fraud may have larger consequences and often requires greater justification. Extensive fraud may result in heightened moral deliberation and second-guessing, in addition to requiring more elaborate excuses and backstories, to reduce the risk of being caught (Olson & Raz, 2021). In each of these cases, one is likely to deliberate among options as the extent of fraud increases, giving attention to more stimuli with actionable potential. In addition, when fraud moves from a binary decision to one involving the extent of fraud, there are often more options from which to choose. The binary case has just two options; however, in the latter case, the extent of fraud can be split among numerous options, each with actionable potential. According to the RAM, this increased actionable potential results in more movement deviation (Welsh & Elliott, 2004). In sum, we hypothesize:

H1b: *The extent of fraud on online forms is positively correlated with the extent of deviation from the shortest path.*

Deception is a complex cognitive process that increases cognitive load—another axiom of deception (Carrión et al., 2010). As discussed, people generate false information when engaging in deception, attempting to minimize the evidence and formulate a credible response (Derrick et al., 2011). Doing both requires behaving strategically, which increases cognitive load and decreases available working memory (Buller & Burgoon, 1996). Reduced working memory slows reaction times (Unsworth & Engle, 2005) and hand movements (Meyer et al., 1988). For example, when visually guiding the hand to a target, the brain has less time to program corrections to one's movement trajectory, thus decreasing movement precision. The brain automatically compensates for the decreased precision by reducing movement speed (Meyer et al., 1988). Because hand-movement speed and precision are inversely related (Plamondon & Alimi, 1997), movement precision improves only if movement speed is reduced. Thus, when a person takes more time to perceive and program corrections, their hands operate better under the constraint of slower reaction times (Meyer et al., 1988). In sum, we hypothesize:

H2a: *When providing fraudulent responses on an online form, people perform slower mouse movements.*

As outlined, a greater extent of fraud leads to more cognitive deliberation as people consider the heightened consequences, increased fraud options, and more elaborate justification and risk mitigation needs. This increased cognitive load decreases available working memory (Buller & Burgoon, 1996), decreases reaction times (Unsworth & Engle, 2005), and ultimately causes slower hand movements (Meyer et al., 1988). In sum, we hypothesize:

H2b: *The extent of fraud on online forms is negatively correlated with the speed of mouse movements.*

Overview of Studies

To test our hypotheses, we conducted two online studies (see Table 2; see https://osf.io/escyg for data files and analysis scripts). Study 1 focused on internal validity using a well-validated visual perception task that incentivized—but did not actively encourage—participants to commit fraud. Study 2 focused on external validity using a real-world insurance claim scenario. In addition, Study 2 allowed us to examine the influence of the extent of fraud on mouse movements. All participants had the freedom to commit fraud to boost their compensation. Fraud increased movement deviation by 16% and decreased speed by 9% in Study 1, and it increased deviation by 49% and decreased speed by 29% in Study 2. Furthermore, the results of Study 2 suggest that the extent of fraud influences deviation and speed.

Table 2. Overview of Studies				
Study	**Task**	**Purpose**	**Obs.**	**Findings**
1	Established visual perception task	Address *internal* validity: explanatory study to test H1a & H2a	1150	Fraud increased deviation by 16% and decreased speed by 9%.
2	Completing insurance claim forms	Address *external* validity: extend Study 1 and test H1b & H2b	550	Fraud increased deviation by 49% and decreased speed by 29%. The extent of fraud increased deviation and decreased speed.

Study 1: Established Cheating Task

Using a well-established task that incentivizes participants to commit fraud, Study 1 shows that deception causes changes in mousing behavior. Fraudulent behavior increases movement deviation and decreases speed.

Materials and Procedure

To examine how fraudulent behavior influences mouse movements, we adapted a cheating task developed by Gino et al. (2010). In this "flexible dot task" (Hochman et al., 2016), participants were asked to truthfully identify which side of a square, divided by a diagonal line, contained more dots (see Figure 1). We used 10 scenarios with randomly generated dots—half with more on the right side.[5] To reduce the likelihood of erroneous responses, we presented only unambiguous scenarios with at least a five-dot difference between the left and right sides. Each scenario appeared for one second; thereafter, participants had to select their answer to the question "Which side contained more dots?" by moving their mouse cursor to the "left" or "right" selection buttons. (See Appendix A for the instructions.)

Participants completed a sequence of 10 unique trials. For each trial, the mouse cursor was anchored in the middle of the screen by requiring participants to click on a "Start" button. After participants clicked, one of the 10 randomly ordered scenarios loaded (see Figure 1). While participants selected their answers, we used JavaScript to record their mouse movements and send the data to a web service developed by the research team for further processing.

We incentivized participants to commit fraud by paying a variable bonus based on their responses. Clicking the "right" button always had a higher payout, even when the left side clearly contained more dots, which encouraged participants to click "right" even when the correct answer was "left." Participants would receive 0.5 pence (approx. US¢0.65) for clicking on "more on left" or 5 pence (approx. US¢6.5) for clicking on "more on right" (see Figure 1). Thus, participants would receive the maximum payout by fraudulently reporting in all trials that more dots were presented on the right.

The task had four possible outcomes, defined by both the dots' locations and the participant's choice to click left or right (Hochman et al., 2016, see Table 3). Only Option 4 was considered fraudulent because only beneficial errors resulted in increased payouts (Hochman et al., 2016). Thus, we compared mouse movements for beneficial errors (i.e., fraud) to those for other responses (i.e., correct hits, correct rejects, and detrimental errors).

Participants

We recruited 150 participants over the age of 18 from the United States, using the participant recruitment platform Prolific.[6] We paid participants £1 for the 10-minute task (equivalent to a £6/US$8 hourly wage), in addition to the variable bonus. In line with previous mouse-tracking studies, we excluded participants whose completion time was longer than three standard deviations from the average (Freeman & Dale, 2013) and excluded anyone accessing the study on a mobile device. This resulted in a sample size of 115 participants with 1,150 valid observations (10 scenarios per participant). The mean age was 30.9 years, and 33.9% of participants were women.

[5] We used three other scenarios as practice scenarios, which we removed for the main analysis. However, the results remain robust, even when including these scenarios.

[6] Online recruitment platforms have been found to be appropriate for random-sample populations (Berinsky et al., 2012). For example, Mason and Suri (2012) found that the behavior of respondents on an online recruitment platform closely resembled that of participants in traditional laboratory experiments.

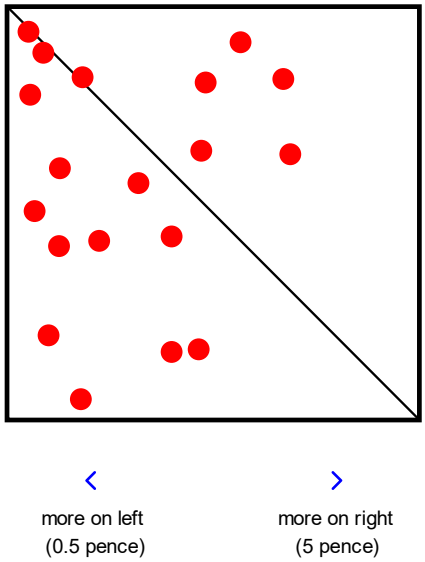


Figure 1. Sample Screen from the Flexible Dot Task

Table 3. Possible Outcomes of the Dot Task

Option	More dots on	Selection	Outcome
1	Left	Left	No fraud (low payout—correct reject)
2	Right	Right	No fraud (high payout—correct hit)
3	Right	Left	No fraud (low payout—detrimental error)
4	Left	Right	Fraud (high payout—beneficial error)

Data and Variables

Our self-developed web service used mouse-movement data[7] to calculate the variables of interest as follows. First, in line with the mouse-movement literature (Hehman et al., 2015), we normalized for users' screen resolutions[8] to ensure that movements were proportionally the same. Second, the web service calculated statistics for deviation and speed.

We created a novel measure of *deviation* that is valid in natural online formats and interactions—where people can have different beginning, intermediate, and end points—by averaging the deviation for each submovement. A submovement is an uninterrupted movement without a meaningful pause (typically a pause greater than 200 ms) or a click. By calculating each submovement deviation and averaging them, the deviation statistic accounts for differences in tasks that would require more or fewer movements. (This is relevant for more realistic settings, such as those of Study 2, in which engaging in fraudulent behavior required moving the cursor over longer distances.)

To calculate submovement deviation, we (1) calculated the *actual* distance traveled by the mouse cursor in that submovement, (2) calculated the *shortest* distance of the submovement (see Figure 2), and (3) divided the actual distance by the shortest distance. We calculated the actual distance by summing the distances between each consecutive *x*-/*y*-coordinate pair in the submovement and the shortest distance as the distance between the submovement's start and end points, a straight line representing the shortest distance required to move between the two points. By dividing the actual distance by the shortest distance, we accounted for the minimum required distance to complete a task and thereby for different task lengths and differences in the mouse cursor's location at the beginning of a task.

[7] Mouse-movement data included the cursor's *x*-/*y*-coordinate pairs and the corresponding timestamps at millisecond granularity.

[8] Screen resolution (measured in pixels) can be detected using simple scripts. Pixels are a unitless measure, meaning a pixel on one screen is not equivalent to a pixel on a different screen with a different resolution. Thus, traveling between two points on a webpage on a low-resolution screen will involve fewer pixels than traveling between two points on a website on a higher-resolution screen of the same size (although the physical distance on the screen might be equal). To account for this, we normalized all movements to a standard 8 × 6 grid (*x*-position × 8/screen width and *y*-position × 6/screen height).

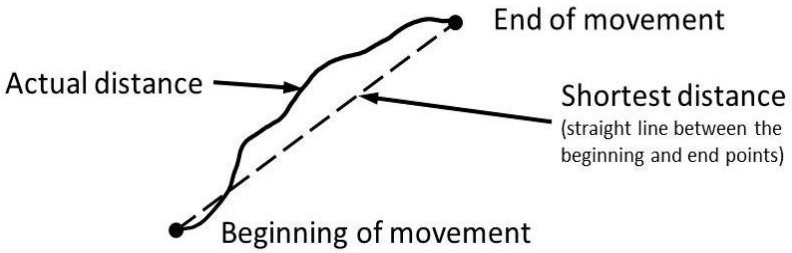


Figure 2. Actual and Shortest Distances

Thus, the higher the averaged ratio, the greater the deviation per unit moved (normalized pixel).[9] Speed was calculated as the actual distance divided by the movement time for each submovement. This value was then averaged across submovements.[10] We log-transformed deviation and speed because both were highly skewed.

Model Specification

We coded fraudulent responses (i.e., beneficial errors) as 1 and others as 0. We specified a linear multilevel regression model to estimate the effect of fraud on our two outcome variables, mouse-movement deviation (H1a) and mouse-movement speed (H2a). Multilevel models account for clustered data structures—such as observations belonging to the same participants—by allowing individual intercepts to vary. We used the following multilevel model:

$$y_{ij} = \alpha + \alpha_j + \beta \cdot \text{fraud}_{ij} + \epsilon_{ij}, \quad (1)$$

where y_{ij} denotes the i-th observation of y (i.e., mouse-movement deviation or speed) for the j^{th} participant; α represents the grand mean, and α_j the varying intercept for the j^{th} participant. β is the effect of fraud (i.e., beneficial error). For the deviation hypotheses (H1), we expected the coefficient to be positive; for the speed hypotheses (H2), we expected the coefficient to be negative.

Model Estimation

We used a Bayesian approach to estimate the models.[11] In a Bayesian procedure, it is necessary to define prior beliefs about the parameters. We used weakly informative priors from the literature (McElreath, 2020). For the intercept, we used a normal distribution with a mean of 0 and a standard deviation of 10, $\alpha \sim N(0,10)$; for all other parameters—the coefficients (β), the residual standard deviation, and for the group-level standard deviation—we used $N(0,1)$.[12] We used *R* (Ihaka & Gentleman, 1996) and the brms package (Bürkner, 2018) to estimate the multilevel models. We ran four chains with 4,000 iterations—2,000 for warm-up and 2,000 for sampling. The diagnostics show that all four chains converged, were well-mixed ($\hat{R} < 1.1$) and were of sufficient size, with an effective sample size (ESS) of > 1,000.

Results

Table 4 presents the summary statistics. Although only 50% of the pictures contained more dots on the right side, in 59% of the observations, participants stated that there were more dots on the right side. In total, 141 of the 1,150 valid observations (12%) were fraudulent.

[9] We calculated *deviation* as follows: $Deviation = \frac{\sum_{i=1}^{s} \frac{\sum_{k=1}^{c-1} \sqrt{(x_{i_k} - x_{i_{k+1}})^2 + (y_{i_k} - y_{i_{k+1}})^2}}{\sqrt{(x_{i_0} - x_{i_c})^2 + (y_{i_0} - y_{i_c})^2}}}{s}$, where *s* denotes the number of submovements, *c* the number of x/y-coordinate pairs in a submovement, *x* the x-coordinate, *y* the y-coordinate, and *t* the epoch timestamp associated with each x/y-coordinate pair.

[10] We calculated *speed* as follows: $Speed = \frac{\sum_{i=1}^{s} \frac{\sum_{k=1}^{c-1} \sqrt{(x_{i_k} - x_{i_{k+1}})^2 + (y_{i_k} - y_{i_{k+1}})^2}}{t_{i_c} - t_{i_0}}}{s}$.

[11] Bayesian approaches update prior beliefs about parameters when new evidence arrives (e.g., data from an experiment) to estimate posterior beliefs. Bayesian methods have some advantages over traditional methods. In recent years, traditional methods have been criticized, especially for leading only to a point estimate and relying too much on *p*-values, making them susceptible to *p*-hacking (Head et al., 2015), and for being difficult to interpret (Nuzzo, 2014). Therefore, recent calls suggest avoiding *p*-values altogether or handling them carefully (Mertens & Recker, 2020). Bayesian methods instead show the complete distribution of parameters with their credible intervals, making them easy to interpret; Bayesian methods can incorporate an existing knowledge base from literature or previous experiments and are suitable for small sample sizes. Finally, Bayesian methods are suitable for our use case—estimating multilevel models—where observations are clustered in participants or scenarios. (See Kruschke et al. 2012, for a discussion of Bayesian methods in management/organizational science.)

[12] We also tried half student-*t* priors for the standard deviation, resulting in similar results.

Table 4. Descriptive Statistics (Study 1)

Panel A: Full dataset					
Statistic	**Observations**	**% / Mean**	***SD***	**Min.**	**Max.**
All observations	1,150				
Fraud					
More dots right (presented)	575	50.0%			
More dots right (reported)	682	59.3%			
Beneficial errors (fraud)	141	12.3%			
Detrimental errors	34	3.0%			
Correct hits	541	47.0%			
Correct rejects	434	37.7%			
Mouse movement					
Deviation (normalized pixels)	1,150	5.16	4.54	2.00	86.13
Speed (normalized pixels/ms)	1,150	3.02	2.21	.68	51.49

Panel B: Split by fraud				
Statistic	**Nonfraud**		**Fraud**	
	Mean	***SD***	**Mean**	***SD***
Mouse movement				
Deviation (normalized pixels)	4.96	(3.02)	6.55	(10.06)
Speed (normalized pixels/ms)	3.05	(2.32)	2.81	(1.11)
Demographics				
Age	30.8	(10.5)	31.5	(10.7)
Gender (female)	33%		30%	
Observations	1,009		141	

Table 5 shows that decisions to commit fraud (1) significantly increased mouse-movement deviation (β = 0.15***; 95%-CI[13]: [0.07, 0.22]; Figure 3a) and (2) significantly decreased mouse-movement speed (β = -0.09***; 95%-CI: [-0.15, -0.03]; Figure 3b). The results are robust when we control for age and gender.[14] For easier interpretation, we converted the model coefficients from the log scale to a percentage scale. Fraudulent responses increased deviation by 16% and decreased speed by 9%. These results are consistent with H1a and H2a.

Discussion

Using an established cheating task, Study 1 demonstrates that fraud significantly increased mouse-movement deviation and decreased mouse-movement speed. Nevertheless, this study has the following limitations: (1) to maximize internal validity, we used an established cheating task (Gino et al., 2010); however the task could be perceived as highly artificial, thus it is unclear whether the results would hold in a more realistic scenario; (2) the results of only one study are difficult to generalize; (3) the dot task required relatively short mouse movements, so it is unclear whether the results would hold in scenarios where more elaborate movements are needed; and (4) Study 1 focused on a binary fraud outcome, which did not allow us to test our hypotheses on the extent of fraud. Thus, in Study 2, we focused on external validity using a more realistic task that allowed the extent of fraud to vary.

Study 2: Insurance Task

Study 2 aimed to extend the results of Study 1 using a task with greater organizational implications and to test Hypotheses 1b and 2b, which relate to how the extent of fraud influences deviation and speed. In Study 2, participants had to file several auto insurance claims, a common online task in which people frequently commit fraud (Dionne & Gagné, 2002).

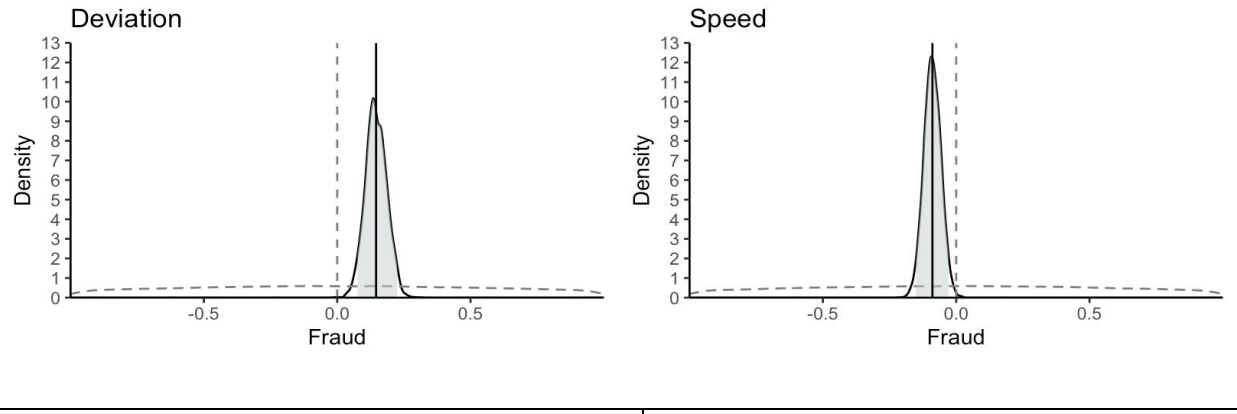
[13] CI refers to "credible interval." Significance levels: *90%, **95%, ***99% represent ranges of the CI, where it does not contain 0.

[14] Deviation: β = 0.15***; 95%-CI: [0.07, 0.22]; Speed: β = -0.09***; 95%-CI: [-0.15, -0.03]. Results remained robust when we removed detrimental errors. To account for task difficulty, we added "scenario" as a random effect to our statistical models. Results remained robust. Deviation: β = 0.15***; 95%-CI: [0.07, 0.23]; Speed: β = -0.09***; 95%-CI: [-0.16, -0.02].

Table 5. The Effect of Fraud on Deviation and Speed (Study 1)

	DV: Deviation		DV: Speed	
	Estimate	**95%-CI**	**Estimate**	**95%-CI**
Fixed effects				
Fraud (beneficial error = yes)	.15***	[.07, .22]	-.09***	[-.15, -.03]
Intercept	1.48***	[1.43, 1.53]	1.00***	[.94, 1.06]
Random effects				
σ residual	.39***	[.37, .41]	.31***	[.30, .32]
σ participant	.26***	[.23, .31]	.32***	[.28, .37]
N observations	1,150		1,150	
N participants	115		115	
R^2	.31		.51	
WAIC	1199.5		684.0	

Note: Results based on linear mixed-effects regressions with a level-two random effect (on the participant level). Significance levels: *90%, **95%, ***99% represent the ranges of the CI, where it does not contain 0 (derived from Ho et al., 2017); R^2 based on Gelman et al. (2019).



(a) Effect on deviation | **(b) Effect on speed**

Note: The dotted distributions show the prior distributions; the solid distributions show the posterior distributions, that is, our estimation results.

Figure 3. Posterior Distributions of the Fraud Coefficient (Study 1)

Materials and Procedure

We asked participants to complete five scenarios of claiming car damages using an online damage report form (see Table 6 for the scenarios; we randomized the order of the scenarios to control for learning and sequence effects).[15] While participants completed the scenarios, we recorded mouse movements using the same JavaScript code and web service as in Study 1. We calculated deviation and speed analogously to Study 1.

We incentivized fraud as follows: At the beginning of each scenario, participants received 2,000 coins as play money; we informed them that the deductible of their insurance contract was 600 coins. Because contracts with deductibles are sometimes perceived as unfair (Miyazaki, 2009), we expected some participants to inflate damages and commit fraud to cover their deductible and increase their final "wealth" (see Appendix A for the instructions).[16] We paid a variable bonus based on their final wealth on a randomly selected scenario at a rate of 10 pence (approx. US¢13)/100 coins. Thus, claiming more than the presented damage resulted in a higher payout.

[15] We pretested the scenarios in a pilot study (see Hibbeln et al., 2014).

[16] In addition, we introduced consequences of being detected to encourage people to provide more thoughtful and realistic inputs. To test if the level of punishment and the likelihood of being caught influence the extent of fraud, we randomly assigned participants to one of two conditions: a low-punishment condition with a low probability of getting caught (400 coins/10%) and a high-punishment condition with a high probability of getting caught (2,000 coins/50%). The low-punishment/low-probability group reported slightly more damages (3.41 vs. 3.25); as the difference was not significant (-0.16; $p = 0.33$), we pooled the data.

Table 6. Overview of Scenarios		
Scenario #	**Repair costs (in coins)**	**Number of accident damages**
1	400	1
2	800	2
3	1,200	3
4	1,600	4
5	2,000	5

Our damage report form required mouse inputs with longer and more complex movements than in Study 1. Participants had to mark the damage locations on the car's rear end, as shown in Figure 4. We started tracking mouse movements when each scenario was displayed and stopped when the participant made a submission. No other inputs were needed for this task. We conducted a laboratory study to pretest the scenarios and to conduct an initial test of our hypotheses. Given that our main study was conducted online and situational factors were thus uncontrollable, this laboratory study with its controlled setting minimized situational influences.[17]

Participants

Using Prolific, we recruited 150 participants who were at least 18 years of age from the United States. The mean age was 35.3 years, and 44.5% of participants were women. We paid £1 for a 10-minute task (equivalent to a £6/US$8 hourly wage), in addition to the variable bonus. In line with Study 1, we excluded participants whose completion time was three standard deviations longer than the average (Freeman & Dale, 2013) and anyone who accessed the study on a mobile device. We also excluded observations of participants who reported fewer damages than presented,[18] resulting in a final sample size of 110 participants with 550 observations (five scenarios per participant).

Data and Variables

Mouse variables: We calculated deviation and speed variables analogously to Study 1. Note that the deviation measure from Study 1 accounts for differences in task size. This is an important distinction: Committing fraud in Study 2 required more mouse movements because participants had to click additional damage locations to commit fraud. The method for calculating deviation in Study 1 provided an unbiased estimate of deviation, that is, unaffected by the requirement to click on more locations to report more damages. In other words, our calculation normalizes both deviation and speed to make them independent of the distance traveled.

Fraud: In each scenario, participants could choose to commit fraud by claiming more damages than presented (see Figure 5). We operationalized fraud as a binary variable and considered any scenario in which more damages than presented were claimed as fraud. We used the "number of additional damages claimed" to operationalize the extent of fraud.

Model Specification and Estimation

We used the same model specification as in Study 1. For the priors, we took advantage of Bayesian methods to include prior knowledge in the estimation. From the laboratory study (see Footnote 17), we had already gained knowledge about the distribution of the parameters: Deviation is positively correlated (0.19) and speed is negatively correlated (-0.20) with fraud. We used this distributional knowledge to derive priors as follows: For the model estimating the effect of fraud on *deviation*, we used $\alpha \sim N(1.87, 0.05)$ for the intercept, $\beta \sim N(0.19, 0.10)$ for the effect of fraud on deviation, $N(0.42, 0.03)$ for population-level errors, and $N(0.16, 0.06)$ for participant-specific errors. For the model estimating the effect of fraud on *speed*, we used $\alpha \sim N(1.10, 0.04)$ for the intercept, $\beta \sim N(-0.20, 0.09)$ for the effect of fraud on speed, $N(0.42, 0.03)$ for population-level errors, and $N(0.13, 0.06)$ for participant-specific errors.[19]

[17] For the lab study, we used the same insurance task as described in Study 2. Thirty-seven participants completed all five scenarios (170 observations). Of the 170 valid observations, 32 (19%) were fraudulent. The results show that fraudulent decisions (i.e., when participants claimed more damages than presented) increased mouse-movement deviation (β = 0.19; 95%-CI: [0.01, 0.38]) and decreased mouse-movement speed (β = -0.20; 95%-CI: [-0.38, -0.02]). Hence, the results are in line with Study 1 and Study 2. See Hibbeln et al. (2014) for further details.

[18] Reporting less damage than was presented means that participants harmed themselves (receiving a lower payout), indicating that the participants did not understand the instructions or did not complete the task carefully.

[19] In a robustness check, we also used uninformative priors. Results remain robust.

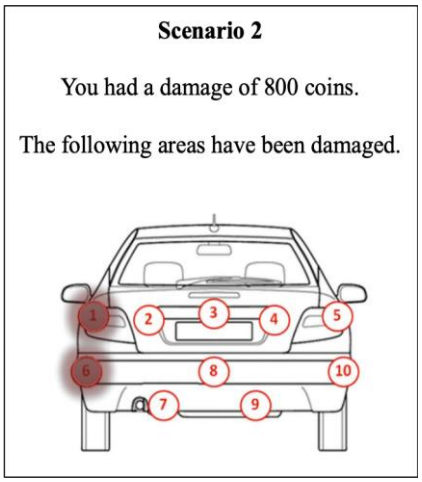


Figure 4. Sample Scenario (repair costs: 800; number of damage locations: 2)

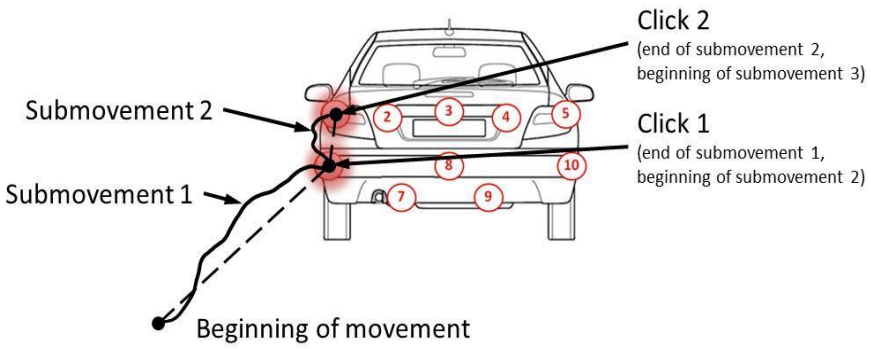


Figure 5. Example of Submovements

Results

Table 7 shows the summary statistics. Although we presented an average of 3.00 damages across five scenarios, participants claimed, on average, 3.29 damages. Of the 550 valid observations, 101 (18%) were fraudulent.

Table 8 shows that (1) fraudulent decisions—i.e., when participants claimed more damages than presented—increased mouse-movement deviation (β = 0.40***; 95%-CI: [0.28, 0.53]; Figure 6a), and (2) fraudulent decisions decreased mouse-movement speed (β = -0.35***; 95%-CI: [-0.45, -0.24]; Figure 6b).

Because nonfraudulent and fraudulent responses differed for age and gender, we conducted several robustness checks; further, in line with Hochman et al. (2016), we accounted for task difficulty; the main results remained qualitatively similar.[20] For easier interpretation, we converted the model coefficients from the log scale to a percentage scale.

[20] We ran several models to check for the robustness of the results. In addition to age and gender, we controlled for the number of cars owned, experience with computers, and experience with filing insurance claims. The results remained robust. Deviation: β = 0.41***; 95%-CI: [0.28, 0.53]; Speed: β = -0.36***; 95%-CI: [-0.47, -0.25]. To account for task difficulty, we added "scenario" as a random effect to our statistical models. The results remained robust. Deviation: β = 0.47***; 95%-CI: [0.36, 0.59]; Speed: β = -0.41***; 95%-CI: [-0.50, -0.31]. We ran another model with the same prior specification as described in Study 1. Using those (uninformative) priors the results were even stronger—fraud increased deviation (β = 0.54*** with 95%-CI: [0.35, 0.73]) and simultaneously decreased speed (β = -0.44*** with 95%-CI: [-0.58, -0.29]).

Table 7. Descriptive Statistics (Study 2)

Panel A: Full dataset

Statistic	Observations	%/Mean	*SD*	Min.	Max.
Fraud					
Damages (presented)	550	3.00	1.42	1	5
Damages (reported)	550	3.29	1.59	1	10
Fraud	101	18%			
Mouse movements					
Deviation (normalized pixels)	550	8.60	14.07	1.00	222.58
Speed (normalized pixels/ms)	550	3.21	4.03	0.44	70.36

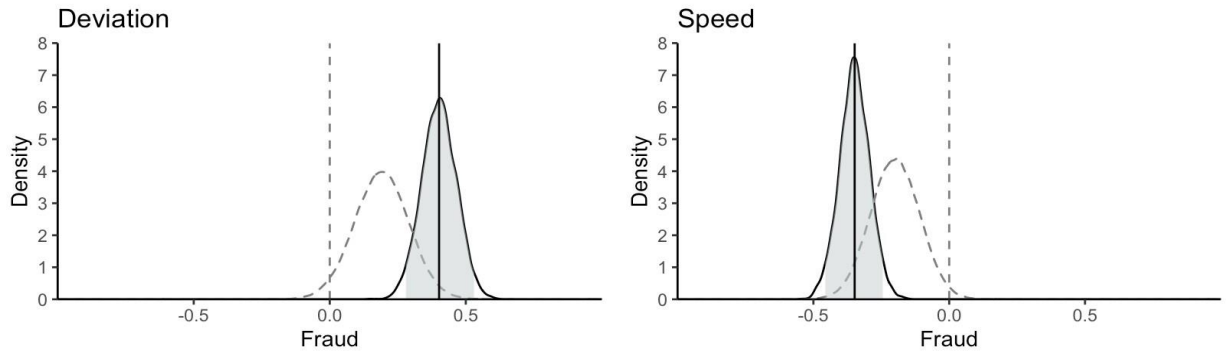
Panel B: Split by fraud

Statistic	Nonfraud		Fraud	
	Mean	*SD*	Mean	*SD*
Mouse movement				
Deviation (normalized pixels)	7.03	(5.03)	15.60	(30.02)
Speed (normalized pixels/ms)	3.48	(4.37)	1.98	(1.32)
Demographics				
Age	35.60	(12.67)	33.87	(10.85)
Gender (female)	49%		26%	
Observations	449		101	

Table 8. The Effect of Fraud on Deviation and Speed (Study 2)

	DV: Deviation		DV: Speed	
	Estimate	95%-CI	Estimate	95%-CI
Fixed effects				
Fraud (= yes)	.40***	[.28, .53]	-.35***	[-.45, -.24]
Intercept	1.74***	[1.68, 1.80]	1.05***	[.99, 1.10]
Random effects				
σ residual	.65***	[.63, .69]	.53***	[.50, .56]
σ participant	.26***	[.19, .32]	.25***	[.50, .31]
N observations	550		550	
N participants	110		110	
R^2	.14		.23	
WAIC	1333.5		990.6	

Note: Results based on linear mixed-effects regressions with a level-two random effect (on the participant level). Significance levels: *90%, **95%, ***99% represent the ranges of the CI, where it does not contain 0; R^2 based on Gelman et al. (2019).



(a) Effect on deviation | **(b) Effect on speed**

Note: The dotted distributions show the prior distributions (from our pilot); the solid distributions show the posterior distributions, that is, our estimation results.

Figure 6. Posterior Distributions of the Fraud Coefficient (Study 2).

Fraudulent responses increased mouse cursor deviation by 49% and decreased speed by 29%. These results are consistent with H1a and H2a. Our results show that when people provide fraudulent responses, their mouse movements exhibit greater deviation and lower speed. (Preliminary results of a further analysis show that mouse movements can also be used to predict fraud.)[21]

The design of the second study, where participants could claim a variable number of damages, allowed us to test H1b and H2b. Here, we hypothesized that the *extent of fraud* is positively correlated with the amount of deviation (H1b) and negatively correlated with mouse movement speed (H2b). To test these hypotheses, we reran our models, using the "number of additional damages claimed" as an independent fraud variable. The results support our hypotheses by showing that a larger extent of fraud increased deviation (β = 0.31***; 95%-CI: [0.21, 0.41]) and decreased speed (β = 0.27***; 95%-CI: [-0.34, -0.19]).[22]

Discussion

The results of Study 2, which used a realistic task requiring more elaborate mouse input, are in line with those of Study 1, which demonstrated that the dynamics of deciding to be fraudulent significantly increased mouse-movement deviation and decreased mouse-movement speed, thus lending further support to Hypotheses 1a and 2a. These results are in line with our pilot study, which found that when providing fraudulent responses, deviation increased by 21%, and speed decreased by 18%. It should be reiterated that both measures—deviation and speed—were standardized. Although in Study 2, engaging in fraudulent behavior required moving the mouse for longer distances (to click on more damage spots than presented), we normalized these distances. The results show that fraudulent inputs exhibited greater movement deviations and lower speeds after controlling for longer distances. Further, the results suggest that the influence on deviation and speed is related to the extent of fraud, confirming Hypotheses 1b and 2b. Together, the results show that fraud influences mouse movements in an organizationally relevant context.

General Discussion

Summary of Results

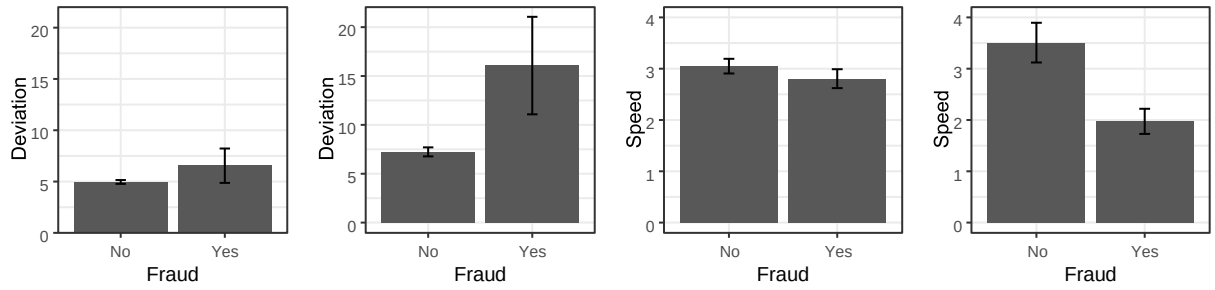
In this research, we drew on the RAM, cognitive dissonance/conflict, and CLT to hypothesize how the fraud decision-making process influences both mouse-movement deviation and speed. The results of two studies show that the decision to commit fraud systematically influences mouse-movement behavior. Committing fraud increased the deviation of mouse movements by 16% in Study 1 and 49% in Study 2 and simultaneously reduced the speed of mouse movements by 9% in Study 1 and 29% in Study 2 (see Figure 7). Further, the results show that not only the decision to commit fraud, but also the extent of fraud influences mouse movements, such that participants committing more extensive fraud displayed increased movement deviation and decreased movement speed.

Limitations and Directions for Future Research

Like all studies, our studies have certain limitations. First, our trace-data method used inputs from a computer mouse. Further research is needed to detect fraud committed using other input devices. Nevertheless, based on the strong relationship between cognitive processing and hand movements (Freeman et al., 2011), our results likely apply to other devices. (Note that voice input—as with Alexa—is beyond the scope of this research.)

[21] To assess the predictive accuracy of the model, we applied a leave-one-out cross-validation (LOO-CV) using the LOO package in R (Vehtari et al., 2017). We calculated both the area under the ROC curve (AUC) and the balanced accuracy. For our analyses, we distinguished between two practically relevant use cases in online contexts: *new visitors*, who are interacting with a particular website for the first time, and *returning visitors*, who have interacted with the website before. For returning users, we used existing user-specific data and considered the users' individual movement baselines. In contrast, for new users, we could only draw conclusions based on aggregated data of all visitors. Because we used a repeated-measures design in our study, we considered both cases by including subject-specific effects for returning visitors in our predictions (which we omitted for new visitors). For returning visitors, the AUC of 0.87 (with a 95%-CI of [0.83, 0.91]) indicates that the predictive accuracy of the model based on *deviation* and *speed* is considerably better than the line of no-discrimination (AUC = 0.50). Our model yields a balanced fraud-prediction accuracy of 80%. For new visitors, we omitted subject-specific effects and obtained significant results by achieving an AUC of 0.70 (with a 95%-CI of [0.65, 0.76]) and a balanced accuracy of 64%. As a robustness check, we compared other machine learning algorithms to a logistic regression (i.e., random forest, XGBoost, *K*-nearest neighbors, and neural network). Out of those, the logistic regression performed best in terms of AUC. Taken together, these results show that mouse-movement deviation and speed significantly improve the prediction of fraud; more data would be needed to confirm these preliminary results.

[22] As a robustness check, we further tested a different operationalization of the "extent of fraud" by considering the ratio of number of additional damages claimed / actual damages. Whereas our previous measure captures the "absolute extent of fraud," this alternative measure may better capture the "perceived extent of fraud" since the participants possibly have paid attention to the actual damages and used it as an anchor to evaluate the extent of fraud. The overall results remain robust: A larger "perceived extent of fraud" significantly increases deviation (β = 0.10***; 95%-CI: [0.07, 0.13]) and decreases speed (β = -0.08***; 95%-CI: [-0.10, -0.06]). We thank Reviewer 3 for this suggestion.



(a) Study 1: Deviation	(b) Study 2: Deviation	(c) Study 1: Speed	(d) Study 2: Speed

Note: deviation: normalized pixels; speed: normalized pixels/ms

Figure 7. The Effect of Fraud on Mouse-Movement Deviation (a & b) & Speed (c & d)

Our tracking technology ultimately captures the location of any pointing device on the screen, including a finger. Some of these devices can capture even more sophisticated trace data than a computer mouse. For example, a touchscreen can capture the force a user applies when interacting with the screen, and in-air sensors capture the *z*-dimension in addition to the *x*- and *y*-dimensions captured by a mouse. Future research should explore how human cognitive and affective states influence a range of trace-data types.

Second, we focused on measures of deviation and speed. Future research should attempt to identify other measures—such as acceleration, deceleration, click delays, and other behavioral characteristics—that could be influenced by changes in increased cognitive load or conflicting cognitions. Likewise, other fraud characteristics can affect mouse movement characteristics; for example, the potential gain, loss, or risk associated with the fraud decisions is likely to amplify or weaken the effects on mouse movements. Future research could further explore these influences by systematically manipulating different fraud-related variables.

Third, future research should examine a broader set of contexts and populations. We used two tasks to test the effects of fraud on mouse movements. Clearly, fraud exists in contexts beyond playing payout games or submitting insurance claims. Consistent with our theory, we would expect amplified results when the probability of being caught or suffering punishment is increased because people will be more likely to double-check, reconsider, hesitate, or even question their actions. Further, future research should more deeply examine ways to distinguish between first-time and returning visitors to optimize the way in which potentially fraudulent interactions can be detected.

Fourth, we predict that fraud may influence mouse-movement deviation and speed through cognitive conflict/dissonance and cognitive load. However, we did not explicitly measure cognitive conflict/dissonance and cognitive load because measuring them is challenging in online settings (we deemed self-report measures not useful for our current research). Past research has used PET (positron emission tomography) and pupil dilation to measure cognitive load (Jonides et al., 1997), as well as fMRI (functional magnetic resonance imaging) to identify activity in the anterior cingulate cortex (ACC), which can indicate cognitive conflict (Braem et al., 2017). Future research could use such neurophysiological tools to further unearth the underlying processes. Likewise, we hypothesized a linear relationship between the extent of fraud (as operationalized by the absolute number of additional damages reported) and mouse movement speed and deviation. Our results were robust to a different operationalization, but no operationalization of the extent of fraud that is based on observable actions will perfectly capture the underlying cognitive processes; thus, future research should test different operationalizations (e.g., by actually measuring the users' perceptions). Moreover, other factors are likely to moderate this relationship. For example, whereas some people might be very cautious in making their claims believable, others might be bolder, deliberating less about potential consequences; in other words, the same extent of fraud might *appear* to be large for some but small for others. Future research could take personality traits and situational factors into account, or attempt to tease out the mechanisms by, for example, carefully manipulating potential consequences.[23]

[23] We thank Reviewer 3 for this suggestion.

Theoretical Contributions

Using trace data such as mouse movements to understand hidden cognitive and emotional states is a growing research area in the information systems (IS) domain. Hibbeln et al. (2017) examined the effects of negative emotion on mouse movements, and Jenkins et al. (2019) used the RAM and CLT to explain why an orienting response results in changes to people's mouse movements when providing a predetermined response. Our study applied the RAM, CLT, and cognitive dissonance theory to explain how mouse movements can reveal the cognitive dynamics of a decision to be fraudulent in real time. When completing online forms using a mouse, if people decide to be fraudulent, they experience cognitive and moral conflict as they formulate a response. As they consider the different options to create a credible response, they interact with targets with actionable potential, resulting in larger deviations from the shortest mouse-movement path. Further, their brains compensate for a higher load on working memory by adjusting the speed of mouse movements. Thus, our research demonstrates that mouse movements can illuminate the fraud decision-making process—both relating to binary fraud decisions and the extent of fraud. Using two vastly different studies (in terms of task and setting), we demonstrate that these effects are robust in highly controlled and more realistic tasks.

We extend prior research and theory on mousing behavior in IS to realistic, nonsanctioned fraud scenarios. Sanctioned deception is often associated with low motivation of participants, lack of ecological validity, and low stakes (Buckley, 2012). Jenkins et al. (2019) found that mouse movements indicate deception in sanctioned experiments, but limited research has validated that fraud and deception influence mouse movements in nonsanctioned contexts. We extend prior work by collecting data on participants freely choosing to act fraudulently and show that mouse movements can provide insights into the fraud decision-making process in a nonsanctioned situation.

Further, we extend prior research by examining how the extent of fraud influences mouse-movement speed and deviation. In the fraud context, prior research has primarily focused on how a binary outcome of fraud influences behaviors (Jenkins et al., 2019)—whether or not the fraud occurred—irrespective of the extent of the fraudulent act. However, the *extent* of fraud can vary greatly, particularly in the insurance fraud context, and little research has investigated how the extent of fraud can be captured by analyzing behavioral traces. This paper addresses this gap by both theoretically explaining and empirically demonstrating how the extent of fraud influences the degree of changes to mouse-movement deviation and speed.

On a larger scale, our work contributes to evidence linking hand motions captured through mouse movements to various emotional and cognitive processes (Freeman et al., 2011; Hibbeln et al., 2017; Jenkins et al., 2019). What is particularly exciting is that these movements reflect both actual behavior and behavioral changes, measured within an information technology usage context. Our work suggests that analyzing trace data as an actual measure of usage could enrich other areas of IS research, in which perceptions of states and behaviors are measured post hoc to interaction, and methods based on trace data such as mouse movements can complement other methods. Clearly, this approach suggests numerous future research opportunities. Although all research methods have strengths and weaknesses (Dennis & Valacich, 2001), tracking mouse movements provides potential benefits that can enhance the insight obtained in various studies (Kieslich et al., 2020). For instance, early behavioral studies used observational case studies, which offer extensive behavioral information on limited cases. Unfortunately, case studies are labor intensive, time-consuming, and often not generalizable to broader populations. Survey studies have emerged to overcome some case method limitations, allowing researchers to quickly collect data from a broad population. Mouse tracking can be embedded into a survey or a broad range of experimental and observational tasks to capture rich observational traces to reveal the (hidden) cognitive processes of respondents answering questions or completing online tasks. In sum, mouse tracking is an exciting research method that captures rich observational trace data at scale.

To enable capturing such trace data at scale, we formulated a valid measure of mouse-movement deviation for natural online interactions. One commonly used measure of deviation in the previous research is the graded motor response analysis (Freeman & Dale, 2013; Jenkins et al., 2019; McKinstry et al., 2008). This assumes that all users have the same mouse-movement start and end points. This is not always applicable in natural online settings, as people's movements may have different beginning, intermediate, and ending points. We created a measure that is valid in these more natural interactions by averaging the deviation in each submovement. As this measure is normalized for interactions of different lengths, researchers and practitioners can compare deviations across different users.

Managerial Implications

Fraud is ubiquitous, with substantial social and organizational costs. As organizations move their interactions online, detecting potential fraud on online data collection forms is increasingly important. Mouse tracking is a low-cost and highly scalable detection method. Capturing mouse movements does not require special computer hardware; movements can be

collected in a web browser using JavaScript code. This paper provides a method for detecting theoretically sound and validated cues of fraud on online forms without substantial investment. Many of today's leading organizations are augmenting costly human-based fraud detection processes using machine learning algorithms (such as anomaly detection). Yet it is impossible to detect fraudulent behavior with complete accuracy, and such automation can fail due to poor input data, opaque algorithms, and other reasons. As current approaches are costly and prone to errors, organizations could complement existing fraud detection methods by analyzing the input of a process based on mouse movements. By doing so, they could derive a confidence score as an additional dimension and automatically flag suspicious behavior for further auditing using algorithms to identify data entry anomalies. Our procedure would make the costly auditing process more effective and efficient by allowing companies to focus scarce resources on the cases that are most likely to be fraudulent, potentially resulting in tremendous savings to organizations, customers, and society in general.

Beyond detecting and flagging potentially fraudulent user submissions, organizations could devise real-time follow-up questions or other inputs to increase result confidence. Early detection of potentially fraudulent cases could be used to diagnose issues as they surface. In case unusual inputs are detected, but mouse cursor movements are not indicative of fraud, a system could alert users of potentially erroneous inputs (both detrimental and beneficial errors). This could reduce costs for organizations and increase customer satisfaction if a user is alerted of a detrimental error.

Managers should note that our studies used a somewhat simple design, in that the pages required mouse input only, whereas many online forms require additional user actions, such as keyboard input or file uploads.[24] In production environments, companies seeking to utilize this methodology should devise input forms that allow for capturing sequences of mouse movements (as a separate page or as a section within a page); as with any user interface, there is no one-size-fits-all design/layout, and A/B testing should be used to assess how far the chosen layout affects mouse movement dynamics. In addition to analyzing mouse movements, companies could analyze keystroke dynamics to analyze and detect anomalies. A combination of approaches would likely increase the confidence in detecting cases that require further auditing; future research could test the benefits of using multiple sources of trace data to detect fraud. Likewise, we only used one layout/task in Study 2. Future research should test other layouts and tasks to test the robustness of our method in different settings.

While our method for detecting fraud is promising, we call on researchers and practitioners to follow best practices for the ethical use of behavioral data. For example, one issue is the potential to misclassify users; thus, it is important to account for individual differences. Likewise, we recommend that researchers and practitioners use some type of technique to baseline individuals in order to account for normal variations in behavioral data that are not associated with the fraudulent outcome.

In our research, we focused on the effects of cognitive dissonance/conflict and cognitive load on mouse movement characteristics in the context of fraud. However, people experience such dissonance/conflict or heightened cognitive load in situations beyond fraud. For example, in educational contexts, it is imperative not to overload learners; consequently, it is important to adapt material and presentation to provide optimal stimulation and reduce factors that could contribute to overloading learners. Computer-based training approaches could include our method to detect potential overload, possibly adapting content or presentation to maximize the learning outcome. For example, during test taking, mouse movements could be analyzed to discern whether test takers' mistakes can be attributed to lack of knowledge or extraneous factors leading to cognitive overload. Likewise, our method could be used to detect cognitive dissonance in various settings, such as in helping people make better choices (e.g., "digital nudging," Schneider et al. 2018), different approaches to behavior modification, and public health campaigns (such as vaccination campaigns). Thus, our method is likely to have a broad range of potential applications; obviously, future research should thoroughly test the efficacy of our method in different settings.

Conclusion

This research demonstrates how capturing and analyzing trace data can help obtain near-real-time insights into online fraud decisions. Building on cognitive- and neuroscience, we posited that the fraud decision-making process has physiological and psychological side-effects resulting in predictable changes in hand movements. These predictable changes can be captured by monitoring fine-grained trace data as a person uses a computer mouse. Our two studies support our hypotheses and suggest that mouse-movement traces exhibit meaningful differences when committing fraud and are influenced by the extent of fraud. Our findings have implications for creating scalable, cost-effective algorithms to detect potential fraud as it occurs in online contexts. Our method of analyzing actual user behaviors in real time has the potential to complement other behavioral methods, in the context of fraud and in a variety of other contexts and settings.

[24] We thank Reviewer 3 for pointing this out.

Acknowledgments

All authors contributed equally; the authorship order is reversed alphabetically. The authors thank Isabel Eilers and Viktoria Menge from the University of Braunschweig as well as Daniel Boekhoff and Phillip Lassen from Peaks & Pies GmbH for outstanding research assistance. They also thank participants of research seminars at Erasmus University Rotterdam, ESCP Business School Berlin, University of Cologne, University of Lüneburg, University of Mannheim, and the University of Munich for their valuable comments. They also gratefully acknowledge the support from the University of Liechtenstein. Finally, the authors thank the senior editor, associate editor, and the anonymous reviewers for their valuable comments on earlier versions of this manuscript.

References

Banholzer, N., Feuerriegel, S., Fleisch, E., Bauer, G. F., & Kowatsch, T. (2021). Computer mouse movements as an indicator of work stress: Longitudinal observational field study. *Journal of Medical Internet Research, 23*(4), Article e27121.

Berente, N., Seidel, S., & Safadi, H. (2019). Data-driven computationally intensive theory development. *Information Systems Research, 30*(1), 50-64.

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis, 20*(3), 351-368.

Braem, S., King, J. A., Korb, F. M., Krebs, R. M., Notebaert, W., & Egner, T. (2017). The role of anterior cingulate cortex in the affective evaluation of conflict. *Journal of Cognitive Neuroscience, 29*(1), 137-149.

Buckley, J. P. (2012). Detection of deception researchers needs to collaborate with experienced practitioners. *Journal of Applied Research in Memory and Cognition, 1*(2), 126-127.

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 6*(3), (203-242.

Bürkner, P. C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *R Journal, 10*(1), 395-411.

Carneiro, N., Figueira, G., & Costa, M. (2017). A data mining based system for credit-card fraud detection in e-tail. *Decision Support Systems, 95*), 91-101.

Carrión, R. E., Keenan, J. P., & Sebanz, N. (2010). A truth that's told with bad intent: An ERP study of deception. *Cognition, 114*(1), 105-110.

Coalition Against Insurance Fraud. (2021). *The impact of fraud.* https://insurancefraud.org/fraudstats/.

Crowe Global (2021). *Fraud costs the global economy over US$5 trillion.* https://www.crowe.com/global/news/fraud-costs-the-global-economy-over-us$5-trillion, 2021-05-23.

Dale, R., Kehoe, C., & Spivey, M. J. (2007). Graded motor responses in the time course of categorizing atypical exemplars. *Memory and Cognition, 35*(1), 15-28.

Dennis, A. R., & Valacich, J. S. (2001). Conducting experimental research in information systems. *Communications of the Association for Information Systems, 7*, 1-42.

Derrick, C., Douglas, Jenkins, J. L., & Nunamaker, J. F. (2011). Design principles for special purpose, embodied, conversational intelligence with environmental sensors (SPECIES) agents. *AIS Transactions on Human-Computer Interaction, 3*(2), 62-81.

Derrick, D. C., Meservy, T. O., Jenkins, J. L., Burgoon, J. K., & Nunamaker, J. F. (2013). Detecting deceptive chat-based communication using typing behavior and message cues. *ACM Transactions on Management Information Systems, 4*(2), 1-21.

Dionne, G., & Gagné, R. (2002). Replacement cost endorsement and opportunistic fraud in automobile insurance. *Journal of Risk and Uncertainty, 24*(3), 213-230.

FBI (2021). *Insurance fraud*. Retrieved March 23, 2001, from https://www.fbi.gov/stats-services/publications/insurance-fraud.

Freeman, J. B., & Ambady, N. (2011). When two become one: Temporally dynamic integration of the face and voice. *Journal of Experimental Social Psychology, 47*(1), 259-263.

Freeman, J. B., & Dale, R. (2013). Assessing bimodality to detect the presence of a dual cognitive process. *Behavior Research Methods, 45*(1), 83-97.

Freeman, J. B., Dale, R., & Farmer, T. A. (2011). Hand in motion reveals mind in motion. *Frontiers in Psychology, 2*(59), https://doi.org/10.3389/fpsyg.2011.00059

Gelman, A., Goodrich, B., Gabry, J., & Vehtari, A. (2019). R-squared for Bayesian regression models. *The American Statistician, 73*(3), 307-309.

Gino, F., Norton, M. I., & Ariely, D. (2010). The counterfeit self: The deceptive costs of faking it. *Psychological Science, 21*(5), 712-720.

Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., & Jennions, M. D. (2015). The extent and consequences of p-hacking in science. *PLoS Biology, 13*(3), e1002106.

Hehman, E., Stolier, R. M., & Freeman, J. B. (2015). Advanced mouse-tracking analytic techniques for enhancing psychological science. *Group Processes and Intergroup Relations, 18*(3), 384-401.

Hibbeln, M., Jenkins, J. L., Schneider, C., Valacich, J. S., & Weinmann, M. (2014). Investigating the effect of insurance fraud on mouse usage in human-computer interactions. In *Proceedings of the International Conference on Information Systems*.

Hibbeln, M., Jenkins, J. L., Schneider, C., Valacich, J. S., & Weinmann, M. (2017). How is your user feeling? Inferring emotion through human-computer interaction devices. *MIS Quarterly, 41*(1), 1-21.

Ho, Y.-C. C., Wu, J., & Tan, Y. (2017). Disconfirmation effect on online rating behavior: A structural model. *Information Systems Research, 28*(3), 626-642.

Hochman, G., Glöckner, A., Fiedler, S., & Ayal, S. (2016). "I can see it in your eyes": Biased processing and increased arousal in dishonest responses. *Journal of Behavioral Decision Making, 29*(2-3), 322-335.

Hunter, M. (2015). Tax-refund fraud to hit $21 billion, and there's little the IRS can do. CNBC. https://www.cnbc.com/2015/02/11/tax-refund-fraud-to-hit-21-billion-and-theres-little-the-irs-can-do.html.

Ihaka, R., & Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics, 5*(3), 299-314.

Jenkins, J. L., Proudfoot, J. G., Valacich, J. S., Grimes, G. M., & Nunamaker, J. F. (2019). Sleight of hand: Identifying concealed information by monitoring mouse-cursor movements. *Journal of the Association for Information Systems, 20*(1), 1-32.

Jonides, J., Schumacher, E. H., Smith, E. E., Lauber, E. J., Awh, E., Minoshima, S., & Koeppe, R. A. (1997). Verbal working memory load affects regional brain activation as measured by PET. *Journal of Cognitive Neuroscience, 9*(4), 462-475.

Kieslich, P. J., Schoemann, M., Grage, T., Hepp, J., & Scherbaum, S. (2020). Design factors in mouse-tracking: What makes a difference?" *Behavior Research Methods, 52*(1), 317-341.

Kruschke, J. K., Aguinis, H., & Joo, H. (2012). The time has come: Bayesian methods for data analysis in the organizational sciences. *Organizational Research Methods, 15*(4), 722- 752.

Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods, 44*(1), 1-23.

McElreath, R. (2020. *Statistical rethinking: A Bayesian course with examples in R and Stan.* CRC Press.

McKinstry, C., Dale, R., & Spivey, M. J. (2008). Action dynamics reveal parallel competition in decision making. *Psychological Science, 19*(1), 22-24.

Mertens, W., & Recker, J. (2020). New guidelines for null hypothesis significance testing in hypothetico-deductive IS research. *Journal of the Association for Information Systems, 21*(4), 1072-1102.

Meyer, D. E., Abrams, R. A., Kornblum, S., Wright, C. E., & Smith, J. E. (1988). Optimality in human motor performance: Ideal control of rapid aimed movements. *Psychological Review, 95*(3), 340-370.

Miyazaki, A. D. (2009). Perceived ethicality of insurance claim fraud: Do higher deductibles lead to lower ethical standards? *Journal of Business Ethics, 87*(4), 589-598.

Morrow, S., & Maynard, N. (2021). *Online payment fraud research report.* Juniper Research. https://www.juniperresearch.com/researchstore/fintech-payments/online-payment-fraud-research-report.

Nuñez, J. M., Casey, B. J., Egner, T., Hare, T., & Hirsch, J. (2005). Intentional false responding shares neural substrates with response conflict and cognitive control. *NeuroImage, 25*(1), 267-277.

Nuzzo, R. (2014). Statistical errors. *Nature, 506*, 150-152.

Olson, J. A., & Raz, A. (2021). Applying insights from magic to improve deception in research: The Swiss cheese model. *Journal of Experimental Social Psychology, 92*, Article 104053.

Plamondon, R., & Alimi, A. (1997). Speed/accuracy trade-offs in target-directed movements. *Behavioral and Brain Sciences, 20*(2), 279-349.

Schneider, C., Weinmann, M., & vom Brocke, J. (2018). Digital nudging: Guiding online user choices through interface design. *Communications of the ACM, 61*(7), 67-73.

Song, J.-H., & Nakayama, K. (2008). Target selection in visual search as revealed by movement trajectories. *Vision Research, 48*(7), 853-861.

Thomas, J., & Biros, D. (2011). A conceptual model of real world, high stakes deception detection. In *Proceedings of the Annual Hawaii International Conference on System Sciences*.

Unsworth, N., & Engle, R. W. (2005). Individual differences in working memory capacity and learning: Evidence from the serial reaction time task. *Memory and Cognition, 33*(2), 213- 220.

Valacich, J. S., Jenkins, J. L., Hariri, S., & Howie, J. (2013). Identifying insider threats through monitoring mouse movements in concealed information tests. In *Proceedings of the Hawaii International Conference on System Sciences.*

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing, 27*(5), 1413-1432.

Welsh, T. N., & Elliott, D. (2004). Movement trajectories in the presence of a distracting stimulus: Evidence for a response activation model of selective reaching. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 57*(6), 1031-1057.

About the Authors

Markus Weinmann is a professor of business analytics at the University of Cologne and an affiliate member at the Erasmus Research Institute of Management (ERIM), Erasmus University Rotterdam. His research concerns the opportunities and challenges of the digital economy—in particular, using data to infer how people judge, decide, and behave on digital platforms. Areas of application include crowdfunding, crowdsourcing, and online ratings. Dr. Weinmann's research has been published in leading journals, including *Information Systems Research*, *Management Science*, *Marketing Science*, and *MIS Quarterly*.

Joseph (Joe) S. Valacich is the Munsinger Professor of Entrepreneurship and Innovation in the MIS Department at the University of Arizona. His primary research focus is on human-computer interaction (HCI), cybersecurity, and e-business. Dr. Valacich has published more than 110 scholarly articles in numerous prestigious journals, including *Academy of Management Journal*, *Accounting Review*, *Information Systems Research*, *Journal of Applied Psychology*, *Journal of the Association for Information Systems*, *Journal of Management Information Systems*, *Management Science*, *MIS Quarterly*, *Organizational Behavior and Human Decision Processes*, and many others.

Christoph Schneider is an assistant professor in the Department of Operations, Information, and Technology at IESE Business School. His research interests include human-computer interaction, electronic commerce, and computer-mediated collaboration. His research has appeared in leading journals including *Information Systems Research*, *Management Science*, and *MIS Quarterly*.

Jeffrey L. Jenkins is an associate professor of information systems at Brigham Young University. He has developed, validated, and patented algorithms to detect cognitive and emotional reactions using human-computer interaction devices. He co-founded Neuro-ID, Inc.—a university tech startup based on his research. His work has been published in the top information systems journals, including the *Journal of Management Information Systems*, *MIS Quarterly*, *Information Systems Research*, *Journal of the Association of Information Systems*, and others.

Martin T. Hibbeln is a professor of finance in the Mercator School of Management at the University of Duisburg-Essen, Germany. His research interests include financial data analytics, financial risk management, and financial technology. Dr. Hibbeln's research has been published in leading finance and information systems journals, including *Journal of Financial Intermediation*, *Journal of Banking & Finance*, *Journal of Risk and Insurance*, and *MIS Quarterly*.

Appendix A

Study Instructions

Study 1: Flexible Dot Task

Your Task

- You will see a box with multiple red dots in it (see example below [Figure 1 in main paper]).
- You will only see the dots for about 1 second, and once they disappear your task is to indicate whether there were more dots on the right side of the square or on the left side of the square by clicking on the left or right button (note that sometimes a dot will be on the line between the two parts).
- This study is expected to take approximately 10 minutes to complete.
- Your participation is completely voluntary and you may quit the study at any time.

Payment

- To motivate your efforts, we will pay you based on the buttons you click (not based on whether you are correct or not)!
- Because most people can more easily estimate the number of dots on the left side, we will pay you 0.5 pence for each trial that you answer as having more dots on the left side, and 5 pence for each trial that you answer as having more dots on the right side (see illustration).

Study 2: Insurance Task

Situation

- You have an **initial wealth of 2,000 coins**.
- For your vehicle, you purchased an **insurance policy with a deductible of 600 coins**. In other words, if the damage to your car was 1,000 coins, you would be responsible for paying the first 600 coins, and the insurance company would pay the remaining 400 coins.
- Imagine that recently, you had an **accident** when backing up in a small parking spot and **damaged the rear end of your car**.
- Now, you have to **file an insurance claim** on the **insurance company's website**.

Example

- Before the accident, you have an initial wealth of 2,000 coins.
- The repair of the damage to your vehicle costs 1,000 coins; thus, your wealth is reduced to 1,000 coins.
- Assume you file a claim of 1,300 coins. Given your deductible of 600 coins, the insurance would pay you 700 coins (1,300 coins – 600 coins).
- Your final wealth after receiving payment from the insurance company would be 1,700 coins.

Your Task

- You will be presented with 5 different scenarios for which you will have to file insurance claims. In each scenario, the damage to your vehicle is different.
  - In each scenario, you have an initial wealth of 2,000 coins. The higher your claim, the higher the payment from the insurance company. In other words, your final wealth depends on the amount you claim from the insurance company.
  - *Group 1:* The insurance company screens 10% of all insurance claims. When the insurance company detects cheating by the insured, this leads to a punishment of 400 coins.
  - *Group 2:* The insurance company screens 50% of all insurance claims. When the insurance company detects cheating by the insured, this leads to a punishment of 2,000 coins.