A Multi-Experimental Examination of Analyzing Mouse Cursor Trajectories to Gauge Subject Uncertainty

Full Paper

Jeffrey L. Jenkins, Ross Larsen, Robert Bodily, Daniel Sandberg, Parker Williams, Steve Stokes, Spencer Harris
Brigham Young University
jjenkins@byu.edu

Joseph S. Valacich
University of Arizona
valacich@arizona.edu

Abstract

Providing information online is pervasive in human-computer interactions. While providing information, people may deliberate their responses. However, organizations only receive the end-result of this deliberation and therefore have no contextual information surrounding the response. One type of contextual information includes knowing people's response uncertainty while providing information. Knowing uncertainty allows organizations to weigh responses, ask follow-up questions, provide assistance, or identify problematic instructions or responses. This paper explores how mouse cursor movements may indicate uncertainty in a human-computer interaction context. Specifically, it hypothesizes how uncertainty on multiple-choice questions influences a mouse-movement statistic called area-under-the-curve (AUC). We report the result of two studies that suggest that AUC is higher when people have moderate uncertainty about an answer than if people have high or low uncertainty. The results suggest a methodology for measuring uncertainty to facilitate multi-method research and to assess data in a pragmatic setting.

Keywords

Mouse cursor tracking, uncertainty, area-under-the-curve, adaptive computers

Introduction

Providing information online is a catalyst for countless processes. For example, job applications, product reviews, online questionnaires, credit applications, and educational quizzes are only a few examples of the diverse set of online systems that gather important information. In many contexts, people rehearse their responses before submitting information (Dennis et al. 2008). Organizations, however, often only receive the end result of that rehearsal process. In other words, one can see the final information provided, but has little contextual information about the cognitive processes people engaged in when determining their responses. For example, a person might respond to an online questionnaire, but may be highly uncertain in their answer. This uncertainty may be caused by many reasons, including a lack of knowledge to answer a question, low response confidence, unclear instructions, or even hesitancy to provide truthful information. However, this contextual information is often lost in an online context (Derrick et al. 2013), and thereby limits organizations' abilities to ask follow-up questions, weigh responses, provide clarifications, or otherwise gain insight to improve data quality.

In this paper, we propose that monitoring mouse-cursor movements can provide insight into people's level of uncertainty when providing information online during human-computer interactions. Mouse cursor movements have been suggested to provide "high-fidelity, real-time motor traces of the mind [that] can reveal 'hidden' cognitive states that are otherwise not availed by traditional measures" (Freeman et al. 2011 p. 2). Past research has shown that mouse cursor tracking can provide insight into a variety of human states, including attention (Song and Nakayama 2006), emotion (Grimes et al. 2013), and cognitive conflict (McKinstry et al. 2008). We extend this research to explore how mouse-cursor movements provide insight into people's level of uncertainty.

To begin the investigation of measuring uncertainty through mouse-cursor movements in human computer interactions, we limit the scope of this paper to people's uncertainty while answering multiple-choice questions in an online context. Multiple-choice questions are used extensively to gather information from people online and require people to move the mouse cursor to select an answer. Such questions therefore provide rich data for analyzing mouse-cursor movements and therefore an appropriate initial context for investigating uncertainty. In summary, we address the following research question: how do mouse-cursor movement characteristics correlate with uncertainty as people answer multiple-choice questions online?

Literature Review

Mouse cursor tracking as a scientific methodology was originally explored as a cost-effective alternative to eye tracking to denote where people devote their attention in a human-computer interaction context (Byrne et al. 1999; Chen et al. 2001; Guo and Agichtein 2010). For example, research has shown that eye gaze and mouse-cursor movement patterns are highly correlated with each other (Chen et al. 2001; Guo and Agichtein 2010; Pan et al. 2004). When scanning search results, the mouse often follows the eye and marks promising search hits (i.e., the mouse pointer stops or lingers near information), suggesting where people devote their attention (Rodden et al. 2008). Likewise, people often move their mouse while viewing web pages, suggesting that the mouse may indicate where people focus their attention (Mueller and Lockerd 2001). In selecting menu items, the mouse often tags potential targets (i.e., hovers over a link) before selecting an item (Cox and Silva 2006). Monitoring where someone clicks can also be used to assess the relevance of search results (Huang et al. 2011). Finally, by continuously recording mouse position, the awareness, attraction, and avoidance of content can be assessed (e.g., people avoiding ads, not looking at text because of frustration, or struggling to read the text) (Navalpakkam and Churchill 2012). Consequently, mouse tracking is often applied as a usability assessment tool for visualizing mouse-cursor movements on webpages (Arroyo et al. 2006; Lagun and Agichtein 2011), and to develop heat maps indicating where people devote their attention (Atterer and Lorenzi 2008; Lettner and Holzmann 2012).

As the ability for more fine-grained measurement and analysis of mouse-cursor movements improved, research expanded the use of mouse cursor tracking to explore a more diverse set of neuromotor and psychological responses. In a concise review of mouse tracking literature, Freeman and Ambady (2011 p. 1) suggest that the "movements of the hand…offer continuous streams of output that can reveal ongoing dynamics of processing, potentially capturing the mind in motion with fine-grained temporal sensitivity." Accordingly, hundreds of recent studies have chosen mouse tracking as a methodology for studying various cognitive and emotional processes. For brevity, relevant studies are summarized in Appendix A.

Uncertainty is likely to impact both how people devote their attention and cognition. To explain how uncertainty influences people's attention and cognition while answering multiple-choice questions, we draw on the Metamemory framework (Flavell and Wellman 1975). The framework explains the process people engage in when answering a multiple-choice question (Figure 1). It posits that after people are given a question, they make a Feeling of Knowing judgment. Feeling of Knowing is the degree to which people believe they know or could find the answer after reading the question but before attempting to answer it. If people feel that they do not know and cannot find the answer, they immediately choose/ guess an answer and / or proceed to the next question (the right path at the first decision in Figure 1). However, if people feel that they know or could find the answer, they proceed to a search strategy stage (the left path at the first decision in Figure 1) (Nelson 1990).

In the search strategy stage, people look for a possible answer. If a potential answer is found, people subconsciously assess how confident they are that their answer is correct. If they are confident, they choose the answer and move on to the next question. However, if people are uncertain, they assess their willingness to search longer. If people are willing to search longer, they return to the Feeling of Knowing stage and repeat the process of searching for a potential answer and assessing their confidence. This cycle repeats until the Feeling of Knowing exceeds the threshold for searching longer. The search loop will continue until their confidence is high enough to answer the question, or until people's willingness to search is depleted and they consequently choose an answer and / or move to the next question (Nelson 1990). This is consistent with research on decision heuristics and the unconscious mind explaining the trade-off between effort and satisfactory responses. Namely, a trade-off exists between effort and the outcome, such that when the effort exceeds the utility of finding a better solution, people will satisfy with the suboptimal outcome (Kahneman 2011; Kahneman and Tversky 1979; Tversky and Kahneman 1981).

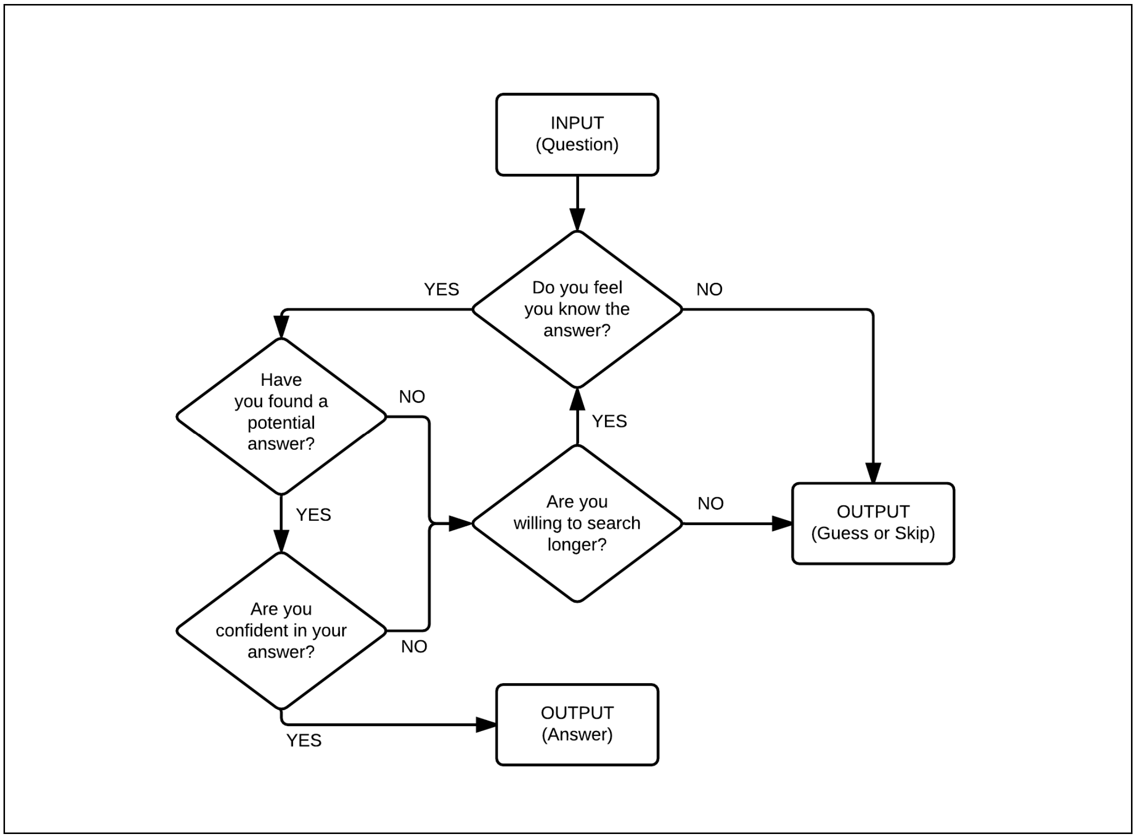

Figure 1. The Retrieval Stage of the Metamemory Framework (Nelson 1990)

Thus, the Metamemory framework suggests at least three different outcomes that may occur depending on the level of uncertainty: a) a person knows the answer and answers quickly, b) a person is not sure of the answer but searches for a solution, and c) a person does not know the answer at all and guesses (or skip the question).

First, if people think they know the answer, they will have a high Feeling of Knowing and will continue to the left (see Figure 1) to search for the answer. They will find an answer that they are confident is correct, and will immediately answer the question. Second, if people are not sure they know the answer but think they can figure it out, they may have a medium level of Feeling of Knowing. They will continue to the left as well. As they search and find answers, they will assess whether they think the answer is a plausible one and if they are confident in their answer. If they answer no to either of those questions, they move on to decide whether they are willing to search longer for an answer. The amount of time that they spend in this search loop depends on whether the Feeling of Knowing exceeds the threshold for searching longer. The greater the Feeling of Knowing and the lower the confidence in their answers will result in a greater amount of time spent in the search loop. Finally, if people do not know an answer at all, they will have low Feeling of Knowing and proceed to the right on the first decision to guess or to skip the question.

Based on these three possible outcomes, we next hypothesize how uncertainty will influence people's attentional and cognitive states and thereby mouse-cursor movements.

Theory

In this section we hypothesize how uncertainty will influence a mousing statistic called area-under-the-curve (AUC) while people answer multiple-choice questions. AUC is a measure of how much people deviate from the most direct path of responding to a question (Freeman and Ambady 2010). For example, the location of the mouse when a question is displayed can be conceptualized as a starting point (x-,y-coordinate) for an interaction. The answer that a person ultimately picks can be conceptualized as an ending point of the interaction. Between the starting and ending points, a straight line can be drawn. In mouse-cursor tracking literature, this straight line is often referred to as the *idealized response trajectory* (Hehman et al. 2014). AUC refers to the geometric area between the actual mouse trajectory and the idealized response trajectory (Figure 2); it is a measure of total deviation from the idealized response trajectory.

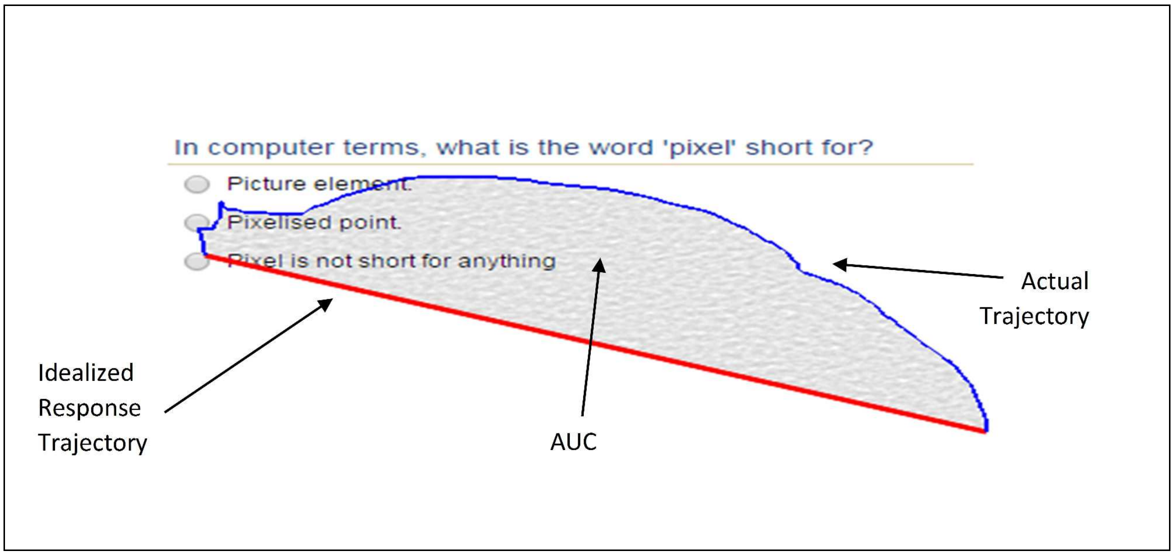

Figure 2. Area-Under-The-Curve

We hypothesize that AUC will be greater for moderate uncertainty than for high uncertainty and low uncertainty. Under conditions of moderate uncertainty, people engage in a search process for the correct answer. As a part of this process, they may find a potential answer, assess their confidence in their answer, consider another possible answer if they are still uncertain, and continue this process until an acceptable answer is chosen. As people consider different possible answers, it will cause their mouse movements to deviate from the idealized response trajectory for two reasons: attentional shifts and cognitive conflict. First, as people shift their attention from answer to answer and repeatedly to the question text in search for a possible answer, it will cause people's mouse movements to deviate from the most direct path. Namely, the mouse often follows the eyes vertically, less frequently follows the eyes horizontally, and marks promising results, suggesting where people devote their attention (Rodden et al. 2008). Thus, as people consider different answers and information, their mouse may deviate toward that information and away from the most direct path.

Second, on a more subtle level, the cognitive conflict from deciding between multiple answers causes people to deviate from their intended movement trajectory and therefore also the idealized response trajectory. All stimuli with actionable potential that capture a person's attention will prime movement responses (Georgopoulos 1990; Song and Nakayama 2006; Song and Nakayama 2008; Tipper et al. 1997). In our context, stimuli with actionable potential include possible answers because they are possible destinations that compete for one's attention. To *prime* a behavior refers to subconsciously programming a response (transmitting nerve impulses to the hand and arm muscles) to move toward or away from the stimulus. This causes the hand to deviate from its intended movement, as the observed hand movement is a product of all primed responses, both intended and non-intended (Welsh and Elliott 2004).

As different answers repeatedly capture a person's attention and cause cognitive conflict under conditions of moderate uncertainty, people will deviate from the idealized response trajectory. However, in conditions of low uncertainty, people will not focus their attention as broadly and will experience less cognitive conflict. They will move with less distraction to choose an answer. As such, we hypothesize that they will deviate less from the idealized response trajectory and have lower AUC. In summary,

H1: Moderate uncertainty will result in greater AUC than low uncertainty.

Likewise, we predict that moderate uncertainty will result in greater AUC than high uncertainty. Under conditions of high uncertainty, people will experience low Feeling of Knowing. If it is low enough (i.e., no Feeling of Knowing), they will not even enter a process to search for a correct answer, and will guess or skip the question. However, even under conditions of low Feeling of Knowing, people's search stage may be abrupt as they do not have confidence in any answer. Compared to people with moderate uncertainty who may elaborate in a search stage to find a correct answer, people with high uncertainty will have a shorter search stage. As such, their mouse movements will deviate less from a direct path of selecting an answer, and therefore have lower AUC. In summary,

H2: Moderate uncertainty will result in greater AUC than high uncertainty.

We also predict that people with high uncertainty will have greater AUC than people with low uncertainty. Unless people with high uncertainty have no Feeling of Knowing (in which case they may immediately guess or skip the question), they will make some attempt at answering the question. This may occur through attempting to eliminate possible answers, derive cues from the text, or use other techniques. As this information deviates one's attention from a single answer and causes cognitive conflict, we would expect the mouse trajectory to vary from the idealized response trajectory. Conversely, people with low uncertainty will more immediately select and answer and move to the next question. Thus, they would likely deviate less from the idealized response trajectory than people with high uncertainty. In summary,

H3: High uncertainty will result in greater AUC than low uncertainty.

Methodology

To test our hypotheses, we conducted two studies. In the first study, we manipulated uncertainty and explored whether the manipulation predicted AUC. In the second study, participants reported their level of uncertainty on questions, and we regressed this independent variable on AUC. Table 1 summarizes the two studies.

	Study 1	Study 2
Participants	102 undergraduate students	32 participants recruited via snowball sampling technique
Experiment Design	Single-factor repeated-measure design with 3 levels (high, moderate, and low uncertainty)	Observational study exploring the relationship between self-reported levels of uncertainty and AUC
Questions	Questions were designed to illicit high, moderate, and low uncertainty. Participants answered 3 questions in each category.	20 multiple-choice practice questions were chosen from online sources (e.g., SAT, ACT). Questions were chosen equally from the 4 lower levels of Bloom's Taxonomy to induce variance in uncertainty. Participants answered all 20 questions.
Findings	H1 and H2 were supported. H3 was not supported.	H1, H2, and H3 were supported.

Table 1. Overview of Experimental Studies

Study 1

To test our hypotheses, Study 1 deployed a single-factor repeated-measure experiment design with three levels: low uncertainty, moderate uncertainty, and high uncertainty. Participants answered three questions in each category. The repeated measure design allowed us to control for naturally occurring individual differences among participants. For example, participants have inherent differences in fine motor control and mouse sensitivity. In our analysis, we can control for these differences by treating the participant id as a random variable and AUC as a fixed variable in a mixed-effects model. Three of the questions are shown in Table 2.

Low Uncertainty	Moderate Uncertainty	High Uncertainty
Which country borders the United States to the North? • Mexico • Cuba • <u>Canada</u> • Russia	A bat and ball cost \$1.10 The bat costs one dollar more than the ball. How much does the ball cost? • <u>\$0.05</u> • \$0.20 • \$0.10 • \$0.15	What is the top speed that Usain Bolt has been clocked at? • 32mph • 15mph • <u>28mph</u> • 22mph

Table 2. Sample Survey Questions

After answering each question, participants answered the following question on a seven-point Likert scale (from strongly disagree to strongly agree): I was uncertain about my answer to this question. The question is used as a manipulation check in the analysis to explore whether we successfully manipulated uncertainty.

Participants

Participants were recruited from an undergraduate information systems course at a large private university. All participants were sent a link to an online survey to complete on their own computers at their convenience. For participating, students were offered extra credit equivalent to .25% of a percentage point. One-hundred-and-two people participated in the study. One subject did not complete the last question. Thus, we have 917 valid responses (102 participants x 9 questions – 1 no response). Approximately 73% of participants were male. The average age was 23.9.

Measuring Mouse Movements

A JavaScript code based on the jQuery library (http://jquery.com/) was imbedded in the survey for the experiment to capture mouse movements as participants moved their mouse from the starting location to an answer. A jQuery-based script was utilized as jQuery is free, widely available, and compatible with all major web browsers. Thus, it can easily be implemented in almost all types of webpages for research or pragmatic purposes. Using the jQuery ".mousemove()" function, the code captured the x-, y-coordinate pairs and timestamps at millisecond intervals[1] as the participants moved the mouse from the anchoring position to the answer. The script then sent this data to a web service, developed by the research team, for further analysis. The web service rescaled the x-, y-coordinate pairs to an 8 x 6 grid: the x-axis goes from -4 to 4, and the y-axis goes from -3 to 3. The mouse's starting position is mapped at coordinate (0, 0). This technique normalizes the data to help account for different screen resolutions, and facilitates the analysis of cursor movements even if the mouse starting location is different. From this normalized data, AUC was calculated by subtracting the area of the actual mouse movement trajectory from the area of the idealized response trajectory, and then taking the absolute value of that difference. The area of the idealized response trajectory is found by calculating the area of the right triangle that connects the beginning point (0,0) and the ending point (x,y) on the diagonal. The area of the actual trajectory was calculated through a Riemann Sums bootstrapping technique.

Analysis

We first explore whether our manipulations of uncertainty were successful. An ANOVA indicated that the three question groups did have a significant difference in reported uncertainty, $f(2,915) = 94.985$, $p < .001$. A post-hoc analysis with a Bonferroni adjustment indicated that people had significantly less uncertainty for the low uncertainty questions than the moderate uncertainty questions ($p < .001$), less uncertainty for the moderate uncertainty questions than the high uncertainty questions ($p<.001$), and less uncertainty for the low uncertainty question than the high uncertainty questions ($p<.001$).

To test the hypotheses, we utilized a within-subject ANOVA. This allows us to account for individual (participant-level) differences in mousing (mouse sensitivity, individual skills, etc.) while comparing treatment groups. First, we constructed a linear mixed-effects model using the nlme package in R using the following model,

$$auc_i = \alpha_i + \beta \cdot treatment_i + \varepsilon_i$$

whereas i is the index of the individual, α_i is the individual-specific effect, and treatment is the condition the participant was in (low, moderate, or high uncertainty). AUC was the dependent variable. Afterwards, an ANOVA was performed on the mixed model using the multcomp in R, with a Bonferroni post-hoc comparison. The ANOVA indicated that there was a difference between our treatment groups, $f(2, 914)= 6.855$, $p < .001$. The post-hoc comparison indicated that the moderate uncertainty questions had significantly higher AUC than the low uncertainty questions ($p < .01$) and the high uncertainty questions ($p<.01$). Thus, H1 and H2 were supported. However, the AUC on the low uncertainty questions was not significantly different from the higher uncertainty questions ($p > .05$). Thus, H3 was not supported. The model had an $R2$ of .127.

---

[1] Typically at a rate higher than 70 Hz (70 times per second).

In further exploring this non-significant result, we explored whether total response time was different among the treatment groups using the same mixed-effects model but with response time as the dependent variable. We found that both low uncertainty and high uncertainty had significantly lower response times than the moderate uncertainty group, $f(2,914)=8.013$, $p < .001$ and $f(2,914)=6.939$, $p < .001$ respectively. However, there was no difference between the low and high uncertainty groups, $f(2,914)=1.068$, $p > .05$. This lends support that if people have low or high uncertainty, they similarly respond by moving quickly and directly to an answer without much elaboration (either answering correctly or randomly guessing).

Study 2

Study 2 was an observational study with no discrete manipulations of uncertainty. Participants completed a diverse set of questions with varying degrees of difficulty, and reported their level of uncertainty immediately following each question. We then explored the relationship between the self-reported uncertainty and participant's AUC.

We selected 20 multiple-choice questions from various online resources for the test (e.g., the SAT, the California Achievement Test). To specify a diverse set of questions, we categorized each question into Blooms Taxonomy (Bloom et al. 1956), and chose 5 questions each that roughly fall in each of the first four categories: remember, understand, apply, and analyze. We were not able to find adequate multiple-choice questions in the last two categories of Bloom's taxonomy (evaluate or create), and thus we limit the scope of our paper to the lower four categories. This resulted in 20 questions total. The diverse set of questions promote an adequate level of uncertainty; the mean level of self-reported uncertainty was 5.453 (sd=1.600).

After answering each question, participants answered a series of questions regarding their response, including the following question to assess uncertainty on a seven-point Likert Scale (ranging from Definitely Incorrect to Definitely Correct): Rate how correct you think your answer is on the following scale. This response was used as an independent variable in our analysis.

Participants

The research team recruited participants using a snowball sampling technique. All participants were sent a link to the survey to be completed at their own convenience on their personal computers. Thirty two people participated in the study. Nine of the participants only submitted partial data (did not complete the entire study). Furthermore, limitations on participants' devices (e.g., a person using a cell phone to take the survey with no computer mouse) precluded other responses. As a result, we had 398 total valid responses to questions. The average age of participants was 32.4. Approximately 53% of participants were female.

Analysis

Mouse movements were analyzed using the same system as in Study 1. To test our hypotheses, we specified a linear mixed-effects model as follows:

$$auc_i = \alpha_i + \beta_1 \cdot levelOfUncertainty_i - \beta_2 \cdot levelOfUncertainty_i^{\,2} + \varepsilon_i$$

whereas i is the index of the individual, α_i is the individual-specific effect, and levelOfUncertainty is the level of uncertainty the person reported. Furthermore, as we expect a non-linear relationship between uncertainty and AUC (people with low and high uncertainty are likely to have lower AUC than people with moderate uncertainty), we included the negative square of uncertainty to help test the hypotheses. AUC was the dependent variable.

The results indicated that both the main effect ($\chi^2(1) = 3.96$, $p <.05$) and negative-squared effect ($\chi^2(1) = 2.92$, $p <.05$) of uncertainty significantly influenced AUC. These results lend support to our hypotheses. Namely, uncertainty is significantly correlated with AUC overall. However, the inverse u-shaped curve is significant—people with low and high uncertainty have lower AUC than people with moderate uncertainty. The model had an R2 of .103.

As a supplemental analysis, we explored whether the level of uncertainty similarly predicts response time (like in Study 1). The same analysis was conducted as previously, but with response time as the dependent variable rather than AUC. The results indicated that neither the main effect ($\chi^2(1) = 0.33$, $p > .05$) nor the negative-squared effect ($\chi^2(1) = 0.29$, $p > .05$) of uncertainty were significant in predicting response time. Thus, the level uncertainty (as opposed to a dichotomous variable in study 1) reflected in AUC but not response time in Study 2.

Discussion

The purpose of this research was to explore the following research question: how do mouse-cursor movement characteristics correlate with uncertainty as people answer multiple-choice questions online? To answer this question, we created hypotheses relating to how low, moderate, and high uncertainty will influence a mousing statistic called area-under-the-curve based on the Metamemory framework and theory on attention. We conducted two studies to test our hypotheses. Study 1 manipulated the three levels of uncertainty through discrete treatments. The results supported that moderate uncertainty led to greater AUC than both the low and high uncertainty conditions (H1 and H2 supported). However, we did not find a significant difference in AUC between the low and high uncertainty conditions (H3 not supported). Future research should investigate this non-supported relationship in greater detail. One possibility is that our high uncertainty questions led to a very low Feeling of Knowing and thus people never entered the search loop and immediately guessed on the question. Likewise, our low uncertainty questions may have been so easy that participants immediately knew the answer. Our supplemental analysis of time supports these explanations.

To help explore this in greater detail, we conducted a second study in which a diverse set of questions were used to manipulate uncertainty. After each question, we had participants report their level of uncertainty on their answer. Our results of this study lend support to our hypotheses. The main effect of reported uncertainty on AUC was significant (lending support to H3). Likewise, the negative-squared of uncertainty was also significant, suggesting that an inverse u-shaped curve also exists (lending support to H1 and H3). We next discuss the theoretical and practical implications of these results.

Implications for Research

First, our research provides a possible novel methodology for measuring uncertainty as people answer multiple-choice questions in research. Specifically, we found that AUC is significantly correlated with people's uncertainty in answering questions. This can be used to help measure uncertainty in multi-method research, used as a manipulation check in uncertainty studies, or used to measure the quality of people's responses to questions. Further, a supplemental analysis in experiment 2 indicated that the AUC of one's movement may provide information that is not readily available by simply measuring response time. Thus, from a grand perspective, our results echo the statement of Freeman et al. (2011 p. 1) that mouse cursor tracking "offer[s] continuous streams of output that can reveal ongoing dynamics of processing, potentially capturing the mind in motion with fine grained temporal sensitivity."

Unlike previous mouse cursor tracking literature, we find that AUC can be used to measure cognitive states in the wild. Traditional measures of AUC are often used to indicate deviation from the idealized response trajectory in highly controlled experiments—e.g., moving the mouse from the bottom of the screen to choose between two competing responses located in the upper corners of the screen. However, rarely has AUC been used in scenarios where people's mouse cursor position may start at different places on the screen, and have multiple destinations. In our experiment, we account for this less controlled environment by dynamically recording each participant's starting position when a question is first shown. We then calculate the ending point by recording the location of the last answer clicked-on by participants before submitting. Thus, we contribute to mouse cursor tracking in HCI by demonstrating that an idealized response trajectory can be dynamically calculated for each participant in the wild, even with different starting locations.

Finally, we integrate theory on attention and the Metamemory framework to help predict mouse-cursor movements. Under conditions of moderate uncertainty, the Metamemory framework suggests that people will engage in an iterative search process. Our results support this proposition—people under moderate uncertainty deviate more from the most direct path in answering than under high and low uncertainty, thus lending support of such a process. Thus, our results suggest that the Metamemory framework may be a useful theory for understanding mouse movement responses to multiple-choice questions in human-computer interactions.

Implications for Practice

Our results have several implications for practice. As these are our first attempts to measure uncertainty through mouse-cursor movements, we chose a simple scenario to begin our exploration—uncertainty as people answer multiple choice quiz questions. This, in itself, has implications for information learning assessments in organizational or educational settings. For example, measuring uncertainty in an educational setting may provide much useful information. High AUC as an indicator of moderate uncertainty may indicate that a question is cognitively challenging, that a person lacks mastery of a subject topic, or even that a question is poorly worded or hard to understand. Conversely, low AUC may indicate two extremes: mastery of a topic or no knowledge of a topic (and thereby guesses). Although AUC in itself is not definitive, it can be coupled with other information (whether a person got an answer correct or incorrect) to provide insight into an interaction, or it can trigger the need to ask follow-up questions about a topic or about a particular question.

Another potential application of our research is to measure uncertainty in an e-commerce setting. AUC metrics can be collected unobtrusively within an ecommerce site, which in turn, can reveal buying uncertainty. An accurate measure of buying uncertainty can then be used to improve website design or increase a site's conversion rate. For example, organizations can use this additional piece of data (buying uncertainty) to trigger interventions to secure sales (e.g., offer assistance, discounts, provide more information, etc.)

Further, our results set a foundation for much broader implications in human-computer interactions. Providing information through forms and surveys online is pervasive. As previously mentioned, job applications, product reviews, online questionnaires, credit applications, and compliance surveys are only a few examples of the diverse set of online systems that gather important information. Currently, however, organizations often only receive a person's final response, and have no information about the cognitive process that people went through to conclude the answer. Our approach suggests that contextual information about an interaction (like the level of uncertainty while answering a question) may be gleaned from people's mouse-cursor movements. This enhances an organization's ability to trigger follow-up questions, weigh responses, provide clarifications, or otherwise gain insight to improve data quality.

Limitations and Future Research

Our research is an early but fundamental examination of the influence of uncertainty on mouse-cursor movements. As such, the research has many limitations that need to be addressed in future research. First, our research only examined the influence of uncertainty on AUC. However, many other mousing statistics exist, such as movement speed, acceleration, stance, idle time, maximum deviation, and direction changes (Hehman et al. 2014). Examining these other statistics may also provide information about uncertainty, and even help differentiate between high and low levels of uncertainty.

Both studies have limitations. In Study 2, we had several participants quit the quiz early, possibly because of its length. Furthermore, technical limitations on several of the personal computers inhibited data collection. This data was not available for analysis. Further, study 2 correlates uncertainty with people's self-report measure of uncertainty. Often people overestimate their certainty and are hesitant to admit their limitations (e.g., their uncertainty). Thus, our results may be subject to a response bias.

Behaviors in both studies may be influenced by the types of questions asked. In this study, we manipulated questions based on Blooms Taxonomy. Some questions were there inherently longer than others, required computational calculations, or had other differences. This may account for why response time was not an indicator of uncertainty—i.e., different questions required different amounts of time to read and complete. AUC on the other hand is not dependent on the length or complexity of the question. Thus, it is likely less biased by the question type than response time.

Furthermore, future research should explore how different characteristics of the test-taking environment will influence mousing behavior. For example, rewards systems may influence mouse movements (e.g., whether a person gets rewarded for answering a question correctly, or punished for answering incorrectly). As another example, the timing of tests may influence movements (e.g., whether the test is timed or not). Future research should explore these moderating factors.

References

Arroyo, E., Selker, T., and Wei, W. 2006. "Usability Tool for Analysis of Web Designs Using Mouse Tracks," *CHI '06 Extended Abstracts on Human Factors in Computing Systems*, Montreal, Canada: ACM, pp. 484-489.

Atterer, R., and Lorenzi, P. 2008. "A Heatmap-Based Visualization for Navigation within Large Web Pages," *Proceedings of the 5th Nordic conference on Human-Computer Interaction*, Lund, Sweden: ACM, pp. 407-410.

Barca, L., and Pezzulo, G. 2012. "Unfolding Visual Lexical Decision in Time," *PloS One* (7:4), p. e35932.

Bartolotti, J., and Marian, V. 2012. "Language Learning and Control in Monolinguals and Bilinguals," *Cognitive Science* (36:6), pp. 1129-1147.

Bloom, B.S., Engelhart, M.D., Furst, E.J., Hill, W.H., and Krathwohl, D.R. 1956. "Taxonomy of Educational Objectives: The Classification of Educational Goals," in *Handbook I: Cognitive domain*. New York: David McKay Company.

Bruhn, P. 2013. "Emergence of Spontaneous Anticipatory Hand Movements in a Probabilistic Environment," *Advances in Cognitive Psychology* (9:2), pp. 62-73.

Byrne, M.D., Anderson, J.R., Douglass, S., and Matessa, M. 1999. "Eye Tracking the Visual Search of Click-Down Menus," in: *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. Pittsburgh, Pennsylvania, USA: ACM, pp. 402-409.

Chen, M.C., Anderson, J.R., and Sohn, M.H. 2001. "What Can a Mouse Cursor Tell Us More?: Correlation of Eye/Mouse Movements on Web Browsing," in: *CHI '01 Extended Abstracts on Human Factors in Computing Systems*. Seattle, Washington: ACM, pp. 281-282.

Cloutier, J., Freeman, J.B., and Ambady, N. 2014. "Investigating the Early Stages of Person Perception: The Asymmetry of Social Categorization by Sex vs. Age," *PloS One* (9:1), p. e84677.

Cox, A.L., and Silva, M.M. 2006. "The Role of Mouse Movements in Interactive Search," *28th Annual Meeting of the Cognitive Science Society*, Vancouver, British Columbia, pp. 1156-1161.

Dale, R., Kehoe, C., and Spivey, M.J. 2007. "Graded Motor Responses in the Time Course of Categorizing Atypical Exemplars," *Memory & Cognition* (35:1), pp. 15-28.

Dale, R., Roche, J., Snyder, K., and McCall, R. 2008. "Exploring Action Dynamics as an Index of Paired-Associate Learning," *PLoS One* (3:3), p. e1728.

Dennis, A.R., Fuller, R.M., and Valacich, J.S. 2008. "Media, Tasks, and Communication Processes: A Theory of Media Synchronicity," *MIS quarterly* (32:3), pp. 575-600.

Derrick, D.C., Meservy, T.O., Jenkins, J.L., Burgoon, J.K., and Nunamaker Jr, J.F. 2013. "Detecting Deceptive Chat-Based Communication Using Typing Behavior and Message Cues," *ACM Transactions on Management Information Systems (TMIS)* (4:2), pp. 62-81.

Dshemuchadse, M., Scherbaum, S., and Goschke, T. 2013. "How Decisions Emerge: Action Dynamics in Intertemporal Decision Making," *Journal of Experimental Psychology* (142:1), pp. 93-100.

Duran, N.D., Dale, R., and McNamara, D.S. 2010. "The Action Dynamics of Overcoming the Truth," *Psychonomic Bulletin & Review* (17:4), pp. 486-491.

Farmer, T.A., Cargill, S.A., Hindy, N.C., Dale, R., and Spivey, M.J. 2007. "Tracking the Continuity of Language Comprehension: Computer Mouse Trajectories Suggest Parallel Syntactic Processing," *Cognitive Science* (31:5), pp. 889-909.

Faulkenberry, T.J. 2013. "How the Hand Mirrors the Mind: The Embodiment of Numerical Cognition," *Proceedings of the 40th Annual Meeting of the Research Council on Mathematics Learning*, Tulsa, Oklahoma, p. 205.

Flavell, J.H., and Wellman, H.M. 1975. "Metamemory," *Annual Meeting of the American Psychological Association*, Chicago, p. 66.

Freeman, J.B. 2014. "Abrupt Category Shifts During Real-Time Person Perception," *Psychonomic Bulletin & Review* (21:1), pp. 85-92.

Freeman, J.B., and Ambady, N. 2009. "Motions of the Hand Expose the Partial and Parallel Activation of Stereotypes," *Psychological Science* (20:10), pp. 1183-1188.

Freeman, J.B., and Ambady, N. 2010. "MouseTracker: Software for Studying Real-Time Mental Processing Using a Computer Mouse-Tracking Method," *Behavior Research Methods* (42:1), pp. 226-241.

Freeman, J.B., and Ambady, N. 2011. "When Two Become One: Temporally Dynamic Integration of the Face and Voice," *Journal of Experimental Social Psychology* (47:1), pp. 259-263.

Freeman, J.B., Ambady, N., Rule, N.O., and Johnson, K.L. 2008. "Will a Category Cue Attract You? Motor Output Reveals Dynamic Competition Across Person Construal," *Journal of Experimental Psychology* (137:4), pp. 673-690.

Freeman, J.B., and Dale, R. 2012. "Assessing Bimodality to Detect the Presence of a Dual Cognitive Process," *Behavior Research Methods* (45:1), pp. 1-15.

Freeman, J.B., Dale, R., and Farmer, T.A. 2011. "Hand in Motion Reveals Mind in Motion," *Frontiers in Psychology* (2:59), pp. 1-6.

Georgopoulos, A.P. 1990. "Neurophysiology of Reaching," in *Attention and Performance XIII,* M. Jeannerod (ed.). Hillsdale, NJ: Lawrence Erlbaum Associates Inc., pp. 227-263.

Grimes, G.M., Jenkins, J.L., and Valacich, J.S. 2013. "Assessing Credibility by Monitoring Changes in Typing Behavior: The Keystroke Dynamics Deception Detection Model," *HICSS-46 Symposium on Credibility Assessment and Information Quality in Government and Business*, Maui, Hawaii.

Guo, Q., and Agichtein, E. 2010. "Towards Predicting Web Searcher Gaze Position from Mouse Movements," *CHI'10 Extended Abstracts on Human Factors in Computing Systems*, Austin, TX: ACM, pp. 3601-3606.

Hehman, E., Stolier, R.M., and Freeman, J.B. 2014. "Advanced Mouse-Tracking Analytic Techniques for Enhancing Psychological Science," *Psychological Science* (20:10), pp. 1183–1188.

Huang, J., White, R.W., and Dumais, S. 2011. "No Clicks, No Problem: Using Cursor Movements to Understand and Improve Search," *SIGCHI Conference on Human Factors in Computing Systems*, Vancouver, BC, Canada: ACM, pp. 1225-1234.

Kahneman, D. 2011. *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.

Kahneman, D., and Tversky, A. 1979. "Prospect theory: An Analysis of Decision Under Risk," *Econometrica: Journal of the Econometric Society* (47:2), pp. 263-291.

Lagun, D., and Agichtein, E. 2011. "ViewSer: Enabling Large-Scale Remote User Studies of Web Search Examination and Interaction," *SIGIR*, Beijing, China, pp. 365-374.

Lettner, F., and Holzmann, C. 2012. "Heat Maps as a Usability Tool for Multi-Touch Interaction in Mobile Applications," *Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia*, Ulm, Germany: ACM, pp. 1-2.

Maehr, W. 2008. *eMotion: Estimation of User's Emotional State by Mouse Motions*. Saarbrücken, Germany: VDM Verlag.

Martens, M.A., Hasinski, A.E., Andridge, R.R., and Cunningham, W.A. 2012. "Continuous Cognitive Dynamics of the Evaluation of Trustworthiness in Williams Sndrome," *Frontiers in Psychology* (3:160).

McKinstry, C., Dale, R., and Spivey, M.J. 2008. "Action Dynamics Reveal Parallel Competition in Decision Making," *Psychological Science* (19:1), pp. 22-24.

Metcalfe, J., Eich, T.S., and Miele, D.B. 2013. "Metacognition of Agency: Proximal Action and Distal Outcome," *Experimental Brain Research* (224:1), pp. 1-12.

Mueller, F., and Lockerd, A. 2001. "Cheese: Tracking Mouse Movement Activity on Websites, A Tool for User Modeling," in: *CHI '01 Extended Abstracts on Human Factors in Computing Systems*. Seattle, Washington: ACM, pp. 279-280.

Navalpakkam, V., and Churchill, E. 2012. "Mouse Tracking: Measuring and Predicting Users' Experience of Web-Based Content," in: *2012 ACM annual conference on Human Factors in Computing Systems*. Austin, Texas, USA: ACM, pp. 2963-2972.

Nelson, T.O. 1990. "Metamemory: A Theoretical Framework and New Findings," *Psychology of Learning and Motivation* (26:1), pp. 125-173.

Palmer, C.J., Paton, B., Barclay, L., and Hohwy, J. 2013. "Equality, Efficiency, and Sufficiency: Responding to Multiple Parameters of Distributive Justice During Charitable Distribution," *Review of Philosophy and Psychology* (4:3), pp. 1-16.

Pan, B., Hembrooke, H.A., Gay, G.K., Granka, L.A., Feusner, M.K., and Newman, J.K. 2004. "The Determinants of Web Page Viewing Behavior: An Eye-Tracking Study," in: *2004 Symposium on Eye Tracking Research & Applications*. San Antonio, Texas: ACM, pp. 147-154.

Papesh, M.H., and Goldinger, S.D. 2012. "Memory in Motion: Movement Dynamics Reveal Memory Strength," *Psychonomic Bulletin & Review* (19:5), pp. 906-913.
Rodden, K., Fu, X., Aula, A., and Spiro, I. 2008. "Eye-Mouse Coordination Patterns on Web Search Results Pages," in: *CHI '08 Extended Abstracts on Human Factors in Computing Systems*. Florence, Italy: ACM, pp. 2997-3002.
Rodrigues, M., Gonçalves, S., Carneiro, D., Novais, P., and Fdez-Riverola, F. 2013. "Keystrokes and Clicks: Measuring Stress on E-learning Students," in *Management Intelligent Systems,* J. Casillas, F.J. Martinez-Lopez, R. Vicari and F.D.l. Prieta (eds.). Switzerland: Springer, pp. 119-126.
Shah, K.N., Rathod, K.R., and Agravat, S.J. 2014. "A Survey on Human Computer Interaction Mechanism Using Finger Tracking," *arXiv preprint arXiv:1402.0693*).
Solman, G.J.F., Cheyne, J.A., and Smilek, D. 2012. "Found and Missed: Failing to Recognize a Search Target Despite Moving It," *Cognition* (123:1), pp. 100-118.
Song, J.H., and Nakayama, K. 2006. "Role of Focal Attention on Latencies and Trajectories of Visually Guided Manual Pointing," *Journal of Vision* (6:9), pp. 982-995.
Song, J.H., and Nakayama, K. 2008. "Target Selection in Visual Search as Revealed by Movement Trajectories," *Vision Research* (48:7), pp. 853-861.
Spivey, M.J., Grosjean, M., and Knoblich, G. 2005. "Continuous Attraction Toward Phonological Competitors," *National Academy of Sciences of the United States of America* (102:29), pp. 10393-10398.
Tipper, S.P., Howard, L.A., and Jackson, S.R. 1997. "Selective Reaching to Grasp: Evidence for Distractor Interference Effects," *Visual Cognition* (4:1), pp. 1-38.
Tower-Richardi, S.M., Brunyé, T.T., Gagnon, S.A., Mahoney, C.R., and Taylor, H.A. 2012. "Abstract Spatial Concept Priming Dynamically Influences Real-World Actions," *Frontiers in Psychology* (27).
Tversky, A., and Kahneman, D. 1981. "The Framing of Decisions and the Psychology of Choice," *Science* (211:4481), pp. 453-458.
Wang, Q., Taylor, H.A., and Brunyé, T.T. 2012. "When Going the Right Way is Hard to Do: Distinct Phases of Action Compatibility in Spatial Knowledge Development," *Acta Psychologica* (139:3), pp. 449-457.
Weaver, S.M., and Arrington, C.M. 2013. "Tracking the Multitasking Mind," *Zeitschrift für Psychologie* (221:1), pp. 51-60.
Weis, L. 2012. "The Honest Hand - How Computer Mouse Trajectories Might Capture Deceptive Processes," University of Aberdeen, p. 17.
Welsh, T.N., and Elliott, D. 2004. "Movement Trajectories in the Presence of a Distracting Stimulus: Evidence for a Response Activation Model of Selective Reaching," *The Quarterly Journal of Experimental Psychology* (57:6), pp. 1031-1057.
Wojnowicz, M.T., Ferguson, M.J., Dale, R., and Spivey, M.J. 2009. "The Self-Organization of Explicit Attitudes," *Psychological Science* (20:11), pp. 1428-1435.
Yu, Z., Wang, F., Wang, D., and Bastin, M. 2012. "Beyond Reaction Times: Incorporating Mouse-Tracking Measures into the Implicit Association Test to Examine its Underlying Process," *Social Cognition* (30:3), pp. 289-306.
Zimmermann, P., Gomez, P., Danuser, B., and Schär, S. 2006. "Extending Usability: Putting Affect into the User-Experience," *Proceedings of NordiCHI'06*, pp. 27-32.
Zimmermann, P., Guttormsen, S., Danuser, B., and Gomez, P. 2003. "Affective Computing-A Rationale for Measuring Mood with Mouse and Keyboard," *International Journal of Occupational Safety and Ergonomics* (9:4), pp. 539-551.
Zushi, M., Miyazaki, Y., and Norizuki, K. 2012. "Web Application for Recording Learners' Mouse Trajectories and Retrieving their Study Logs for Data Analysis," *Knowledge Management & E-Learning: An International Journal* (4:1), pp. 37-50.

Appendix A: Summary of Relevant Studies Measuring Cognitive / Emotional Responses through Mouse Cursor Movements

Cognitive process examined through mouse / hand movements	Citation
Attitude formation, concealment of racial prejudices	(Wojnowicz et al. 2009)
Attraction toward distracting stimuli	(Song and Nakayama 2006; Song and Nakayama 2008)
Decision conflict	(McKinstry et al. 2008; Palmer et al. 2013)
Decision making	(Dshemuchadse et al. 2013; Martens et al. 2012)
Deception	(Duran et al. 2010; Weis 2012)
Detection of dual cognitive processing	(Freeman and Dale 2012)
Dynamic competition in classification tasks	(Dale et al. 2007; Freeman and Ambady 2009; Freeman and Ambady 2011; Freeman et al. 2008)
Emotional reactions	(Maehr 2008; Rodrigues et al. 2013; Zimmermann et al. 2006; Zimmermann et al. 2003)
Increased cognitive processing	(Freeman and Ambady 2011)
Language learning, processing, or interpretation	(Barca and Pezzulo 2012; Bartolotti and Marian 2012; Farmer et al. 2007; Spivey et al. 2005)
Learning	(Dale et al. 2008; Zushi et al. 2012)
Mathematical processing	(Faulkenberry 2013)
Memory recall	(Papesh and Goldinger 2012)
Metacognition	(Metcalfe et al. 2013)
Perception formation of people	(Cloutier et al. 2014; Freeman 2014)
Semantic Priming	(Shah et al. 2014)
Search / Recognition	(Solman et al. 2012)
Spatial knowledge development	(Wang et al. 2012)
Subconscious / Implicit / Anticipatory processing	(Bruhn 2013; Tower-Richardi et al. 2012; Yu et al. 2012)
Task switching	(Weaver and Arrington 2013)

Table A1: Summary of Relevant Studies Measuring Cognitive / Emotional Responses