Elsevier Editorial System(tm) for Applied Ergonomics

Manuscript Draft

Manuscript Number: JERG-D-16-00457R3

Title: Effects of Monetary Reward and Punishment on Information Checking Behaviour: An Eye-Tracking Study

Article Type: Full Length Article

Keywords: Error; reward; punishment; data-entry

Corresponding Author: Dr. Simon Y. W. Li, PhD

Corresponding Author's Institution: Lingnan University

First Author: Simon Y. W. Li, PhD

Order of Authors: Simon Y. W. Li, PhD; Anna L Cox; Calvin Or; Ann Blandford

Abstract: The aim of the present study was to investigate the effect of error consequence, as reward or punishment, on individuals' checking behaviour following data entry. This study comprised two eye-tracking experiments that replicate and extend the investigation of Li, Cox, Or, and Blandford (2016) into the effect of monetary reward and punishment on data-entry performance. The first experiment adopted the same experimental setup as Li et al. (2016) but additionally used an eye tracker. The experiment validated Li et al. (2016) finding that, when compared to no error consequence, both reward and punishment led to improved data-entry performance in terms of reducing errors, and that no performance difference was found between reward and punishment. The second experiment extended the earlier study by associating error consequence to each individual trial by providing immediate performance feedback to participants. It was found that gradual increment (i.e. reward feedback) also led to significantly more accurate performance than no error consequence. It is unclear whether gradual increment is more effective than gradual decrement because of the small sample size tested. However, this study reasserts the effectiveness of reward on data-entry performance.

Highlights

- Direct comparison of the effect of reward and punishment on data-entry performance.
- Monetary reward and punishment improve accuracy compared to no incentives.
- Incremental reward is more effective than punishment.
- Accurate performance does not necessarily depend on checking frequency or duration.

Title Page

Effects of Monetary Reward and Punishment on Information Checking Behaviour: An Eye-Tracking Study

Simon Y. W. Li [a], Anna L. Cox [b], Calvin Or [c] & Ann Blandford [b]

[a] Department of Applied Psychology, Lingnan University, WYL Building, 8 Castel Peak Road, Tuen Mun, Hong Kong SAR

[b] UCLIC, University College London, 2nd floor, 66 – 72 Gower Street, London, WC1E 6EA, UK (anna.cox@ucl.ac.uk; a.blandford@ucl.ac.uk)

[c] Department of Industrial & Manufacturing Systems Engineering, University of Hong Kong, Haking Wong Building, Pokfulam Rd., Hong Kong SAR (klor@hku.hk)

Corresponding author: Simon Y. W. Li

Email: simonli2@ln.edu.hk

Postal address: WYL-306, WYL Building, Lingnan University, 8 Castle Peak Road, Tuen Mun, Hong Kong SAR

Abstract

The aim of the present study was to investigate the effect of error consequence, as reward or punishment, on individuals' checking behaviour following data entry. This study comprised two eye-tracking experiments that replicate and extend the investigation of Li, Cox, Or, and Blandford (2016) into the effect of monetary reward and punishment on data-entry performance. The first experiment adopted the same experimental setup as Li et al. (2016) but additionally used an eye tracker. The experiment validated Li et al. (2016) finding that, when compared to no error consequence, both reward and punishment led to improved data-entry performance in terms of reducing errors, and that no performance difference was found between reward and punishment. The second experiment extended the earlier study by associating error consequence to each individual trial by providing immediate performance feedback to participants. It was found that gradual increment (i.e. reward feedback) also led to significantly more accurate performance than no error consequence. It is unclear whether gradual increment is more effective than gradual decrement because of the small sample size tested. However, this study reasserts the effectiveness of reward on data-entry performance.

(Word count for Abstract: 186)

Keywords

Error; reward; punishment; data-entry

1. Introduction

The ability to detect and correct one's errors is an important aspect of human performance. This ability becomes especially important in work contexts that are mission or safety critical. For example, in healthcare, a widely cited report estimated that about 44,000 to 98,000 hospital deaths each year are results of various kinds of medical errors (Kohn, Corrigan, & Donaldson, 2000) . Moreover, safety-critical errors can often happen in routine tasks such as data-entry, which has been ranked as the fourth-leading cause of medication errors by the U.S. Pharmacopeia in 2003 ("Data entry is a top cause of medication errors," 2005).

Traditional research on error detection has focused on whether or not errors are detected, and whether certain error types (e.g. slips vs mistakes) are easier to be detected than others (Sellen, 1994; Zapf, Maier, Rappensperger, & Irmer, 1994). A number of theoretical models of error detection have been proposed. For example, Reason (1990) described three main ways in which errors get detected: (1) by monitoring one's own performance; (2) by cues or feedback provided in the environment; and (3) by other people. Sellen (1994) proposed a similar framework and suggested that error detection can occur via (1) the incorrect actions themselves; (2) consequences from the incorrect actions; (3) external constraints in the environment; and (4) other people. More recently, Blavier, Rouy, Nyssen, and De Keyser (2005) proposed a model of error detection based on prospective memory and emphasised the importance of intention formation and retention when detecting errors. Despite their different theoretical orientation, these error detection models share a common idea that regular checking of one's own performance forms an important part of the detection process. This is supported by empirical evidence from laboratory (e.g. Allwood, 1984) and observational studies (e.g. Nyssen & Blavier, 2006).

Human-computer interaction (HCI) researchers have also investigated data-entry performance. A number of studies have examined the effectiveness of various data-entry methods by comparing double entry (same set of data entered twice), read aloud (data checked while another person reads them out loud), and visual checking (data checked by sight) (Barchard & Pace, 2011; Barchard & Verenikina, 2013); consistently, double entry has been found to result in the most accurate performance. However, whilst Barchard and colleagues' studies provide answers to the research questions they set, the findings do not address what motivates checking in the first place.

As previous research has implicated the essential role of checking in error detection, in this paper, we carried out two eye-tracking experiments to investigate how checking might be motivated to enhance error detection. The questions we ask are: can one motivate checking behaviour in terms of monetary reward and punishment? If so, can checking enhance data-entry performance?

The role of motivation has long been recognised in theoretical discussions of human error (e.g. Lourens, 1990) and it may take many different forms. For example, accountability can be a form of motivation, which determines whether or not one is held accountable for one's errors; and this can have implications for an organisation's safety culture and its workers' attitudes towards errors (Dekker, 2009; Woods, Dekker, Cook, Johannesen, & Sarter, 2010). Error consequences in terms of reward and punishment are another form of motivation. It was found that punishment that did not have any direct consequence was ineffective in reducing errors (Back, Cheng, Dann, Curzon, & Blandford, 2007). However, when it had an actual cost on participants' performances, punishment resulted in fewer errors in procedural task performance (Brumby, Cox, Back, & Gould, 2013).

Li et al. (2016) carried out a study on data-entry, focusing on how to motivate checking behaviour by imposing error consequences in terms of monetary reward and punishment. They designed a computer-based data-entry task in which there were two panels, each with eight information fields, on the computer screen. Participants were required to transcribe textual or numerical information from the left panel to the right panel. Each information field on the left panel was covered by a grey box, and when the participants wanted to look at the to-be-transcribed information, they had to hover the mouse over the grey box to reveal the information. This paradigm allowed Li et al. to measure the number and duration of uncovering actions made by the participants; these measurements were used to quantify their checking behaviour. Error consequences were manipulated such that in the Reward condition, participants were informed that if they correctly transcribed all the trials, they would receive extra payment; in the Punishment condition, if the participants made even one error in any of the trials, they would receive a reduced payment; participants in the Control condition were paid a fixed amount for completing all the transcription trials. One of the main findings was that reward and punishment resulted in more accurate performance than no consequence; however, reward and punishment did not lead to different performance levels (see Table 1). The other finding of the study was that monetary reward and punishment motivated participants to engage in more frequent and longer checking behaviour than no consequence at all.

< Insert Table 1 here >

In order to further investigate the effects of reward and punishment on checking behaviour in terms of eye-movement data, we employed an eye tracker and an experimental task similar to that used in Li et al. (2016) . The experimental task was modified so that there were no grey boxes on the left panel of the computer screen; this way, participants could check the to-be-transcribed information by eye gazes only rather than mouse movement as in

Li et al.. By adopting an eye tracker, our first experiment is designed to examine the effect of reward and punishment under more ecologically valid checking conditions. Research on interaction behaviour suggests that people are sensitive to information access cost. Gray and Fu (2004) found that, when given a choice, people consistently opt for an interaction strategy that involves the least cognitive and physical effort even when the strategy might lead to suboptimal task performance. As our current experimental paradigm imposed less checking effort (visual) than Li et al. (manual mouse movement), we expect there would be an overall increase of checking behaviour in our first experiment relative to Li et al..

In terms of theoretical formulation in our study, it is worth noting that we did not adopt a framework such as the Tversky and Kahneman (1991) work on gain/loss framing because that is applied to decision-making tasks and our domain of interest is not decision making but data entry, which is a procedural activity. While the effect of loss aversion can be found in decision making under uncertainties, we do not see it as a suitable theoretical framework for our data-entry task. Therefore, we appeal to empirical findings from other domains in the following.

Findings from neuroscience suggest that a neural signal called error-related negativity (ERN) is sensitive to monetary gains (Stürmer, Nigbur, Schacht, & Sommer, 2011) and losses (Potts, 2011) and that different neural circuits are responsible for reward and punishment (Wrase et al., 2007; Yeung & Sanfey, 2004). Moreover, recent evidence suggests that reward is more effective than punishment in improving motor memory (Abe et al., 2011). Reward has also been found to improve creativity (Eisenberger & Cameron, 1996; Eisenberger & Rhoades, 2001) and motivation (Eisenberger, Rhoades, & Cameron, 1999; Hendijani, Bischak, Arvai, & Dugar, 2016). Although none of the outlined studies directly compared rewards to punishments (e.g. Eisenberger et al., 1999; Hendijani et al., 2016) or investigated the effect on data-entry performance (Abe et al., 2011), these studies suggest that reward is better than punishment at improving human performance in a number of domains. One of the novelties of our study is the investigation of motivation (in terms of reward and punishment) on data-entry performance, and, to the best of our knowledge, there are no existing studies in HCI that can provide us with direct predictions. Therefore, it is necessary to look for theoretical support from studies in other domains and test whether their findings can be generalised to ours. We examined the effect of reward and punishment in the first experiment and predicted that reward consequence would result in better data-entry performance than punishment consequence.

In the second experiment, we manipulated error consequences so that they were associated with each individual trial and that immediate performance feedback, in terms of payment increment (i.e. reward) or decrement (i.e. punishment), was also provided to the participants. Therefore, in the second experiment, we predicted that reward, in the form of gradual increment, would result in more accurate data-entry performance and more rigorous checking behaviour than punishment in the form of gradual decrement.

2. Experiment 1

This experiment was designed to examine the effect of reward and punishment on checking behaviour by using an eye tracker. Checking behaviour was quantified in terms of eye fixations and fixation duration. In Li et al. (2016), the experimental task involved the, previously mentioned, grey-box paradigm and checking was carried out by the participants through mouse movement. Therefore, checking, as performed in the current task paradigm, was less effortful when compared to Li et al. (2016).

Three hypotheses were tested in the current experiment: first, based on the effect of information access cost (Gray & Fu, 2004), we predicted an overall increase in the number of checks and a decrease in check duration when compared to Li et al. (2016). Because when checking is easier (as in the current experiment), participants might check more often but each check is shorter in duration due to the ease of information access. Second, the current experiment was expected to partially reproduce Li et al.'s results, namely that reward and punishment result in more accurate data-entry performance and more rigorous checking than no error consequences at all. Third, based on various empirical studies in the literature (Abe et al., 2011; Eisenberger & Cameron, 1996; Eisenberger & Rhoades, 2001; Eisenberger et al., 1999; Hendijani et al., 2016), reward is expected to lead to more accurate performance and more rigorous checking than punishment.

2.1 Method.

2.1.1 Participants.

Sixty participants (age range: 17 – 23 years; mean age: 19.9 years) from Lingnan University were recruited for the experiment. Data from 2 participants were discarded because of technical faults in recording their eye movement data; a further 2 participants were recruited to replace them. There were a total of 44 females and 16 males.

2.1.2 Design.

The experiment was a between-subject design with one factor, which is error consequence. This factor had three levels: Control, Reward and Punishment. As in Li et al. (2016), the Control condition had no error consequence, i.e. participants were informed that they would get HK$100 (~US$13) for completing the task and there was no mention of any consequences associated with errors committed. In the Reward condition, participants were informed about the reward consequence: they would get paid HK$ 60 (~US$8) for completing the experiment but would get an extra HK$40 (~US$5) if no errors were made. In the Punishment condition, participants received the punishing consequence instruction: they would get paid HK$100 (~US$13) for completing the study but would lose HK$40 (~US$5) if they made even one error. When recruiting participants for the experiment, the recruitment poster mentioned that the incentive payment would be HK$60 – 100 (~US$8 – $13) depending on task performance; there was no mention of reward or punishment in the recruitment advertisement.

The dependent variables were error rate, number of checks and duration of checks. An error was defined as a mismatch between the transcribed and the target information of each of the eight information fields. There were 40 trials in total for each participant, giving a total of 320 opportunities for making errors for each participant. An error rate was calculated for each participant by dividing the number of errors committed by the number of error opportunities.

Number of checks was defined as the number of eye fixations on a target information field. An eye fixation was counted as a check when the following two criteria were satisfied: 1) the corresponding text box of the information field (in the right panel of the screen) was not empty; and 2) each subsequent fixation on the target information field (in the left panel) was at least as long as a pre-specified duration, which was set to 100 ms and 500 ms. The 100-ms criterion is based on eye tracking research in reading, which suggest that fixation durations can range from 100 ms to over 500 ms (Rayner, 1978; Rayner & Duffy, 1986). For completeness, the 500-ms criterion was set in order to make a direct comparison with Li et al. (2016), which defined a check (by uncovering a greyed out text box) as an uncovered target for at least 500 ms. We used a shorter fixation (100 ms) for the current analysis because participants' checks could be shorter when they checked by eye gaze rather than by uncovering grey boxes. Duration of checks was measured in ms.

2.1.3 Materials.

The experimental task was a computer-based task programmed in Visual Basic.NET. The Tobii T120 Eye Tracker was used; it has a refresh rate of 120 Hz and is integrated in a 17” TFT monitor display with a resolution of 1280 × 1024 pixel. The experimental task was programmed to accommodate the eye tracker’s screen size and resolution. Each participant sat across from the computer screen at a distance of about 19 inches.

The eye tracking data were captured by the eye tracker’s software – Tobii Studio (version 3.2). Eye fixation data were extracted using Tobii Studio’s in-built ClearView Fixation Filter with *velocity threshold* set at 50 pixels/sample and *duration threshold* set at 500 ms and 100 ms because two sets of analyses were carried out (see the Results section).

2.1.4 Task.

The experimental task was a routine data-entry task, in which participants were required to transcribe textual and numeric information from one part of the screen to another. The task was framed as a Library task and participants were asked to transcribe information from eight different fields: *title*, *catalog ID*, *year*, *volume*, *inclusive pages*, *subject author* and *max*. *waiting time*. Figure 1 shows a screen shot of the task: the left panel contains the target information to be transcribed and the right panel contains text boxes for participants to type in the corresponding information. All actions performed by each participant were captured by the experimental task software in an action log. A data analysis script, written in Python, combined the eye tracking data and the action log data for overall data analysis.

< Insert Figure 1 here >

2.1.5 Procedure.

The experiment was carried out at the Psychology Laboratory at Lingnan University. An experimenter briefed a participant about the purpose of the experiment which described using a simple data-entry task to study how people use computers to transcribe simple textual and numeric information, and that an eye tracker was used to collect their eye tracking data. The experimenter did not mention studying the effect of reward and punishment at this point.

The experimenter then demonstrated a trial to the participant. In the demonstration, the experimenter emphasised that the transcribed information has to be exactly the same as the target information to be counted as correct. The experimenter also showed the participant, on printed laminated slides, examples of incorrect transcriptions due to errors in upper/lower cases, spaces, and punctuation.

The participant was given two practice trials in which he/she was to carry out the task without the experimenter's help. At the end of the practice trials, the participant was given feedback about his/her performance. Any errors made were pointed out to the participant and the correct transcription was explained. The experimenter showed the error examples on the laminated slides to the participant again to remind him/her about the potential sources of error.

The experimenter randomly assigned the participant to one the three conditions and gave the corresponding instruction to the participant. The participant would receive one of the following verbal instructions according to his/her assigned condition:

Control condition: "*When you finish transcribing all the trials, we will pay you $100 for your time*."

Reward condition: "*When you finish transcribing all the trials, we will pay you $60 for your time*. *However, if you make no mistakes in all of the trials, we will pay you an extra $40*. *And you will end up with $100*."

Punishment condition: "*When you finish transcribing all the trials, we will pay you $100 for your time*. *However, if you make just one mistake in any of the trials,* *we will take $40 off your payment. And you will end up with $60 only.*"

The experimenter also informed the participant that the eye tracker, which was integrated in the display monitor, would be switched on for the testing phase of the experiment and his/her eye movement and fixations would be recorded. The participant went through a calibration exercise for the eye tracker. The experimenter assisted the participant using Tobii Eye Tracker's calibration procedure (the nine-dot procedure). Once the calibration was successful, the experimenter asked the participant to try not to make any sudden large head movement throughout the experiment as this would disrupt the recording of the eye tracking data. The experimenter then initiated the testing phase of the experimental software and left the room.

The participant was required to carry out 40 trials for the entire experiment. The participant was allowed to take a short break for every 10 trials but he/she was reminded not to move away from the screen monitor or make large head movement. Finally, when the participant completed all the trials, the experimenter debriefed the participant and explained the real objective of the study was to study the effect of error consequence on data-entry performance.

2.2 Results.

Data from the 60 participants were analysed using a one-way ANOVA with Error Consequence as the independent variable on each of the three dependent variables: error rate, number of checks, and duration of checks. For the dependent variables: number of checks and duration of checks; eye fixations data were extracted using Tobii's ClearView Fixation Filter set at 100 ms and 500 ms. For each of the analyses, two planned comparisons between the three conditions were performed: i) Control vs Reward and Punishment combined; and ii) Reward vs Punishment. The planned comparisons were chosen, on the basis of Li et al. (2016)

findings, to test whether Reward or Punishment results in improved data-entry performance compared to Control; and to test for differences between Reward and Punishment. Table 2 summarises the descriptive results.

< Insert Table 2 here >

2.2.1 Error rate.

There were a total of 246 errors (out of 19,200 error opportunities) across all three conditions giving an overall mean error rate of 1.3% (SD = 1.4). The current error rate is comparable to Li et al. (2016) Experiment 1 (1.2% (SD = 1.5)).

With regard to the effect of Error Consequence, the one-way ANOVA yielded a significant effect, $F(2, 57) = 5.319$, $p = .008$, $\omega^2 = .13$[1]. Two planned comparisons were performed: the first one compared Control to the two experimental conditions combined (Reward and Punishment) and found a significant effect, $t(57) = 3.261$, $p = .002$ (two-tailed), $r = .4$[2]; the second contrast compared Reward with Punishment and found a non-significant effect, $t(57) = -0.078$, $p = .938$ (two-tailed). These results are consistent with those obtained in Li et al. (2016) Experiment 1.

2.2.2 Number of checks.

Fixation of 100 ms. The number of checks violated the homogeneity of variances assumption, $F(2, 57) = 5.834$, $p = .003$. Therefore, Welch's F correction was used; the effect of Error Consequence on the number of checks was not significant, $F(2, 37.105) = 0.697$, $p = .505$.

---

[1] For ω^2, effect sizes of .01, .06 and .14 are suggested as small, medium and large respectively (Kirk, 1996).

[2] For r, effect sizes of .1, .3 and .5 are suggested as small, medium and large respectively (Cohen, 1992).

Fixation of 500 ms. When fixations were extracted using the 500-ms filter, the effect of Error Consequence on the number of checks was also not found to be significant, $F(2, 57) = 0.011$, $p = .989$. Therefore, no further comparison was performed on the number of checks between the three conditions.

2.2.3 Duration of checks.

Fixation of 100 ms. The effect of Error Consequence on duration of checks was not found to be significant, $F(2, 57) = 0.245$, $p = .784$.

Fixation of 500 ms. With the fixation criterion set to 500 ms, the effect of Error Consequence on duration of checks was also found not significant, $F(2, 57) = 1.155$, $p = .322$.

2.3 Discussion.

The first hypothesis predicted more checks but each check shorter in the current experiment when compared to Li et al. (2016) Experiment 1. Results from cross-experiment comparisons should be interpreted with caution because of potential unanticipated variability between experiments. Therefore, we have only carried out an informal rather than statistical comparison between the current results (Table 2) and Li et al.'s Experiment 1 (Table 1) on the number of checks and duration. In Li et al.'s Experiment 1, the mean number of checks and duration, across all conditions, were 6.5 and 1911.9 ms respectively. When the 100-ms fixation filter was used for analysis, the mean number of checks and duration were 24.8 and 206.2 ms respectively. This is consistent with our prediction and can be accounted for by the information access cost: when checking was less effortful (as in our current experiment) participants checked more often but with shorter duration due to the ease of information access.

The second hypothesis predicted reward and punishment leading to more accurate data-entry performance *and* more rigorous checking than no error consequences at all. With respect to data-entry accuracy, the current results replicate those of Li et al.'s Experiment 1 (Table 1) suggesting that error consequences, either as reward or punishment, led to fewer errors relative to no consequence at all; the magnitude of error rate reduction is about twofold: 0.9% in Reward and Punishment vs 2% in Control. Furthermore, the overall error rates obtained in the current experiment (1.3%) and in Li et al.'s Experiment 1 (1.2%) are similar in magnitude.

However, the checking behaviour part of the hypothesis is not supported by the current results (see Table 1 & 2). The current data suggest that error consequences, in terms of reward or punishment, did not motivate participants to carry out more and longer checks than nil error consequence. This finding is in contrast to Li et al.'s results which suggest that reward and punishment motivate participants to perform more and longer checks. The current result implies accurate data-entry performance may not necessarily depend on checking frequency and duration. The use of an eye tracker in the current experiment removed the need to uncover grey boxes to carry out checking, as in Li et al., and allowed participants to perform more "natural" checking by eye fixations. Research in reading has found that limited information can be picked up and processed by peripheral vision (Chung, 2004; Latham & Whitaker, 1996; Rayner, 1975). It could be that when a participant looked at, for example, item #2 in the list of to-be-transcribed information on the left panel, peripheral vision has helped the participant pick up information about item #3. Therefore, each eye fixation might have also aided assessing the accuracy of more than the item in focus.

Regarding the third hypothesis that reward would result in more accurate performance and more rigorous checking than punishment; the current result does not support the hypothesis. The non-significant results between the reward and punishment conditions might be due to the design of the reward and punishment mechanism. The reward and punishment consequences were implemented such that one incorrect trial would result in being punished (or not rewarded). Because of this one-shot effect of reward and punishment, it may be that participants who were aware they had made an error then decided it was not worth their time and effort to perform more checks. This would explain why there was no difference observed between reward and punishment in terms of data-entry performance and checking behaviour. In order to test this explanation, Experiment 2 was designed to implement incremental reward and decremental punishment in order to remove the one-shot reward and punishment design.

3. Experiment 2

The objective of this second experiment was to examine how the graduated approach to reward or punishment, and the inclusion of performance feedback would alter the effect of monetary reward and punishment on data-entry performance and checking behaviour. Based on findings from neuroscience studies that suggested different neural circuitry for responding to reward and punishment (Wrase et al., 2007; Yeung & Sanfey, 2004) and behavioural studies that found reward effective in improving performance such as creativity (Eisenberger & Cameron, 1996; Eisenberger & Rhoades, 2001) and motivation (Eisenberger et al., 1999; Hendijani et al., 2016; Lopez, 1981), we hypothesized that reward, in the form of gradual increment, would result in more accurate data-entry performance than punishment in the form of gradual decrement. Regarding checking behaviour, with respect to the finding from the previous experiment, we expected improved data-entry performance to be independent of the number of checks and duration (as indicated by the eye-tracking data). In other words, we predicted there would be no difference in the number and duration of checks between reward, punishment and the control condition.

Two changes were made to the experimental task. The first one was to change from a one-shot approach to reward or punishment to a graduated approach. In the previous experiment, the one-shot approach was that participants receive the entire reward (+HK$40 (+US$5)) or punishment (-HK$40 (-US$5)) on the basis of one single incorrectly transcribed trial. In the current experiment, the reward and punishment were calculated on a trial-by-trial basis, i.e., participants received part of the overall reward (+HK$1(+US$0.1)) or punishment (-HK$1(-US$0.1)) based on the performance of each individual trial. The second change was to add in immediate feedback for the participants to monitor their performance. The independent variable remained the same as in the previous experiment, namely, error consequences in terms of reward or punishment.

3.1 Method.

3.1.1 Participants.

Sixty participants (age range: 17 – 26 years; mean age: 20.5 years) from Lingnan University were recruited for the experiment. There were 42 females and 18 males.

3.1.2 Design.

The same experimental design, dependent and independent variables as Experiment 1 were used. The one exception was that the reward and punishment manipulations were implemented on a per trial basis so that the consequence of making a single error had a significantly smaller impact on the overall reward or punishment.

3.1.3 Materials and task.

The same basic materials as Experiment 1 were used. However, the experimental task software was re-programmed to incorporate the two changes: 1) each trial, depending on whether it is correctly or incorrectly transcribed, had an impact on the payment in the Reward and Punishment conditions; 2) a display was added to the bottom of the screen (Figure 1) to indicate to a participant the payment they would receive (i.e. performance feedback).

3.1.4 Procedure.

The same procedure as Experiment 1 was adopted except that the participants were informed about the display at the bottom of the screen, which indicated their payment amount. Furthermore, participants in the Reward and Punishment condition received the following verbal instructions:

Reward condition: "*When you finish transcribing all the trials, we will pay you $60 for your time. However, for each correctly transcribed trial, you will get an extra $1, which you can see at the bottom of the screen. Therefore, you could end up with a maximum of $100.*

Punishment condition: "*When you finish transcribing all the trials, we will pay you $100 for your time. However, for each incorrectly transcribed trial, you will lose $1, which you can see at the bottom of the screen. Therefore, you could end up with $60 only.*

3.2 Results.

Data from 60 participants were analysed using a one-way independent ANOVA with the same independent and dependent variables as Experiment 1. Table 3 shows the descriptive results across all conditions.

< Insert Table 3 here >

3.2.1 Error rate.

The total number of errors was 249 (out of 19,200 error opportunities) and the overall mean error rate was 1.3% (SD = 1.3). The overall mean error rate is the same as that of the previous experiments.

The obtained error rates violated the homogeneity of variances assumption ($F(2, 57) = 7.316$, $p = .001$); therefore, the Welch's F correction was used. The differences in error rates between the three conditions were significant, $F(2, 30.405) = 4.77$, $p = .016$. Planned comparison of Control with the two experimental conditions was not significant, $t(24.756) = 1.583$, $p = .126$ (two-tailed). But the planned comparison between Reward and Punishment was significant, $t(24.965) = -2.183$, $p = .039$ (two-tailed), $r = .4$. Reward resulted in fewer errors than Punishment. This result of error rates is different from that of the previous experiments and of Li et al.'s Experiment 1.

3.2.2 Number of checks.

Fixation of 100 ms. The number of checks did not significantly differ among the three conditions when the fixation criterion was set to 100 ms, $F(2, 57) = 0.621$, $p = .541$.

3.2.3 Duration of checks

Fixation of 100 ms. There were no significant differences in the duration of checks across the three conditions, $F(2, 57) = 0.389$, $p = .679$.

3.3 Discussion.

The current experiment obtained the same overall error rate (1.3%) as the previous experiment. Although only an informal comparison was carried out because of constraints of cross-experiment comparison, the result provides an indication of reliability of the experimental paradigm.

There were two findings obtained in the current experiment. Firstly, the result on error rate was as predicted: the reward consequence led to fewer errors than no consequence. More specifically, the reward consequence reduced the error rate to about a half (0.8%) compared with no consequence (about 1.7%). This result suggests that when each individual trial was associated with a consequence and immediate performance feedback was provided, reward was effective in reducing errors.

With regard to the effect of punishment, the result suggests that it did not significantly improve performance accuracy (error rate of 1.4%) when compared to no consequence. However, we cannot be confident about the non-significant effect of punishment because of the modest sample size in this experiment (N = 60; 20 in each group). A post-hoc power analysis was carried out using G*Power (Faul, Erdfelder, Buchner, & Lang, 2009). Based on the error rate means, the power analysis revealed that, for a between-groups effect size obtained in the present experiment (f = 0.31), a sample of size of 105 (35 in each group) would be needed to achieve statistical power at the recommended .8 level. Therefore, the effect of punishment with immediate feedback remains inconclusive at present.

However, the beneficial effect of the reward consequence in improving data-entry accuracy is consistent with other findings on enhancing creativity (Eisenberger & Cameron, 1996; Eisenberger & Rhoades, 2001) and motivation (Eisenberger et al., 1999; Hendijani et al., 2016; Lopez, 1981). In the current experimental setup, the display of immediate performance feedback and the magnitude of the starting monetary amount are displayed at the bottom of the computer screen. Participants in the Reward condition began the experiment with a monetary amount of HK$60 (~US$8), and the amount would increase by HK$1 (~US$0.1) for each correctly transcribed trial. The potential of being able to earn up to a total of HK$100 (~US$13) and the trial-by-trial update of one's incremental success had probably motivated the participants to transcribe accurately. In contrast, participants in the

Punishment and the Control conditions were initiated with HK$100 (~US$13) displayed at the bottom of their screen. It is possible that the display of HK$100 (~US$13) might have encouraged participants to become less vigilant than the Reward participants in monitoring their own performances. It is worth noting that the average error rate for Punishment was only 1.4% (i.e., on average, less than one trial among the 40 trials would be incorrect), this means that the monetary display in this condition had continuously displayed close to HK$100 (~US$13) for the entire period of the experiment. A loss of HK$1 (~US$0.1) might not be as effective a motivator for better performance when compared to a steady increment of gain as experienced by participants in the Reward condition.

Note that the superior effect of gradual reward is based on its incremental incentive and not only on feedback. The punishment (i.e. gradual decrement) condition also contained performance feedback for participants and any comparison between the two conditions (reward vs punishment) would have controlled for any difference in feedback effect. As such it seems unlikely that the effect of gradual incentive was confounded by the presence of feedback.

The second finding suggests that there were no significant differences in the number and duration of checks, as indicated by eye fixations and duration, between the three error consequence conditions. This result replicates the one from the previous experiment suggesting that data-entry accuracy does not necessarily depend on the frequency or duration of eye fixations when performing checks.

The current results lead us to conclude that the graduated approach to reward and punishment and the immediate performance feedback were able to differentiate between the effects of reward and punishment. When the performance feedback contained gradual increment (as in the Reward condition) it resulted in more accurate data-entry performance than when the feedback displayed gradual decrement (as in the Punishment condition). Moreover, the improved data-entry accuracy was not associated with more frequent and longer eye fixations in checking.

4. General discussion

Experiment 2 shows that a reward consequence in the form of increment was able to motivate more accurate performance than no consequence at all. This finding complements Li et al. (2016) who found the same result when using the one-shot approach to assigning reward or punishment. This finding introduces a new way of conceptualising reward and punishment as gradual increment and decrement feedback respectively; and we conclude that reward with (as in Experiment 2) or without (as in Experiment 1) gradual increment feedback are both more effective than no consequence in encouraging accurate data-entry performance. Although we cannot draw conclusions about the effect of punishment with decrement feedback from Experiment 2, results from Experiment 1 suggest that punishment was able to lead to more accurate data entry than no consequence at all. However, further research is needed with an adequate sample size to compare the effect of gradual reward increment with punishment decrement. Nevertheless, the conceptualisation of increment and decrement feedback brings out a practical issue relating to establishing a healthy safety culture in organisations. Li et al. (2016) point out that if an organisation wants its workers to honestly report their errors or mistakes in order to learn from them, then it is important to set up a blame-free, but not accountability-free, environment (Dekker, 2009; Woods et al., 2010). If error consequences are associated with punishment then workers are likely to cover up any committed errors in order to avoid the punishment; and, consequently, the organisation will not be able to learn from its actions. Therefore, the current finding implies that it is more effective, in the long-run, to reward desired behaviour with increment than to punish undesired behaviour with decrement especially when error reporting is an important part of an organisation's safety operation.

Previous research on data-entry performance (e.g. Barchard & Pace, 2011) has shown that even when participants were asked to visually check the data they entered they still made an average error rate of 0.8% (633.79 errors out of 76860 entries)[3]. It is only when participants were instructed to carry out double entry (i.e. enter the data twice) the average error rate decreased to 0.03% (22.1 out of 81900 entries)[4]. Our results suggest that when visual checking was coupled with motivational incentives, such as rewarding participants, it was effective in reducing error rates to 0.9% (Experiment 1) and 0.8% (Experiment 2). We are not making formal comparisons between our obtained error rates and Barchard and Pace's because there are many differences between the two studies. But the general results do suggest that visual checking, even when coupled with motivational incentives, is still not as effective as double entry in reducing data-entry errors.

Both experiments show that accurate data-entry performance was not related to the number and duration of eye fixations exhibited by the participants. This finding is surprising as it is in contrast to previous research (Li et al., 2016). We conclude that accurate data-entry does not necessarily depend on more frequent and longer checking on the basis that the checking behaviour afforded by eye fixations is more natural and involved less effort than Li et al. (2016), where the paradigm required participants to use mouse movement for checking. This conclusion is, to a degree, counterintuitive as one would expect more frequent and longer checking to lead to more accurate performance. However, the current data suggest that

[3] The 633.79 total number of errors was obtained by average no. of errors (10.39) × no. of participants (61). The total number of entries across all participants was obtained by no. of items entered per data sheet (1 + 1 + (4 × 10) = 42) × no. of data sheet entered by each participant (30) × no. of participants in that condition (61).

[4] The 22.1 total number of errors was obtained by average no. of errors (0.34) × no. of participants (65). The total number of entries across all participants was obtained by no. of items entered per data sheet (1 + 1 + (4 × 10) = 42) × no. of data sheet entered by each participant (30) × no. of participants in that condition (65).

performance accuracy was independent of frequency and duration of checking and this provides further evidence to highlight that visual checking alone may not be effective in detecting errors, which was also suggested by Barchard and Pace (2011). We have identified three possible explanations for this unexpected finding: first, in the current experimental paradigm the to-be-transcribed information was no longer occluded by grey boxes and participants could carry out checking by eye movements. It could be that peripheral vision has helped participants in assessing the accuracy for neighbouring to-be-transcribed information and this, in turn, reduced the overall number and duration of checks. However, the precise role and benefit of peripheral vision in data-entry performance will require further careful experimentation as the current experiments were not designed for such investigation. Second, some participants showed zero fixations in their data and they might have been using an error prevention, rather than error detection, strategy during data entry. Perhaps these participants were not looking at the textbox they were entering data into but were looking at their fingers as they typed. This would imply that some participants might prefer preventing errors to visually checking for errors. The exact nature of error prevention or detection strategies adopted by participants in data entry can only be determined by future studies with appropriately designed experiments, for example, Crump and Logan (2013) examined typing strategies in skilled typists and found that most typists preferred correcting errors as they occurred rather than preventing them. Third, it is possible that accurate data-entry depends on the quality rather than the quantity (as defined by how frequent and how long in the current study) of checks performed. With the current data, we are unable to provide concrete definitions for operationalising checking qualities but they might be related to covert behaviour that cannot be directly measured. Research on metacognition might be related to error detection ability. Metacognition refers to the abilities to monitor, regulate, coordinate and control cognition (Fernandez-Duque, Baird, & Posner, 2000) and it has been studied in relation to error monitoring in decision-making (Yeung & Summerfield, 2012), reading (Pressley & Ghatala, 1990) and education (Hacker, Dunlosky, & Graesser, 1998). Flavell (1979) proposed that metacognition can be affected by metacognitive knowledge about one's cognitive abilities (e.g. "I easily mistype numbers"), about cognitive strategies (e.g. "I will speak out what I type to make what I type is intended") and about tasks (e.g. "entering all the number fields first is faster than switching between number and name fields"). It is possible that the manipulations of reward and punishment in the current experiments have affected metacognitive knowledge that was not measured. Therefore, for future research, devising ways of measuring one's metacognitive knowledge or abilities might provide a way to understand how one's checking quality might differ in relation to reward and punishment consequences.

4.1. Limitations

There are a number of limitations in the current study. Firstly, the practical application of the reward and punishment manipulation may be limited when the correct data values are not known in real data-entry situations. However, Li et al. (2016) suggest that:

> ... in many data-entry situations it is possible to verify one's performance accuracy against the original data source. For example, when entering student marks from exam scripts into university system, it is possible to check the accuracy of the marks entered; when transcribing medication information into a CPOE, it is possible to verify the entered information against the original paper chart. (p. 265)

But the data verification cannot be done in real-time with the data-entry because the verification is necessarily an off-line process. This poses a practical limitation to the immediate feedback mechanism of incremental reward or punishment in Experiment 2 because in real data-entry tasks real-time verification of the entered data is not feasible. Therefore, as Li et al. (2016) point out:

> … if data-entry performance is to be verified in an actual work environment, it is likely to involve a data-entry system to couple with a separate verification mechanism. … Such a verification mechanism is, of course, only possible when double entry is adopted as a data checking method.” (p. 265)

Despite the practical limitation of Experiment 2’s immediate accuracy feedback, the important point of the finding is that reward resulted in more accurate data-entry performance. Future research can investigate the effect of delayed feedback on data-entry performance as it may be possible to implement a verification mechanism to give delayed rather than immediate feedback. This would address the practical limitation of the current manipulation. An alternative solution is to provide training to change the way people enter data, for example, the training may place specific emphasis on forming checking habits. Previous work on error detection suggests that most errors were detected when people routinely checked their performance out of habit (Allwood, 1984; Nyssen & Blavier, 2006).

Secondly, further research is needed to draw valid conclusions on the true effect of punishment with gradual decrement feedback because of the modest sample size tested in Experiment 2.

5. Conclusions

Error detection and correction is an important aspect of human performance and it has direct practical implications to mission- and safety-critical operations. The notion of increment and decrement offers a new way of designing feedback to enhance error detection performance. The current experiments and those in the Li et al. (2016) study are the first series of experiments that directly compare the effect of reward and punishment in monetary terms. Although we cannot draw conclusions about the effectiveness of punishing decrement, the current evidence shows that, when provided with immediate feedback, rewarding increment can be effective in motivating accurate data-entry performance. Moreover, accurate performance does not necessarily depend on overt checking behaviour.

6. Acknowledgements

We would like to thank Sze-yuen Yau and Man-wa Ng for data collection, and Anthony Lam for programming. This work was supported by the Hong Kong Research Grants Council under Early Career Scheme (LU342912).

(Word count for main text: 6,975)

7. References

Abe, M., Schambra, H., Wassermann, Eric M., Luckenbaugh, D., Schweighofer, N., & Cohen, Leonardo G. (2011). Reward improves long-term retention of a motor memory through induction of offline memory gains. *Current Biology, 21*(7), 557-562. doi:10.1016/j.cub.2011.02.030

Allwood, C. M. (1984). Error detection processes in statistical problem solving. *Cognitive Science, 8*(4), 413-437. doi:10.1207/s15516709cog0804_5

Back, J., Cheng, W. L., Dann, R., Curzon, P., & Blandford, A. (2007). Does being motivated to avoid procedural errors influence their systematicity? *People and Computers XX—Engage* (pp. 151-157): Springer.

Barchard, K. A., & Pace, L. A. (2011). Preventing human error: The impact of data entry methods on data accuracy and statistical results. *Computers in Human Behavior, 27*(5), 1834-1839. doi:10.1016/j.chb.2011.04.004

Barchard, K. A., & Verenikina, Y. (2013). Improving data accuracy: Selecting the best data checking technique. *Computers in Human Behavior, 29*(5), 1917-1922. doi:10.1016/j.chb.2013.02.021

Blavier, A., Rouy, E., Nyssen, A.-S., & De Keyser, V. (2005). Prospective issues for error detection. *Ergonomics, 48*(7), 758-781. doi:10.1080/00140130500123670

Brumby, D. P., Cox, A. L., Back, J., & Gould, S. J. J. (2013). Recovering from an interruption: Investigating speed−accuracy trade-offs in task resumption behavior. *Journal of Experimental Psychology: Applied, 19*(2), 95-107. doi:10.1037/a0032696

Chung, S. T. L. (2004). Reading speed benefits from increased vertical word spacing in normal peripheral vision. *Optometry and Vision Science: Official Publication of the American Academy of Optometry, 81*(7), 525-535.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159. doi:10.1037/0033-2909.112.1.155

Crump, M. J., & Logan, G. D. (2013). Prevention and correction in post-error performance: An ounce of prevention, a pound of cure. *Journal of Experimental Psychology: General, 142*(3), 692. doi:10.1037/a0030014

Data entry is a top cause of medication errors. (2005). Retrieved from http://www.amednews.com/article/20050124/profession/301249959/4/

Dekker, S. W. A. (2009). Just culture: Who gets to draw the line? . *Cognition, Technology & Work, 11*(3), 177-185. doi:10.1007/s10111-008-0110-7

Eisenberger, R., & Cameron, J. (1996). Detrimental effects of reward: Reality or myth? *American Psychologist, 51*(11), 1153-1166. doi:10.1037/0003-066X.51.11.1153

Eisenberger, R., & Rhoades, L. (2001). Incremental effects of reward on creativity. *Journal of Personality and Social Psychology, 81*(4), 728-741. doi:10.1037/0022-3514.81.4.728

Eisenberger, R., Rhoades, L., & Cameron, J. (1999). Does pay for performance increase or decrease perceived self-determination and intrinsic motivation? *Journal of Personality and Social Psychology, 77*, 1026-1040. doi:10.1037/0022-3514.77.5.1026

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149-1160. doi:10.3758/brm.41.4.1149

Fernandez-Duque, D., Baird, J. A., & Posner, M. I. (2000). Executive attention and metacognitive regulation. *Consciousness and Cognition, 9*(2), 288-307. doi:10.1006/ccog.2000.0447

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist, 34*(10), 906-911. doi:10.1037/0003-066X.34.10.906

Gray, W. D., & Fu, W.-T. (2004). Soft constraints in interactive behavior: The case of ignoring perfect knowledge in-the-world for imperfect knowledge in-the-head. *Cognitive Science, 28*(3), 359-382. doi:10.1016/j.cogsci.2003.12.001

Hacker, D. J., Dunlosky, J., & Graesser, A. C. (Eds.). (1998). *Metacognition in educational theory and practice*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Hendijani, R., Bischak, D. P., Arvai, J., & Dugar, S. (2016). Intrinsic motivation, external reward, and their effect on overall motivation and performance. *Human Performance, 29*(4), 251-274. doi:10.1080/08959285.2016.1157595

Kirk, R. E. (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement, 56*(5), 746-759. doi:10.1177/0013164496056005002

Kohn, L. T., Corrigan, J. M., & Donaldson, M. S. (2000). *To err is human: Building a safer health system*. . Washington, DC: National Academy Press.

Latham, K., & Whitaker, D. (1996). A comparison of word recognition and reading performance in foveal and peripheral vision. *Vision Research, 36*(17), 2665-2674. doi:10.1016/0042-6989(96)00022-3

Li, S. Y. W., Cox, A. L., Or, C., & Blandford, A. (2016). Effects of monetary reward and punishment on information checking behaviour. *Applied Ergonomics, 53, Part A*, 258-266. doi:10.1016/j.apergo.2015.10.012

Lopez, E. M. (1981). Increasing intrinsic motivation with performance-contingent reward. *The Journal of Psychology, 108*(1), 59-65. doi:10.1080/00223980.1981.9915245

Lourens, P. F. (1990). Theoretical perspectives on error analysis and traffic behaviour. *Ergonomics, 33*(10-11), 1251-1263. doi:10.1080/00140139008925330

Nyssen, A. S., & Blavier, A. (2006). Error detection: A study in anaesthesia. *Ergonomics, 49*(5-6), 517-525. doi:10.1080/00140130600568766

Potts, G. F. (2011). Impact of reward and punishment motivation on behavior monitoring as indexed by the error-related negativity. *International Journal of Psychophysiology, 81*(3), 324-331. doi:10.1016/j.ijpsycho.2011.07.020

Pressley, M., & Ghatala, E. S. (1990). Self-regulated learning: Monitoring learning from text. *Educational Psychologist, 25*(1), 19-33. doi:10.1207/s15326985ep2501_3

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7*(1), 65-81. doi:10.1016/0010-0285(75)90005-5

Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin, 85*(3), 618-660. doi:10.1037/0033-2909.85.3.618

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition, 14*(3), 191-201. doi:10.3758/bf03197692

Reason, J. (1990). *Human error*. New York: Cambridge University Press.

Sellen, A. J. (1994). Detection of everyday errors. *Applied Psychology, 43*(4), 475-498. doi:10.1111/j.1464-0597.1994.tb00841.x

Stürmer, B., Nigbur, R., Schacht, A., & Sommer, W. (2011). Reward and punishment effects on error processing and conflict control. *Frontiers in Psychology, 2*. doi:10.3389/fpsyg.2011.00335

Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics, 106*(4), 1039-1061. doi:10.2307/2937956

Woods, D. D., Dekker, S., Cook, R., Johannesen, L., & Sarter, N. (2010). *Behind human error*. Aldershot, UK: Ashgate.

Wrase, J., Kahnt, T., Schlagenhauf, F., Beck, A., Cohen, M. X., Knutson, B., & Heinz, A. (2007). Different neural systems adjust motor behavior in response to reward and punishment. *NeuroImage, 36*(4), 1253-1262. doi:10.1016/j.neuroimage.2007.04.001

Yeung, N., & Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *The Journal of Neuroscience, 24*(28), 6258.

Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: confidence and error monitoring. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*(1594), 1310-1321. doi:10.1098/rstb.2011.0416

Zapf, D., Maier, G. W., Rappensperger, G., & Irmer, C. (1994). Error detection, task characteristics, and some consequences for software design. *Applied Psychology, 43*(4), 499-520. doi:10.1111/j.1464-0597.1994.tb00842.x

Figure

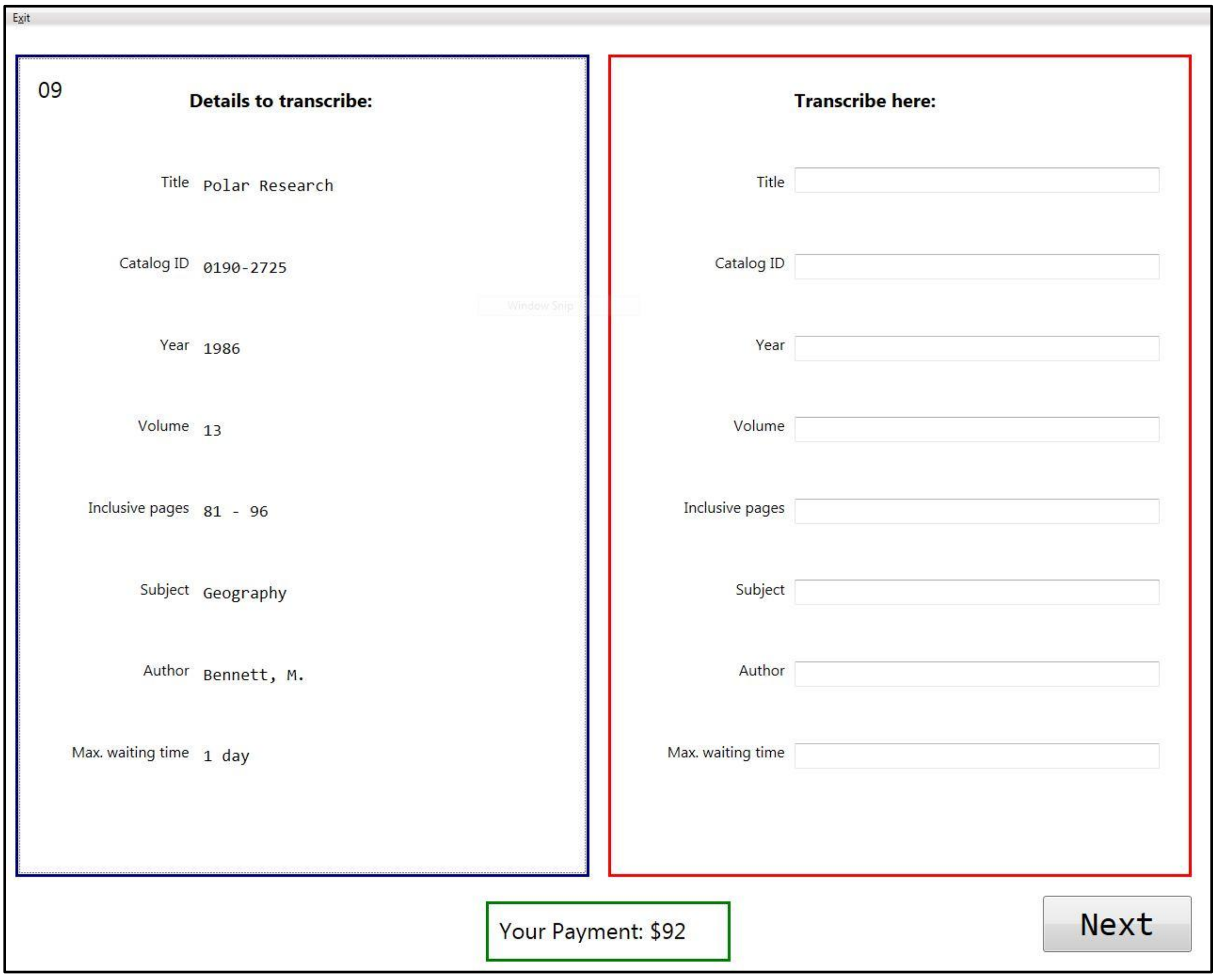

Figure 1: A screenshot of the experimental task in Experiment 2. The payment feedback counter (bottom of the screen) shows increment by HK$1 (~US$0.1) for each correctly transcribed trial (in the Reward condition) or decrement by HK$1 (~US$0.1) for each incorrectly transcribed trial (in the Punishment condition). The experimental task in Experiment 1 has identical appearance but without the payment feedback counter at the bottom of the screen.

Table

Error Consequence conditions [a]	No. of errors [b]	Error rate [c] *M (SD)*	No. of checks *M (SD)*	Duration of checks (ms) *M (SD)*
Control	173	1.8 (1.9)	4.2 (2.6)	1486.3 (512.0)
Reward	94	0.92 (1.2)	7.1 (3.6)	2108.7 (831.4)
Punishment	88	0.98 (1.1)	8.2 (4.0)	2140.7 (978.2)

[a] For each condition, n = 30.
[b] No. of error opportunities for each condition = 9,600.
[c] Error rate is a percentage calculated as a ratio of no. of errors to the no. of error opportunities.

Table 1: Data reproduced from Li et al.'s (2016) Experiment 1

Error Consequence conditions [a]	No. of errors [b]	Error rate [c] *M (SD)*	No. of checks *M (SD)*		Duration of checks (ms) *M (SD)*	
			100-ms fixation	500-ms fixation	100-ms fixation	500-ms fixation
Control	130	2.0 (1.2)	28.4 (14.7)	2.7 (2.1)	218.7 (93.8)	521.5 (281.0)
Reward	57	0.9 (0.9)	22.0 (22.1)	2.7 (2.9)	206.8 (132.3)	415.3 (299.9)
Punishment	59	0.9 (1.6)	24.1 (16.7)	2.6 (3.3)	193.0 (119.4)	383.2 (321.2)

[a] For each condition, $n = 20$.
[b] No. of error opportunities for each participant is 320.
[c] Error rate is a percentage calculated as a ratio of no. of errors to the no. of error opportunities. This error rate is calculated across all participants, therefore, the no. of error opportunities for each condition is 320 × 20 = 6400.

Table 2: Experiment 1 data

Error Consequence conditions [a]	No. of errors [b]	Error rate [c] M (SD)	No. of checks M (SD)	Duration of checks (ms) M (SD)
			100-ms fixation	100-ms fixation
Control	111	1.7 (1.7)	3.1 (5.7)	148.6 (168.0)
Reward	48	0.8 (0.5)	2.3 (3.0)	163.9 (178.0)
Punishment	90	1.4 (1.2)	4.1 (5.7)	198.0 (196.6)

[a] For each condition, $n = 20$.
[b] No. of error opportunities for each participant is 320.
[c] Error rate is a percentage calculated as a ratio of no. of errors to the no. of error opportunities. This error rate is calculated across all participants and conditions, therefore, the no. of error opportunities is $320 \times 20 = 6400$.

Table 3: Experiment 2 data