

Preprint: final version available as: BACK, J., CHENG, W.L., DANN, R., CURZON, P. & BLANDFORD, A. (2006) Does being motivated to avoid procedural errors influence their systematicity? *Proc. HCI 2006 (Vol. 1)*

Does being motivated to avoid procedural errors influence their systematicity?

**Jonathan Back¹, Wai Lok Cheng², Rob Dann²,
Paul Curzon² & Ann Blandford¹**

¹ *UCL Interaction Centre, UCL, Remax House, 31-32 Alfred Place, London WC1E 7DP, UK.*

Email: {j.back, a.blandford}@ucl.ac.uk

² *Department of Computer Science, QMUL, Mile End Road, London E1 4NS, UK.*

Email: {wlc1, rjd1, pc}@dcs.qmul.ac.uk

URL: <http://www.dcs.qmul.ac.uk/research/imc/hum/>

Post-completion error (PCE) is a type of procedural error that occurs even when an individual has the knowledge required to perform the task correctly. After performing a ‘task critical’ step a PCE can occur if an individual forgets to perform a required ‘related step’ before starting a new task. Research has demonstrated that PCE is sensitive to external influences. For example, forgetting to collect the original document after making photocopies is more likely if an individual is thinking about ‘other things’ or is interrupted. Two gaming environments were designed to test the systematicity of PCE where participants were actively trying to avoid them. In both experiments a participant’s score was reset to zero if a PCE occurred. Results showed that this penalty did not significantly influence the systematicity of the error within a gaming session. It was also found that the likelihood of PCEs occurring could be predicted by the intrinsic difficulty associated with performing the ‘task critical’ step or by an individual’s ability to accurately remember information relevant to the task. This implies that even when individuals are motivated to avoid PCEs, user performance remains vulnerable to this error type. Within demanding environments, this vulnerability is likely to be more exposed.

Keywords: post-completion error, procedural error, motivation, priming, games.

1 Introduction

Errors are one measure of the quality of human performance. For example, Miller [1956] identified an important property of working memory by discovering that individuals make errors when recalling more than 7 (+/-2) unrelated elements of information. However, the everyday concept of error presupposes related goals that are cued by the environment. For example, when making coffee for a friend you do not have to recall whether they take sugar until the coffee has brewed. Moreover, the importance that an individual assigns to a particular goal may influence how the goal is performed. If the goal is critical to achieving a task then an individual may be motivated to ensure that performance is error free. This paper reports an empirical investigation into whether motivation within an engaging task paradigm can reduce the systematicity of a specific type of procedural error known as post-completion error (PCE).

The opportunity to make a PCE arises after the execution of a 'task critical' action that achieves a task. If an individual then forgets to perform a required procedural step before commencing a new task then a PCE occurs. Examples include forgetting to collect your change after buying chocolate from a vending machine, forgetting the original document after making photocopies, and forgetting to turn the gas off after cooking a meal. PCE is not associated with a lack of knowledge. This type of error can still occur when individuals have the required procedural knowledge to perform a task correctly. Norman [1981] identified that such 'slips' can be attributed to either a failure to perform necessary attentional monitoring, making an inappropriate attentional check, or forgetting. A PCE occurs when a step is forgotten. This is due to the post-completion step not being a necessary precondition to completing a task. PCEs are caused by a failure in prospective memory. Prospective remembering is the process and skill required to support the fulfilment of an intention to perform a specific action in the future. They are intentions that we can not put into effect at the time we form them [Ellis & Kvavilashvili 2000]. A critical aspect of success is not only recall of the content of the task but also retrieval at an appropriate moment for action. Consequently, there is a move towards understanding how the task environment influences working memory, attention and other cognitive functions.

It is not satisfactory to suggest specific error manifestations are a result of a generally high task performance error rate. It is now recognised that errors in routine interactive behaviour are not the product of some stochastic process, and that causal explanations of human error can be developed. Byrne and Bovair [1997] showed that PCE is sensitive to the working memory demands of the task environment. People make errors frequently, but do not make errors every time. Current experimental research on human error has successfully designed situations where certain types of errors can be provoked [Gray 2000]. This involves the development of procedures that ensure errors are not associated with a lack of procedural knowledge but occur due to the cognitive phenomenon being investigated (training trials ensure participants know what they are supposed to be

doing). Recent research has identified various factors that affect the PCE rate such as, for example, dynamic visual cues [Chung & Byrne 2004]. It has also been shown that PCE occurs in non-routine problem-solving situations [Li et al. 2005]. In an attempt to further examine the robustness of PCE, this paper reports on two experiments that investigated whether motivating a participant to avoid PCE in an engaging task environment influenced systematicity.

2 Experiment 1: Space invaders

In the space invader game experiment¹ participants were required to shoot alien space ships, capture falling aliens, and rescue falling astronauts. A PCE is possible after rescuing an astronaut and then forgetting to change from rescue mode to shooting mode (*see* Figure 1, final instruction). The objective of the game is to obtain a place on the high score table. If a PCE is made, a player's score is reset to zero. Gamers are therefore likely to be motivated to avoid PCEs.











	Move your ship left and right using the arrow keys. Fire your gun using the space bar to shoot aliens.	
	Catch any astronauts falling from the sky, your ship will automatically bring out its catcher attachment.	
	Release astronauts over the yellow and black base.	
	Release aliens over the red and black prison.	
	Once you have released the astronaut, your ship will try to return to its normal state. After a couple of seconds, it will switch to look like the image on the left. When it is in this state, press G to switch back to gun mode.	
	Attempting to fire without pressing G will damage your weapon mechanism for a few seconds and cost you points!	

Figure 1: Space invaders instructions.

Data from twenty Queen Mary, University of London students was used for analysis². A within-subjects experimental design was used. All participants undertook a training level. Participants were required to repeat the training level

¹ This game can be found at www.dcs.qmul.ac.uk/cs4fn/humanerror/pce/

² Selected from a pool of 59 volunteers who played the game. 20 participants were selected on the basis that they completed training level and attempted levels 1-5 (not enough participants attempted level 6 for analysis).

until they successfully rescued two astronauts without making any PCEs (onscreen prompts were provided for the first astronaut rescue only). Level one also required two astronaut rescues. It was possible that participants were still learning how to play so error data was not considered for analysis until a participant started level two. Levels 2-6 required four astronaut rescues per level (i.e. there were four PCE opportunities per level). Each level required all aliens to be destroyed and there was an increase in difficulty between levels (faster aliens that were harder to shoot). Results show that participants systematically forgot to change from rescue mode to shooting mode (by pressing G). The systematicity of this error was 10.9% (levels 2-6). Table 1 shows the mean error rate and standard deviations for levels 2-5 (data not normally distributed). A within-subjects analysis was performed (number of PCEs made per level). A Friedman's 'test of four related samples' (non-parametric) revealed that there was no significant difference associated with PCE rates (Friedman's Chi-Square = 6.41, $df = 3$, $p > .05$). Despite participants being motivated to avoid PCE, the error remained systematic within a gaming session.

	N	Mean	Std. Deviation
Level 2	20	.300 (out of 4)	.571
Level 3	20	.700 (out of 4)	.657
Level 4	20	.350 (out of 4)	.587
Level 5	20	.400 (out of 4)	.503

Table 1: PCE rates for levels 2-5.

Further analysis showed that the error was not stochastic. It was found that the time it took to complete an astronaut rescue influenced the systematicity of the error. When no PCE was made the mean time taken to rescue was 3.74 seconds. However, when a PCE occurred the mean time was 4.60 seconds. A Mann-Whitney U non-parametric two independent samples test was performed (time taken to rescue: when a PCE occurs versus when a PCE does not occur). If an astronaut rescue was more difficult / time consuming (either due to distance to 'yellow and black base', or avoiding alien fire) then this significantly increased the likelihood of a PCE (Mann-Whitney U = 30233.5, Wilcoxon W = 253679.5, $Z = -3.518$, $p < .001$). Implications are discussed in Section 4.

3 Experiment 2: Driving game

In the driving game experiment [Cheng 2006] participants were required to capture cars of a specified colour (e.g. red) by moving the black police car (using the arrow keys). While in the process of capturing red cars participants were also required to remember specific details about cars of another specified colour (e.g. yellow) (*see* Figure 2). Details to remember included: the number of yellow cars, lane most frequently occupied by yellow cars (1-5), letters displayed on the roof of yellow cars, and whether a yellow car was seen before a car of any other colour when a new level was started (five levels in total). After a predetermined number

of red cars were captured (dependent on level difficulty) a participant was required to: move to the kerbside (using the arrow keys), apply the handbrake (pressing 'D'), submit a report form about yellow cars (completing the form and clicking submit), and then release the handbrake (pressing the spacebar) before attempting to move away from the kerbside (using the arrow keys). A PCE was possible after submitting a report if a participant forgot to release the handbrake before attempting to move away from the kerbside. If a PCE was made, a player's score was reset to zero. Specified car colours changed when a level was completed. After participants successfully completed a training session, they were required to play a control version of the game where no PCE cue was provided. Following the control version two different versions of the driving game were played: one where a textual PCE cue was provided (press space to release brake), and one where a symbolic cue was provided (flashing handbrake light). The order in which a participant attempted cue game versions was counterbalanced.

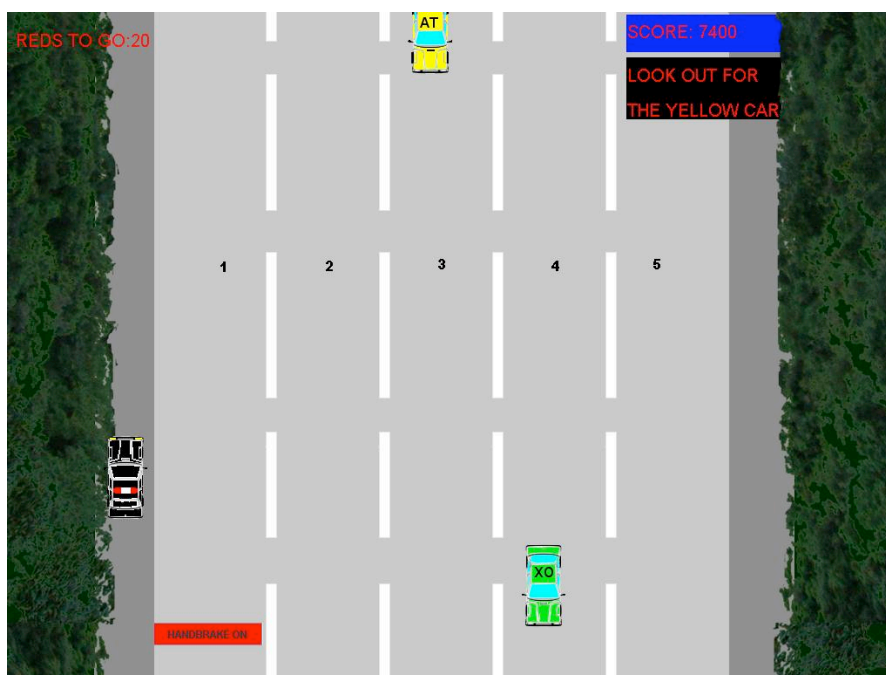


Figure 2: Driving game.

30 Queen Mary, University of London students volunteered. A within-subjects experimental design was used. Participants were required to repeat the training level until they successfully completed the level without making any PCEs (prompts were provided for the first report submission only). Results showed that participants systematically forgot to release the handbrake before attempting to drive away from the kerbside (systematicity with symbolic cue = 9.7%, with textual cue = 5.7%). Table 2 shows the mean error rate alongside mean question answer scores. A Wilcoxon signed rank test determined whether the textual cue

was significantly better at preventing PCE than the symbolic cue (number of PCEs: textual cue game versus symbolic cue game). Surprisingly, results show that the more salient textual cue was not significantly better (Wilcoxon $Z = -2.216$, $p > .05$). Further analysis showed that a participant's ability to answer questions was a good predictor of PCE rates (for both cue game versions: number of PCEs versus number of correct answers). If participants found it easier to remember answers they were less likely to make a PCE. A Wilcoxon signed rank test found a significant relationship (Wilcoxon $Z = -8.249$, $p < .001$).

	N	Mean Error	Std. Deviation	Mean Answer Scores	Std. Deviation
Symbolic	30	.970 (out of 10)	.765	17.73 (out of 40)	3.051
Textual	30	.570 (out of 10)	.626	18.13 (out of 40)	2.662

Table 2: PCE rates and answer scores for driving game versions.

4 Discussion and Conclusions

Previous laboratory studies on PCE have found that although the frequency of the error can be lowered by providing just-in-time cues [Chung & Byrne 2004] or reducing working memory demands [Byrne & Bovair 1997; Li et al. 2005], the error remains systematic. Previous experimentation has not, however, explicitly tried to motivate participants in a way that makes the avoidance of PCE the most important task consideration. Results from the space invaders experiment show that motivating participants in this way does not allow the error to be avoided within a gaming session. Moreover, motivation did not even significantly reduce the frequency of the error between levels. This finding provides further support to the theory that a 'priming process' is needed to facilitate prospective remembering [Ellis & Kvavilashvili 2000] since being motivated to remember to do something is not sufficient.

Altmann and Traflet's [2002] activation based goal memory model suggests that cognition is directed by the most active goal retrieved at any given time (e.g., avoiding alien fire while rescuing an astronaut). If an individual needs to refocus attention to achieve a previously formulated goal (e.g., remembering to hit 'G' to activate the gun after rescuing an astronaut) then this old goal needs to undergo a 'priming process' to become active. Associative links can act as primers (e.g., remembering to clean your shoes may be triggered by an intention to leave for work). Space invader participants usually avoided making a PCE, but sometimes the associated link, remembering to hit 'G' to activate the gun after rescuing an astronaut, was not strong enough if a participant was expending extra effort to avoid alien fire while rescuing an astronaut. It was found that if it took longer than usual to complete an astronaut rescue then the likelihood of a PCE occurring significantly increased. This is a novel finding since previous research on PCE has not considered that the intrinsic difficulty of the 'task critical' step might provide a

causal explanation for why PCEs occur. Findings remain consistent with Byrne & Bovair's [1997] working memory load theory.

The car driving experiment investigated whether the use of different cue types coupled with a participant's goal of avoiding PCEs influenced systematicity. Neither the textual nor symbolic cue allowed participants to avoid PCEs. Although the textual cue reduced the number of PCEs when compared to the symbolic cue, this difference was not significant. Interestingly, it was found that an individual's ability to remember information relevant to the task (in short term memory) could be used to predict the likelihood of PCEs. In this experiment the 'task critical' step was submitting a report. If the process of providing the report, i.e., remembering details about cars of a specific colour, was more difficult, then the strength of the associated link or 'priming process', i.e. remembering to release handbrake after submitting report, was lower and a PCE was more likely. This finding provides further evidence that the intrinsic difficulty of the 'task critical' step can be used as a causal explanation for PCE.

Designing engaging interaction environments might allow for more complex and consequently more useful interactions since a user may feel more confident, expressive, and motivated. Research presented within this paper has demonstrated the importance of minimising the complexity associated with 'task critical' steps that are likely to be used for 'priming processes'. Future research should identify these steps and ensure that allowing more complex interactions do not compromise 'priming processes' causing more procedural errors.

Acknowledgements

Work conducted as part of the HUMAN error Modelling project (HUM), funded by EPSRC (GR/S67494 and GR/S67500). We are grateful to all participants in the studies reported here.

References

- Altmann, E. M. & Trafton, J. G. [2002], Memory for goals: an activation-based model. *Cognitive Science* **26**(1), 39-83.
- Byrne, M. & Bovair, S. [1997], A Working Memory Model of a Common Procedural Error. *Cognitive Science* **21**(1), 31-69.
- Cheng, W. L. [2006], Investigation into the effects of different cues to reduce human errors: post-completion errors. MSci Dissertation, Department of Computer Science, Queen Mary, University of London.
- Chung, P. & Byrne, M. D. [2004], Visual cues to reduce errors in a routine procedural task. *Proceedings of the 26th Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Ellis, J. & Kvavilashvili, L. [2000], Prospective memory in 2000. *Applied Cognitive Psychology*, **14**, s1-s9.

Gray, W. [2000], The nature and processing of errors in interactive behavior. *Cognitive Science*, **24**(2), 205-248.

Li, S. Blandford, A., Cairns, P. & Young, R. M. [2005], Post-completion errors in problem solving. *Proceedings of the 27th Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Miller, G. [1956], Human memory and the storage of information. *IRE Transactions on Information Theory*, **IT-2**(3), 128-137.

Norman, D. [1981], Categorization of action slips. *Psychological Review*, **88**(1), 1-15.