EUROPEAN JOURNAL OF INFORMATION SYSTEMS
https://doi.org/10.1080/0960085X.2018.1475876

Feedback, task performance, and interface preferences

Steven Bellman [a] and Kyle B. Murray [b]

[a]Ehrenberg-Bass Institute, University of South Australia, Adelaide, Australia; [b]Alberta School of Business, University of Alberta, Edmonton, Canada

ABSTRACT

Interface preferences are influenced by switching costs, including cognitive switching costs of thinking and task performance. In this research, we use feedback intervention theory to predict that feedback can have psychological effects that override the lock-in effect of cognitive switching costs on interface preference. To demonstrate this effect, we use normative feedback, which compares a user's task performance to the performance of others. This focuses attention on the user's self-concept and away from task performance. We use a hedonic information systems (IS) interface – an online game – as feedback valence should have a stronger effect on preferences for hedonic IS as opposed to utilitarian IS. Hedonic IS are preferred for their enjoyment value, as well as their productivity, and positive feedback should be more enjoyable than negative feedback. Results from an online experiment that manipulated the presence of feedback and feedback valence, for a sample of 482 users, support our hypotheses. The theoretical and managerial implications of these results are discussed.

ARTICLE HISTORY
Received 21 September 2016
Revised 16 April 2018
Accepted 18 April 2018

KEYWORDS
Interface; preference; human capital model; power law of practice; cognitive lock-in; feedback intervention theory

1. Introduction

People choose interfaces to access information systems (IS) for a variety of reasons (Benyon, Turner, & Turner, 2005). Sometimes, they have no choice; the interface is mandated by company policy (Venkatesh, Morris, Davis, & Davis, 2003). Other times, the choice of interface is limited by the cost of switching to a competitor (Klemperer, 1987), for example, the cost of breaking a contract, or buying new hardware, software, or training. Ultimately, switching costs can lock-in users to an incumbent interface (Carter, Wright, Thatcher, & Klein, 2014; Shapiro & Varian, 1999; Sun et al., 2017).

In this article, we examine a particular type of switching cost – that is, the amount of thinking and time it would take to use a competitor interface, compared to doing the same task with the incumbent interface (Labrecque, Wood, Neal, & Harrington, 2017; Wernerfelt, 1985). Task performance is a measure of the productivity of a task, holding constant time and thinking costs. Even small differences in task performance, saving just a few seconds, can lead to "cognitive lock-in" to an incumbent interface rather than a competitor (Murray & Häubl, 2007). For example, one field study demonstrated that websites that are faster to use attract more visitors and buyers (Johnson, Bellman, & Lohse, 2003).

Managers introducing a new interface, and companies with competing interfaces, have an interest in eliminating cognitive lock-in to an incumbent interface. Prior research has shown that interface preference is driven by perceived ease-of-use, rather than actual ease-of-use measured by objective task performance (Murray & Häubl, 2007, 2011). Perceived ease-of-use is affected by psychological responses. For example, making errors when first using an interface can lead to a lasting perception that the interface is difficult to use, even after actual ease-of-use improves with practice (Murray & Häubl, 2007). Denying users the freedom to choose their interface can lead to reactance, a negative psychological response (Murray & Häubl, 2011). The enforced option is perceived as difficult to use, even after dedicated practice has made its use highly efficient. The present research introduces and tests a third way of generating psychological responses that reduce cognitive lock-in. We show that interface preferences can be manipulated by providing computerised feedback.

There is a long history of using feedback to improve human–computer interaction and aid decision-making (Te'eni, 1992). Feedback has many dimensions, including valence (whether it is positive or negative) and type (eg, whether it is specific or normative) (Kluger & DeNisi, 1996). *Specific feedback* informs users about how to avoid errors and improve their performance on a task. Examples of specific feedback are error messages, such as: "The articles you requested are currently not available" (Murray & Häubl, 2011, p. 962). *Normative feedback* makes users feel good or bad about themselves by comparing their performance to the average user. Examples of normative feedback are telling someone they are doing a "Great job!" or "You seem to have an uncommon ability

CONTACT Steven Bellman steve.bellman@unisa.edu.au
Present affiliation for Kyle B. Murray is Alberta School of Business, University of Alberta, Edmonton, Canada and Monash Business School, Melbourne, Australia

to structure data logically" (Fogg & Nass, 1997, p. 555). The key difference between normative and specific feedback is that specific feedback does not compare a person's performance with the performance of others. For this reason, specific feedback is less likely to distract users from accurately perceiving their task performance when using an IS interface.

Normative feedback is rare relative to specific feedback, which is provided by nearly every interface (eg, the HTTP standard "404 Not Found" message). Increasingly, however, normative feedback is being pushed to users through gamification (Deterding, Dixon, Khaled, & Nacke, 2011), computerised training (Goodman, Wood, & Chen, 2011), online applications, social media, smartwatches (Chuah et al., 2016), fitness monitors, and other emerging communication technologies. Normative feedback is easy to program into an interface if it uses a constant message. In one experiment, when users were given constant messages in the form of praise (positive normative feedback), these users felt better about themselves and their performance, even when told that the feedback had nothing to do with their actual performance (Fogg & Nass, 1997). In this article, when we refer to feedback we mean normative feedback. Compared with specific feedback, we expected normative feedback to be more likely to generate a psychological response that makes users ignore an interface's actual ease-of-use.

Previous demonstrations of the disruptive effects of psychological responses on cognitive lock-in have used a utilitarian interface – for example, a news website (Murray & Häubl, 2007, 2011). Psychological responses to feedback valence should have a greater influence on interface preference for hedonic as opposed to utilitarian interfaces. This is because providing positive feedback makes a task seem more pleasant (Kluger & DeNisi, 1996). Utilitarian interfaces are productivity-oriented, but hedonic interfaces are pleasure-oriented. Intentions to use hedonic IS are influenced by perceived enjoyment, as well as by perceptions of ease-of-use and usefulness (Van Der Heijden, 2004). Perceived enjoyment also influences continuance intentions for hedonic IS (Merikivi, Tuunainen, & Nguyen, 2017). If the perceived enjoyment of using a hedonic IS interface is elevated by offering positive feedback, then intention to use the hedonic interface should also increase. No prior research has tested this chain of effects, so the aim of the present research is to demonstrate that positive feedback increases preference for hedonic IS interfaces. Hedonic IS include many consumer IS ranging from movie websites (Van Der Heijden, 2004) to mobile phones (Venkatesh, Thong, & Xu, 2012) to online games (Merikivi et al., 2017; Murray & Bellman, 2011).

We use an online game in our experiment as a representative hedonic IS interface, for several reasons (Murray & Bellman, 2011). First, like other hedonic experiences, online games are intrinsically rewarding (Hirschman & Holbrook, 1982) – that is, they offer rewards in and of themselves (eg, enjoyment) – as opposed to utilitarian tasks, which are carried out for extrinsic rewards (eg, money). Second, playing online games can be very exciting, compared to utilitarian experiences like reading news, allowing experimenters to observe highly emotional responses (Van Reekum et al., 2004). Third, playing games requires the learning of skills that improve with practice (Holbrook, Chestnut, Oliva, & Greenleaf, 1984; Schilling, Vidal, Ployhart, & Marangoni, 2003). Fourth, online games are played repeatedly, so asking people to play a game several times, and choose which one they would like to play next, is not unusual, and representative of hedonic consumption (Murray & Bellman, 2011). Fifth, a non-student sample is more likely to persist with a long, repetitive task if it is enjoyable. Playing simple, hedonically motivated games avoids fatigue effects while exhibiting the same learning effects as other hedonic activities in a conveniently short period of time (Luo, Ratchford, & Yang, 2013). Finally, there is evidence that hedonic tasks are best learned using massed repetition (such as in our experiment) rather than spaced learning, which is harder to replicate in the lab (Lakshmanan, Lindsey, & Krishnan, 2010).

But first, we need to show that hedonic interface preference is influenced by task performance. It is easy to see that task performance influences preference for productivity-oriented utilitarian interfaces. It is less obvious that task performance influences hedonic interface preference, since the dominant design objective for hedonic IS is to "encourage prolonged rather than productive use" (Van Der Heijden, 2004, p. 695). But prior research has shown that users do prefer hedonic interfaces with more efficient task performance, as they allow users to extract more enjoyment from their usage time (Murray & Bellman, 2011). This is consistent with the human capital model, which explains that a person's "taste" in activities, including hedonic activities like playing online games, is influenced by investments in human capital (consumer knowledge) related to those activities (Becker, 1993; Murray & Häubl, 2003; Ratchford, 2001; Stigler & Becker, 1977). Ratchford (2001) defines human capital as "knowledge or expertise that might make a consumer more productive at producing consumption activities" (p. 400). This distinction between "knowledge" and "expertise" corresponds to Alba and Hutchinson's (1987) distinction between familiarity (knowledge of) and expertise (knowledge how): "Familiarity is defined as *the number of product-related experiences that have been accumulated by the consumer*. Expertise is defined as *the ability to perform product-related tasks successfully*. … In general, increased product familiarity results in increased consumer expertise" (p. 411).

In our research, familiarity corresponds with interface usage and expertise with task performance. We define *interface usage* as the number of hands-on experiences with the interface accumulated by the user. *Task performance*, as above, is the ability to perform interface-related tasks productively. *Interface preference* is likelihood that an interface will be chosen to carry out a task. According to the human capital model, greater familiarity with an activity (eg, hedonic interface usage) improves expertise at performing the activity productively (task performance), which increases the likelihood of engaging in that activity in future (interface preference). For hedonic tasks, performance measures the ability to extract enjoyment from the task, per unit of time (Murray & Bellman, 2011). For example, going to a baseball game might be unpleasant for a novice spectator but enjoyable and exciting for an expert (Luo et al., 2013). Over time, a person's repertoire of activities tends to converge on those activities they have acquired the greatest expertise at performing, and therefore represent the most productive use of their leisure time (Luo et al., 2013; Ratchford, 2001; Stigler & Becker, 1977).

Building on previous research (Murray & Bellman, 2011), we use an online game as our example hedonic interface for the reasons listed above and because it is easy to measure productivity within an online game. When online game usage improves task performance, users can extract more enjoyment per unit of time by scoring more points per second. When given a choice, users prefer the game they have used more often, because it is more productive, measured by points per second (Murray & Bellman, 2011). In the current research, we replicate this effect and show that task performance cognitively locks-in users to a hedonic interface. Moreover, we demonstrate that feedback turns off this cognitive lock-in effect, by distracting users from considering prior usage when choosing between hedonic interfaces.

This research makes two important contributions to the literature on managing IS interface preferences. First, it tests whether computerised feedback can influence users' interface preferences (Fogg & Nass, 1997; Stoll, Edwards, & Edwards, 2016). Second, it draws on two research streams to explain this effect: the human capital model of the influence of task performance on preference (Becker, 1993; Murray & Häubl, 2003; Stigler & Becker, 1977) and feedback intervention theory, which explains the distracting effects of feedback on learning (Kluger & DeNisi, 1996). In the next section, we draw on relevant theory from IS and psychology research to propose three hypotheses, which we test using a controlled online experiment. We conclude by considering the implications of this research for theory and practice, as well as its limitations – including, whether our results apply to other hedonic IS interfaces, or to utilitarian IS interfaces, and whether other types of feedback can similarly loosen or perhaps tighten the cognitive lock.

2. Theory and hypotheses

2.1. *The mediating role of task performance*

Figure 1 shows the research framework for this study. Drawing on the human capital model, we expect that task performance will mediate the relationship between interface usage and interface preference (Murray & Häubl, 2003). Interface usage improves task performance via a practice effect. This improvement in task performance, relative to other interfaces, increases interface preference, so that the most-used interface is chosen more often for future usage. We use feedback intervention theory (Kluger & DeNisi, 1996) to predict that providing feedback will moderate the mediating role of task performance between interface usage and interface preference. *Feedback presence* will distract users from developing and improving task performance, reducing the influence of task performance on interface preference. Instead, interface usage will have a direct effect on interface preference. We further predict that *feedback valence* will have a moderating effect on this direct relationship between interface usage and preference. Since enjoyment is critical for intentions to use hedonic IS (Van Der Heijden, 2004), positive feedback will increase preference for the hedonic interface giving

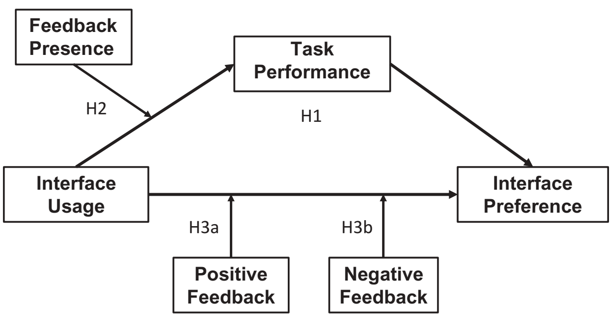

Figure 1. Research framework.

the feedback, reducing the effect of interface usage. Negative feedback will lead to rejection of the interface giving the feedback, again overriding the effect of usage on preference.

The effect of past behaviour (interface usage) on future behaviour (interface preference) could be due to simple stimulus–response learning (Ouellette & Wood, 1998; Pavlov, 1928; Venkatesh et al., 2012; Wood & Rünger, 2016). The human capital model, however, requires an additional mediating variable – task performance – between interface usage and interface preference. Task performance may or may not be accompanied by an increase in cognitive load (Camp, Paas, Rikers, & Van Merrienboerd, 2001). Improvements in task performance, measured by speed or accuracy (Camp et al., 2001), may be due to increased cognitive effort or, alternatively, reductions in cognitive effort that come with the automation of practised skills (Murray & Häubl, 2002, 2007; Shiffrin & Schneider, 1977). According to the human capital model, users prefer to maximise the productivity of their time by choosing the most efficient alternative (Becker, 1993; Murray & Häubl, 2003; Ratchford, 2001; Stigler & Becker, 1977). The cost of switching from an efficient incumbent to a less efficient competitor can be measured by time and thinking costs. If the incumbent's shorter task completion time requires a higher thinking cost, a competitor will still be attractive and the user will not be cognitively locked-in. An important aspect of the human capital model is that the mediator variable – task performance (eg, faster task completion time) – also reflects lower cognitive effort (Murray & Häubl, 2007, 2011).

Prior studies provide empirical evidence for the relationships between usage, task performance, and preference predicted by the human capital model (Murray & Häubl, 2003), even though they do not explicitly refer to the "human capital" model. For example, in a field study where the task performed was visiting an online store, customers were more likely to buy from a store with the shortest task completion time (Johnson et al., 2003). Furthermore, usage improved task performance in line with the power law of practice (Newell & Rosenbloom, 1981), which is used in GOMS (Goals, Operators, Methods, Selection Rules) models to accurately predict usage times for interfaces such as mobile phones (Jastrzembski & Charness, 2007). In two later studies, interface usage was experimentally manipulated, and task completion time measured prior to observing interface preference, to provide causal evidence for the mediating role of task performance (Murray & Häubl, 2007, 2011). In those studies, interface usage was varied by asking different groups of participants to complete different numbers of practice trials with their first interface (the incumbent). The participants then completed the same task with a second interface (the competitor), for one practice trial. The results of both experimental studies showed that the influence of interface usage on interface preference was fully mediated by task performance, measured by relative task completion time (*RTCT*) – that is, if the task took less time to complete using the competitor, relative to using the incumbent, users were more likely to prefer the competitor (Murray & Häubl, 2007, 2011).

The mediating role of task performance between interface usage and interface preference has also been demonstrated for hedonic interfaces. Specifically, Murray and Bellman (2011) showed that task performance mediates the effect of interface usage on a multi-item continuous measure of interface preference. In the hedonic domain, however, task performance is not measured by shorter task completion times, but by greater productivity per unit of time. Murray and Bellman (2011) measured productivity by the amount of time taken to score points. With practice, players took less time to score each point, and consequently scored more points per second. We expect to replicate this result and extend it by using consequential choice as our measure of interface preference (rather than self-reported attitude towards the interface). Therefore, we propose:

H1. Task performance will mediate the relationship between interface usage and interface preference.

2.2. *Effects of feedback on task performance*

As discussed in the introduction, our experiment tested normative feedback, which compares the user to others, rather than specific feedback, because normative feedback is more likely to generate an emotional response. Offering positive normative feedback ("you are doing better than the average person") makes users feel better about themselves and their performance (Fogg & Nass, 1997). The designers of IS interfaces may believe that these positive feelings about performance will translate into positive improvements in actual performance by increasing the motivation to learn (Förster, Grant, Idson, & Higgins, 2001). Increasing the motivation to learn should increase task performance and, as a result, tighten the cognitive lock-in effect. Learning motivation and task performance might also be increased by offering occasional negative feedback, that is, punishments as well as rewards (Finkelstein & Fishbach, 2012; Finkelstein, Fishbach, & Tu, 2017; Zingoni & Byron, 2017). Even neutral normative feedback ("you are performing as well as the average person"), which could be interpreted positively or negatively, might stimulate an increase in learning effort. Prior research, however, indicates that feedback interventions are not always successful. Positive feedback can have positive or negative effects (Kluger & DeNisi, 1996). In fact, one third of feedback interventions have a negative effect on the quality of the actual learning (Kluger & DeNisi, 1996).

These negative effects of feedback can be explained by feedback intervention theory's hierarchy-of-attention model (Kluger & DeNisi, 1996). People can focus their limited attention on just one level of this attention hierarchy and by default their attention is focused optimally at its midpoint. At the bottom of the hierarchy, attention is focused too much on details of the task. At the top of the hierarchy, attention is focused too much on what task performance means for the person's self-concept. Feedback interventions have a negative effect when they distract attention up or down the hierarchy, away from its optimal midpoint. Specific feedback has a negative effect when it focuses attention too far down the hierarchy, on task details (Goodman et al., 2011; Goodman, Wood, & Hendrickx, 2004; Rogers, 2017). Normative feedback has a negative effect when it focuses attention too far up the hierarchy, on reconciling the individual's self-efficacy beliefs with feedback comparing the individual to others (Baadte & Kurenbach, 2017; Bandura, 1986; Kluger & DeNisi, 1996; Maier, Wolf, & Randler, 2016; Vancouver & Tischner, 2004).

Using feedback intervention theory (Kluger & DeNisi, 1996), we expect that feedback will hinder the improvement of task performance because the normative feedback we use should distract attention away from the optimal midpoint of the attentional hierarchy, where learning is maximised. Normative feedback can range in valence from positive praise ("you are performing better than average"), to neutral ("your performance is average"), to negative ("your performance is worse than average"). Positive normative feedback is likely to distract users from learning if it conflicts with their own (realistic) conception of their performance. Praise is also unlikely to increase learning effort, since praise suggests that performance is already above average. For some users, negative feedback will match their self-concept, but for others its mismatch might distract from learning or prompt increased effort. The difference between neutral (average) feedback and a person's positive or negative self-concept will, on average, be smaller than contrasting negative or positive feedback. For this reason, neutral feedback may have a smaller distracting effect on task performance than positive or negative normative feedback.

Hypothesis 1 predicts that interface usage improves task performance, so that task performance has a mediating role between interface usage and interface preference (Ratchford, 2001). Hypothesis 2 predicts that feedback will moderate the mediating role of task performance. A positive moderation effect would increase the rate at which interface usage improves task performance and interface preference. A negative moderation effect would reduce these rates of improvement, so that task performance and interface preference do not improve, or even decline with interface usage. If interface usage no longer improves task performance or interface preference, task performance would no longer mediate between interface usage and interface preference.

Furthermore, if task performance no longer explains interface preference, then interface preference could be directly related to the presence of feedback. For example, task performance could be uniformly low whenever normative feedback is present. We test for this potential direct effect of feedback on task performance. However, it is more likely that interface usage will always have an improving effect on task performance, but the presence of normative feedback will have a negative counter effect, reducing the rate of this improvement. When interface usage has a reduced effect on task performance, task performance will have a weaker mediating role between interface usage and interface preference. Therefore:

H2. Feedback presence will negatively moderate the mediating role of task performance between interface usage and interface preference.

2.3. *Effects of feedback valence on interface preference*

Up to this point, we have considered the mediating role of task performance between interface usage and interface preference, and the moderating effect of the presence of feedback on task performance's mediating role. Now we consider the moderating effects of feedback valence – that is, whether the feedback is positive, negative, or neutral – on the direct relationship between interface usage and interface preference, as shown in Figure 1. We expect that emotional responses to feedback valence will be associated with the interface providing the feedback (Fogg & Nass, 1997; Förster et al., 2001; Stoll et al., 2016). These emotional responses to feedback valence will override the influence of interface usage on interface preference (Murray & Häubl, 2011).

If the incumbent offers positive feedback (eg, praise), that positive feedback will elicit enjoyment (Kluger & DeNisi, 1996). Enjoyment increases intentions to use hedonic IS independently of their perceived ease-of-use or usefulness (Van Der Heijden, 2004). In our experiment, only one interface is associated with feedback valence, the incumbent. The competitor does not offer any form of feedback. This absence of positive feedback may make the competitor seem relatively unpleasant compared with the incumbent. This contrast in emotional responses should make the incumbent more enjoyable and therefore preferred over the competitor, even if competitor has similar task performance after repeated usage trials (Murray & Häubl, 2011).

Hypothesis 2 proposed that feedback presence will negatively moderate the mediating role of task performance between interface usage and interface preference. Now, hypothesis 3 is proposing a similar negative moderating effect of feedback valence. Hypothesis 3 proposes that positive feedback will flatten the relationship between interface usage and preference, by distracting users from considering task performance when choosing. A highly positive emotional response may eliminate the effect of usage altogether, so that feedback valence has a direct effect on preference. The interface giving the positive feedback (in this case, the incumbent) will be preferred despite increasing usage of another interface. But we expect that, as in the case of hypothesis 2, interface usage will still increase interface preference via a simple stimulus–response learning effect, but the slope of this effect will be less steep when one of the interfaces gives positive feedback.

H3a. Positive feedback will negatively moderate the relationship between interface usage and interface preference by favouring the interface giving the feedback.

Negative feedback can induce unpleasant feelings and learned helplessness (Kluger & DeNisi, 1996). We expect that an interface that gives negative feedback will be associated with unpleasant emotion. Another interface that offers no feedback will seem more pleasant in contrast. In a prior experiment, negative emotion was induced by denying freedom of choice between utilitarian interfaces (Murray & Häubl, 2011). This unpleasant negative emotion was associated with the forced-choice incumbent, and so participants switched to the comparatively more pleasant free-choice competitor as soon as they could, despite the incumbent's advantage in task performance (Murray & Häubl, 2011).

Emotional responses are likely to be even more important for hedonic interfaces, because intentions to use hedonic IS are influenced by enjoyment (Van Der Heijden, 2004). We expect that participants will tend to choose the relatively more pleasant interface that does not give negative feedback, ignoring differences in prior usage. There may be evidence for a direct effect of negative feedback on preference, if preference does not change with usage. But again, it is more likely that negative feedback will reduce the slope of the relationship between interface usage and interface preference:

H3b. Negative feedback will negatively moderate the relationship between interface usage and interface preference by favouring the interface that does not give the feedback.

Finally, we expect that receiving neutral feedback is likely to elicit, on average, a neutral emotional response. For some, neutral feedback ("you're average") will be pleasant, while for others it will be unpleasant. But the magnitude of these emotional responses will be relatively small because "average" feedback will not create much of a contrast with either a positive or a negative self-concept. The absence of neutral feedback will similarly elicit a small contrasting emotional response, since for some this absence will be mildly pleasant and for others it will be mildly unpleasant. The presence of neutral feedback should have only a mild distracting effect, and interface usage and task performance should still influence preference. Therefore, the relationship between interface usage and interface preference is likely to be similar in the neutral and no feedback conditions. Although we test for a potential moderating effect of neutral feedback on the relationship between interface usage and interface preference, our theory does not predict a significant result.

The next section describes our experiment, which uses a large non-student sample to test these hypotheses.

3. Method

3.1. *Experimental design*

The experiment closely replicated the procedures used in previous tests of the human capital model (Murray & Häubl, 2003) and related studies (Murray & Bellman, 2011; Murray & Häubl, 2002). Participants practised using one hedonic interface (the incumbent), then practised using a second hedonic interface (the competitor), and finally were asked to choose which interface they would prefer to use again (ie, the incumbent or the competitor). Both the incumbent and the competitor were simple online video games, identical apart from the skills needed to play them. Thus, only differences in usage could explain why one was preferred (Murray & Häubl, 2007, 2011).

The experiment employed a 4 (feedback: no feedback, positive, neutral, or negative feedback) × 2 (interface usage: 1 competitor trial vs. 3 competitor trials) between-participants factorial design. Every participant had the same number of practice trials with the incumbent (9 trials), but half of the participants (depending on random assignment) played only one practice trial with the competitor game (the low interface usage condition), while the other half had three practice trials (the high interface usage condition). Previous research shows that three trials of a competitor interface can sufficiently develop task performance (reduce task completion time) such that the choice-share advantage of an incumbent interface used for nine trials is eliminated (Murray & Häubl, 2007).

Feedback presence and valence were also manipulated by random assignment – feedback was not based on actual performance. This may seem like an unnatural practice, but it has been used successfully in prior feedback intervention research (Fogg & Nass, 1997; Vancouver & Tischner, 2004). Also, to argue that effective feedback requires expensive and complex systems that analyse actual performance, research first needs to demonstrate that simple constant feedback systems do not work. More importantly, manipulating feedback gave us experimental control over temporal precedence, a necessary condition for demonstrating a causal effect, in addition to mere correlation (Cook & Campbell, 1979). If we had provided feedback about actual performance, feedback valence would have been confounded with prior experience and therefore higher starting levels of skill. To further control for prior experience with online games, we measured this variable and used it as a covariate. We did not need to use any other control variables because our use of random assignment successfully controlled for otherwise plausible alternative explanations (eg, Internet speed, computer specification, player's environment, left- vs. right-handedness, preference for one- vs. two-handed controls, etc.) (Cook & Campbell, 1979).

3.2. *Participants*

Four hundred and eighty-two members of an Australian audience panel (age range 20–81, normally distributed, M = 44.68, SD = 13.14, 69% women) participated in the experiment. Each participant received a \$5 (AUD) department store gift card. The sample ranged widely in prior online video game experience (Table 1). Participants received the same reward no matter what their performance to encourage participation from users with little interest in playing online games. A *post hoc* analysis revealed that the sample size was sufficient to detect significant (p < .05) medium-sized effects with over 99% power (Faul, Erdfelder, Lang, & Buchner, 2007).

3.3. *The online game*

The racing car game was designed to be like online video games readily available over the Internet. The object of the game was to score points by shooting enemy cars while avoiding oncoming cars and incoming fire. Critical damage to the player's car ended the game. The car, viewed from above on-screen, could be moved left or right with the mouse.

The incumbent and competitor games differed in a seemingly trivial way: whether the firing trigger was the left button on the mouse (one-handed play) or the keyboard spacebar (two-handed play). One- versus two-handed play made only a marginal difference to game-playing time (20.81 s vs. 17.39 s, p = .07 [Wilcoxon tests used for all these comparisons]), but no difference to score (172.71 vs. 141.73, p = .11), which meant that both games were equally efficient (measured by score per second) prior to random assignment (5.81 vs. 5.99, p = .74). The results reported below are the same if game-interface is included as a factor. However, this interface difference between the games ensured that players acquired skill using the incumbent game that was not transferable to the competitor (Murray & Häubl, 2002, 2007). Playing one-handed first versus two-handed first was counterbalanced.

Table 1. Correlations among variables and descriptive statistics.

		Gender	Age	E	U[a]	RSPS	InterfacePreference[b]
Gender (female = 1, male = 0)		—					
Age (years)		−0.09*	—				
Prior experience (*E* [mean centred])		−0.22***	−0.29***	—			
Interface usage (*U* [1 trial = 0, 3 trials = 1])		−0.06	−0.002	0.03	—		
Task performance (*RSPS*)		0.02	−0.03	0.01	0.16***	—	
Interface preference (incumbent = 0, competitor = 1)		0.02	−0.08	0.06	0.11*	0.18***	—
TOTAL	Mean	0.69	44.68	−0.002	—	−0.62	0.51
N = 482	(SD)	(0.46)	(13.14)	(1.33)	—	(5.25)	(0.50)
1 Competitor trial	Mean	0.71	44.90	−0.03	—	-1.41^{x}	0.45^{x}
n = 228	(SD)	(0.45)	(13.66)	(1.37)	—	(5.39)	(0.50)
3 Competitor trials	Mean	0.66	44.48	0.02	—	0.08^{x}	0.56^{x}
n = 254	(SD)	(0.47)	(12.68)	(1.29)	—	(5.03)	(0.50)
No feedback	Mean	0.69	43.87	0.27^{xyz}	—	−0.67	0.51
n = 178	(SD)	(0.47)	(11.71)	(1.17)	—	(5.43)	(0.50)
Positive feedback	Mean	0.77	42.72	-0.01^{x}	—	-0.14^{y}	0.44
n = 107	(SD)	(0.43)	(13.36)	(1.47)	—	(4.79)	(0.50)
Neutral feedback	Mean	0.62	45.89	-0.11^{y}	—	-1.92^{yz}	0.54
n = 102	(SD)	(0.49)	(14.18)	(1.34)	—	(5.26)	(0.50)
Negative feedback	Mean	0.67	47.11	-0.38^{z}	—	0.32^{z}	0.55
n = 95	(SD)	(0.47)	(13.97)	(1.35)	—	(5.17)	(0.50)

Note. Pairwise correlations. Means in the same column with the same superscript letters are significantly different (p < .05).
SD = standard deviation; *RSPS* = relative score per second (*SPS* for competitor minus *SPS* for incumbent). Zero *RSPS* means that the competitor and the incumbent are equally efficient to use. Positive *RSPS* means that the competitor is more efficient to use than the incumbent.
[a]Coded 0 = 1 competitor trial, 1 = 3 competitor trials, Spearman correlations.
[b] Coded 0 = chose incumbent interface, 1 = chose competitor interface, Spearman correlations.
*p < .05, **p < .01, ***p < .001.

3.4. Procedure

Participants were invited by email to complete an online survey. The survey randomly assigned participants to one of the eight (4 feedback × 2 interface usage) groups. The instructions provided prior to the first incumbent-game trial forewarned all the participants, except those in the no-feedback condition, that they would receive feedback about their performance. If feedback was given, a normative message appeared two-thirds of the way into players' experience with the incumbent game, after the sixth incumbent trial. The positive feedback message was: "You are performing BETTER than other people who have played this game". The other two conditions were created by replacing the words "BETTER than" with "THE SAME as" (neutral) or "WORSE than" (negative). After receiving this feedback, participants played their three remaining incumbent trials and then, depending on group assignment, played either one or three competitor game trials.

3.5. Measures

The game software unobtrusively recorded each usage trial's duration and score achieved. Like previous experiments, we measured the relative task performance of using the competitor interface compared to the incumbent (Murray & Häubl, 2007, 2011). This measure quantifies the relative advantage of the competitor versus the incumbent. If the competitor is more productive than the incumbent, task performance is positive. On the other hand, if the competitor is less productive, task performance is negative (ie, there is a switching cost associated with choosing the competitor).

Building on prior research, we used a measure of task performance that was appropriate for a hedonic task like playing a video game (Murray & Bellman, 2011). For utilitarian tasks, more efficient tasks are completed faster, so task performance can be measured by *RTCT*. *RTCT* is calculated by subtracting the last task completion time for the competitor from the last completion time for the incumbent (Murray & Häubl, 2007). If the competitor is faster to use than the incumbent, *RTCT* will be positive, indicating that the competitor is more likely to be preferred. For hedonic tasks, however, task performance is best captured by rewards per fixed unit of time, such as the number of fish caught per hour, or the number of points scored per second (Murray & Bellman, 2011).

We use the same type of video game that Murray and Bellman (2011) used and, therefore, we use their measure of task performance: relative score per second (*RSPS*). Like *RTCT, RSPS* measures the relative attractiveness of the competitor. *RSPS* was calculated by subtracting the last incumbent game's score per second from the last competitor game's score per second. If the competitor had a higher score per second than the incumbent, *RSPS* would be positive, indicating that the competitor was more likely to be preferred. Both measures of task performance, *RTCT* for utilitarian tasks and *RSPS* for hedonic tasks, have a positive correlation with interface preference when the competitor is more efficient than the incumbent. Task performance should have no effect on interface preference if the competitor and the incumbent are equally efficient (both *RTCT* and *RSPS* would be zero). In Murray and Bellman's (2011) study, *RSPS* mediated the effect of interface usage on interface preference, measured by self-reported attitude towards the competitor game.

Murray and Bellman (2011) considered total score as an alternative measure of task performance. But high scores do not necessarily indicate high productivity, as a less skilled player could score a high number of points very slowly. This was demonstrated in the two experiments reported by Murray and Bellman (2011), in which groups differed in usage (playing 1 game vs. 10 games) and preference (more learning increased preference), but did not differ in total points scored. When productivity was measured by points scored per second, however, the 10-game groups significantly improved their productivity, comparing their first and tenth games. Similarly, another potential measure of task performance, longer game times, could be generated by low-productivity, low-scoring performance. Again, in Murray and Bellman's (2011) experiments, there was no difference in game time across groups that differed in usage, preference, and task performance measured by score per second. Finally, it might seem that point scoring is a measure of instant gratification rather than task performance. But when the enjoyment of a hedonic task derives from satisfying the objective of a game, and that objective is to score points, then points per second provide a useful measure of the productivity of task performance.

After completing the final competitor trial, each participant answered survey items taken, for comparison purposes, from related experiments (Murray & Bellman, 2011; Murray & Häubl, 2007, 2011). First, participants revealed their subjective utility for the competitor game by completing a realistic behavioural choice measure of interface preference (Luce, 1959; McFadden, 1981, 1986; Morales, Amir, & Lee, 2017; Murray & Häubl, 2011). Specifically, they were shown images from the two games they had played (identified by their background colour, black or grey) and chose the game they would prefer to play again (incumbent = 0, competitor = 1). It could be argued that this binary choice measure was inappropriate, especially for games: users would naturally vary their choices, so it would have been better to measure preference by highest percentage over repeated choices. However, research shows that choice experiment results are consistent whether the dependent variable is one choice or repeated choices (Brouwer, Dekker, Rolfe, & Windle, 2010).

Next, participants completed a multiple-item self-reported measure of prior experience (Murray & Bellman, 2011) (see Appendix 1). The validity of this measure was assessed using SmartPLS. Composite reliability (CR) was 0.86 and average variance extracted (AVE) was 0.68, both exceeding the conventional CR > 0.70 and AVE > 0.50 criteria (Bagozzi & Yi, 1988). In addition, the square root of AVE (0.82) exceeded all the correlations between prior experience and other variables measured in this experiment (Table 1), indicating discriminant validity (Fornell & Larcker, 1981). To rule out alternative explanations for our results, participants answered an open-ended question intended to elicit hypothesis guessing. The use of different sources for our predictor and criterion variables (manipulation, log files, discrete choice, and self-report scales) minimises the possibility that our research findings are affected by common method bias (Podsakoff, MacKenzie, & Podsakoff, 2012).

3.6. *Analysis*

We used logistic regression to estimate our main results, as our dependent variable was binary (interface preference = 0 [incumbent] or 1 [competitor]). Appendix 2 shows that the same pattern of significant estimates is found when experiment-wise error is controlled using SmartPLS, although SmartPLS should not be used for models with binary endogenous variables (Hair, Sarstedt, Ringle, & Mena, 2012). The highest variance inflation factor (VIF) was associated with an interaction term, interface usage × negative feedback (VIF = 3.2), and was below the maximum recommended level (VIF = 10) (Myers, 1986). Tests confirmed that our results were not influenced by the age or gender composition of our sample. As Table 1 shows, age and gender were related to prior experience, and we controlled for prior experience in our analyses.

4. Results

4.1. *Manipulation checks*

First, we confirmed that interface usage increased task performance in line with the predictions of the power law of practice for skilled tasks (Murray & Häubl, 2007). Practice reduces utilitarian task completion time as users learn to spend less time unproductively. For hedonic tasks, task completion time may not reduce with practice, but unproductive time does go down (Murray & Bellman, 2011). We predicted average time to score a point, and based on that model, unproductive (residual) time as a function of usage (Murray & Bellman, 2011). The decrease in unproductive time, with practice, was better modelled by a power function (R^2 = .148, vs. .081 [linear] and .080 [exponential]). The power law coefficient for unproductive time was significantly negative ($\alpha = -.08$, $t(253) = -3.87$, $p < .001$ [1-tailed t-test]), indicating that players' performance became more efficient with practice (Newell & Rosenbloom, 1981). Evidence for automatisation and therefore increasingly lower cognitive load comes from decreasing variance in unproductive time, across people (Logan, 1988). A power function was the best fit to the standard deviations of unproductive time (R^2 = .349, vs. .162 [linear] and .156 [exponential]).

Finally, the increase in task performance (measured by score per second) with practice was also better modelled by a power function (R^2 = .798, vs. .641 [linear] and .575 [exponential]). We conducted further tests of practice using the competitor interface. These tests did not assume that learning progressed linearly in log-log space, as learning of new technology can be discontinuous (Lakshmanan & Krishnan, 2011). Participants who used the competitor three times achieved a higher score per second in their third game compared to their first (M_1 = 6.32 vs. M_3 = 7.58, $p < .001$ [Wilcoxon test]). Usage therefore increased the relative task performance (*RSPS*) of the competitor compared to the incumbent (Table 1; $p < .001$). In addition, competitor usage increased interface preference, measured by the proportion preferring the competitor over the incumbent (Table 1; $\chi^2(1) = 6.00$, $p = .01$).

A second set of checks investigated whether in this study, as in previous studies (Murray & Häubl, 2007, 2011), participants did not attribute their interface preference to the amount of experience they had with the two interfaces, nor to the interface-specific skills they had acquired. We analysed the answers participants gave to the open-ended question designed to detect hypothesis guessing. Only 14% of participants in the no-feedback condition mentioned the firing–trigger differences between the two games as a potential explanation for the results of the study, and no participant guessed the hypotheses.

4.2. *Overall results*

Hypothesis 1 predicted that task performance will mediate the relationship between interface usage and interface preference. Because our measure of interface preference is a binary choice, we used a non-parametric bootstrapping test for mediation (Preacher & Hayes, 2008), rather than the Sobel test, which assumes a normally distributed dependent variable. The results showed that task performance (*RSPS*) significantly mediated the effect of interface usage on interface preference. The indirect effect's 95% confidence interval (.02 to .20) did not include zero. Our regression results show that in this study, task performance had only a partially mediating effect. The effect of interface usage on interface preference is still significant after controlling for the mediating effect of task performance (Table 2), because of the effects of feedback (as predicted by H2 and H3).

Hypothesis 2 predicted that feedback presence will negatively moderate the mediating role of task performance between interface usage and interface preference. As predicted, the presence of the moderator had a negative effect that reduced the size and significance of task performance's mediating effect (Hayes, 2015). When the moderator was absent (the no-feedback condition), the mediating effect of task performance was significant (95% confidence interval [CI] = .02 to .40). However, when the moderator (feedback) was present, the mediating effect of task performance was not significant (95% CI = − .002 to .19, combining all three feedback valence conditions). The mediating effect of task performance was not significant whether the feedback was positive (−.07 to .21), neutral (−.09 to .27), or negative (−.06 to .69).

Further evidence for hypothesis 2 comes from tests of whether the moderating effect of feedback changed the input path to the mediator, as illustrated in Figure 1, or the output path. An input path effect is a Model 2 moderated mediation effect, while an output path effect is a Model 3 effect (Preacher, Rucker, & Hayes, 2007). A Model 2 effect is distinguished from a Model 3 effect by the significance of the interaction between the independent variable and the moderator, and the insignificance of the interaction between the moderator and the mediator (Preacher et al., 2007). In a separate test of the interface preference regression model, the interactions between task performance (the mediator) and the three levels of the moderator (positive, neutral, and negative feedback) were insignificant (all *ps* > .158). On the other hand, Table 2 shows a significant interaction between the independent variable (interface usage) and one level of the moderator (positive feedback), which is our evidence for hypothesis 3 below. In the no-feedback condition, interface usage had a significant effect on task performance (p = .017). However, when feedback was present, the effect of interface usage on task performance was not significant (all *ps* > .267 for positive, neutral, and neutral feedback). This is consistent with our theory, which anticipated that because three competitor usage trials no longer improved task performance significantly when feedback was provided, task performance no longer mediated the relationship between interface usage and interface preference.

Table 2. Regression results.

Independent variables	Dependent variable: Task performance[a]	Dependent variable: Interface preference[b]
Intercept	−1.67**	−0.40
Interface usage[c] (*U*)	1.88*	0.76*
Prior experience (*E* [mean centred])	0.06	0.12
Positive feedback (*POSITIVE*)	1.09	0.32
Neutral feedback (*NEUTRAL*)	−0.91	0.28
Negative feedback (*NEGATIVE*)	1.10	0.56
U × *POSITIVE*	−0.97	−1.17*
U × *NEUTRAL*	−0.61	−0.05
U × *NEGATIVE*	−0.29	−0.71
Task performance		0.06**

Note. Logistic regression used for binary dependent variable, interface preference.
[a]Measured by relative task completion time (score per second [*SPS*] for competitor minus *SPS* for incumbent). Zero *RSPS* means that the competitor and the incumbent IS are equally efficient to use. Positive *RSPS* means that the competitor is more efficient to use than the incumbent.
[b]Coded 0 = chose incumbent interface, 1 = chose competitor interface.
[c]Coded 0 = 1 competitor trial, 1 = 3 competitor trials.
*p < .05, **p < .01, ***p < .001

Hypothesis 3a predicted that positive feedback will negatively moderate the relationship between interface usage and interface preference by favouring the interface giving the feedback. This hypothesis was confirmed by the significant negative interaction between positive feedback and interface usage in the regression model predicting interface preference (Table 2). Follow-up cross-tabulation tests (Figure 2) showed that, as predicted by hypothesis 1, interface usage significantly increased interface preference for the competitor in the no-feedback condition, from 40% to 60% ($\chi^2(1)$ = 7.12, p = .008). But as predicted by hypothesis 3a, positive feedback given by the incumbent had a negative, slope flattening effect on this interface usage effect. The relatively less pleasant competitor was less preferred than the incumbent, whether it was used once (47%) or three times (40%). This meant that the effect of usage on preference was no longer significant when positive feedback was present ($\chi^2(1)$ = 2.33, p = .13).

Hypothesis 3b predicted that negative feedback will negatively moderate the relationship between interface usage and interface preference by favouring the interface that does not give the feedback. The negative interaction between usage and negative feedback predicted by hypothesis 3b was not significant in the interface preference model in Table 2. However, Figure 2 shows that negative feedback had the predicted negative moderating effect on the relationship between interface usage and interface preference. When the incumbent gave negative feedback, the relatively more pleasant competitor was preferred more than the incumbent, whether it was used once (53%) or three times (56%). As when positive

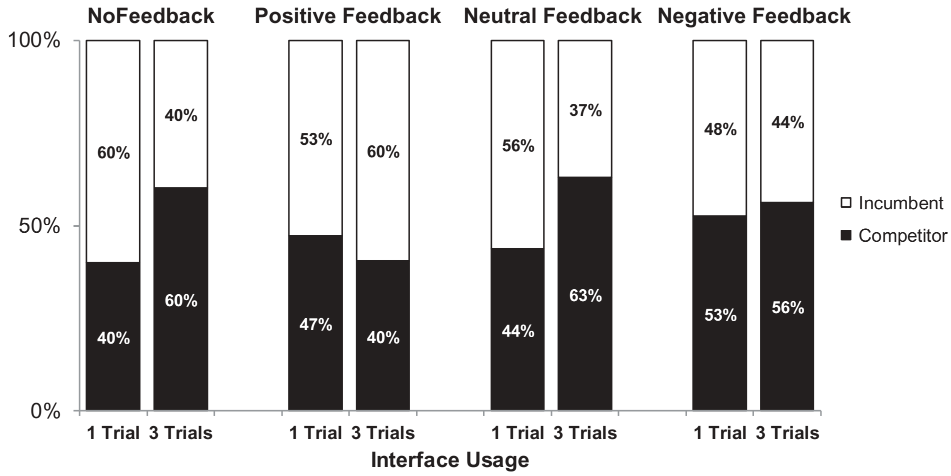

Figure 2. Interface preference (choice shares) by feedback valence and interface usage.

feedback was present, the effect of usage on preference was not significant when negative feedback was present ($\chi^2(1) = .09$, $p = .36$). Although we did not expect that neutral feedback would have a negative moderating effect on the relationship between usage and preference, we did test for this possibility. Figure 2 shows that when neutral feedback was present, interface usage increased preference for the competitor from 44% to 63% ($\chi^2(1) = 4.18$, $p = .04$). This result is similar to the relationship between usage and preference in the no-feedback condition.

Finally, we tested two alternative explanations for our results: (1) feedback has a direct negative effect on task performance and (2) feedback has a direct effect on interface preference. First, we looked at the direct effects of feedback on task performance in Table 2. None of these was significant. Furthermore, in Table 1, task performance was not significantly different when feedback was present (whether positive, neutral, or negative), compared to the no-feedback condition. Finally, those participants who used the competitor three times significantly improved their task performance (score per second) whether normative feedback was present ($M_1 = 6.55$ vs. $M_3 = 7.74$, $p = .001$) or not present ($M_1 = 5.91$ vs. $M_3 = 7.30$, $p = .013$). These results rule out a direct effect of normative feedback on task performance.

Further evidence for the moderating effect of normative feedback comes from the differences between the three feedback valence conditions. Interface usage significantly improved task performance in the neutral feedback condition ($M_1 = 6.03$ vs. $M_3 = 7.35$, $p = .040$), but not in the positive feedback condition ($M_1 = 7.11$ vs. $M_3 = 8.52$, $p = .073$) or the negative feedback condition ($M_1 = 6.53$ vs. $M_3 = 7.38$, $p = .171$). These differences in the effect of usage on performance produced differences between valence conditions in relative task performance (Table 1). But these main effects of valence were not significant after controlling for usage in the regression results (Table 2). Second, positive and negative feedback had what amounted to direct effects on interface preference, as preference was flat across both levels of interface use (Figure 2). However, there were no significant direct effects of feedback valence on interface preference in the regression results, after controlling for usage and the negative interaction between positive feedback and usage (Table 2).

5. Discussion

5.1. *Theoretical contributions*

This research shows that feedback can be used to influence IS interface preference. It extends previous research showing how usage improves task performance, which explains preferences for IS interfaces (Johnson et al., 2003; Murray & Häubl, 2002, 2007, 2011). According to the human capital model, users prefer activities that maximise the productivity of their time (Becker, 1993; Murray & Häubl, 2003; Ratchford, 2001; Stigler & Becker, 1977). Prior research has shown that psychological responses, such as being denied freedom of choice, can override this preference for more productive interfaces (Murray & Häubl, 2011). The results of the present research show that feedback can provoke similar psychological responses that distract users from considering task performance when choosing between IS interfaces.

In our experiment, we maximised the chances of finding a distracting effect of feedback by using positive and negative normative feedback, and a hedonic IS interface, represented by an online game. Hedonic IS are chosen for their enjoyment value, as well as for their ease-of-use and productivity (Van Der Heijden, 2004; Venkatesh et al., 2012). Therefore, positive and negative feedback should directly affect the enjoyment of using hedonic IS. Normative feedback, which compares task performance with others' performance, can distract from learning by focusing too much attention on the user's self-concept (Baadte & Kurenbach, 2017; Kluger & DeNisi, 1996; Vancouver & Tischner, 2004). In our experiment, we found that normative feedback reduced the improving effect of interface usage on task performance, so that task performance no longer mediated the effect of usage on preference. Furthermore, people preferred the interface associated with positive feedback and rejected the interface associated with negative feedback, which weakened the effect of usage on preference.

To confirm this theoretical explanation, we tested two alternative explanations for our results: that feedback has direct effects on (1) task performance or (2) on interface preference. First, there was no evidence of a direct effect of feedback that turned off task performance improvement. When the incumbent offered feedback, competitor usage still improved task performance. Second, some of our results hint at a direct effect of feedback on interface preference. When positive or negative feedback was given by the incumbent, the competitor's preference was constant across both levels of usage. However, these direct effects were not significant after controlling for a significant moderating effect of positive feedback. Overall, the results of this study provide support for our theory – feedback moderates task performance's mediating role between interface usage and preference, rather than directly influencing task performance and interface preference.

5.2. *Limitations and future research*

Like all research, this study had limitations, which suggest directions for future research. The research framework for this study contributes to theory by identifying opportunities for future research from the intersection of two extensive research streams, the human capital model (Becker, 1993; Murray & Häubl, 2007; Stigler & Becker, 1977) and feedback intervention theory (Kluger & DeNisi, 1996). The hypothesised moderating effects of feedback on the relationships between interface usage, task performance, and interface preference were not fully supported by significant interaction effects, although tests for moderated mediation and moderation of choice probabilities were significant. Future research should attempt to extend our results by further elaborating on the current study's support for our hypotheses.

One explanation for not finding significant interaction effects is that feedback affected both interfaces in our experiment. Feedback was delivered during usage of the first interface, and so affected task performance and preference for the incumbent, as well as for the second interface, the competitor. Future research should test the effects of offering feedback only during usage of the competitor. Our experimental results, using an online video game and a non-student sample, have high internal and external validity. These results, however, would benefit from replication in field experiments and in the workplace. Similarly, other research could examine the extent to which our results generalise to different types of feedback (other than normative), to hedonic IS interfaces apart from online video games, and to preferences for utilitarian IS interfaces, such as online stores (Johnson et al., 2003) and news websites (Murray & Häubl, 2007).

Specific feedback, such as error messages, is more common than normative feedback in IS interfaces. Whether specific feedback has a similarly distracting effect on attention to task performance most likely depends on the detail in these specific feedback messages (Rogers, 2017). Specific feedback that is too detailed focuses attention at the base of the attention hierarchy, away from its optimal midpoint where learning is maximised (Kluger & DeNisi, 1996). Specific error messages can also have effects on perceived ease-of-use that persist despite improvements in objective ease-of-use measured by task performance (Murray & Häubl, 2007). On the other hand, specific feedback designed to improve performance should have a positive moderating effect, increasing the mediating effect of task performance (Goodman et al., 2011, 2004). Specific feedback can also have positive or negative valence, and the valence of specific feedback may have motivational effects that also improve task performance (Finkelstein & Fishbach, 2012; Förster et al., 2001).

We used an online video game to replicate previous research showing that usage and task performance influence preferences for hedonic interfaces (Murray & Bellman, 2011). For hedonic activities, task performance is not measured by shorter task completion times, but by productivity per unit of time. Games are useful examples of hedonic activities, as their productivity per unit of time can be easily measured by score per second. For other hedonic IS, such as movie websites (Van Der Heijden, 2004), productivity per unit of time may be more difficult to measure. Alternatively, if the task performance of a hedonic IS cannot be directly measured, it could be inferred by observing task preferences over time (Luo et al., 2013). The effect of feedback manipulations on these preferences could be explained by their effects on perceived ease-of-use and actual ease-of-use (task performance) (Murray & Häubl, 2007, 2011). Finally, although feedback effects on enjoyment may not influence utilitarian interface preference (Johnson et al., 2003; Murray & Häubl, 2007), normative feedback and overly specific feedback (Goodman et al., 2011, 2004) should still distract from learning, negatively moderating the mediating role of task performance between utilitarian interface usage and preference. We hope that this research provides a starting point for discussion and future research related to the effects of feedback on IS adoption.

5.3. *Practical implications*

Our results show how feedback can influence user preference for IS interfaces. These results are based on tests using hedonic IS interfaces, specifically online video games and, therefore, should be of interest to managers in the $100B video games industry (Takahashi, 2014). However, future research is likely to confirm that the distracting effects of normative feedback (Kluger & DeNisi, 1996), and feedback that is too specific (Goodman et al., 2011, 2004; Rogers, 2017), also apply to utilitarian IS interface preferences. Thus, our results have the potential to apply to the IS industry generally. IS interfaces compete to cognitively lock-in their incumbent users and unlock the users of other interfaces (Murray & Häubl, 2011). Previous research has shown that cognitive lock-in can be overridden by psychological responses, such as to the denial of freedom of choice (Murray & Häubl, 2011). Denying freedom of choice among IS interfaces is difficult when users can bring their own devices to work. This research investigated the potential for a different strategy: the use of computerised feedback.

These results suggest how feedback can be used to influence user preference for incumbent or competing interfaces. Normative feedback distracts from learning the skills needed to improve task performance when using a new interface. An incumbent can use normative feedback to distract users from noticing the task performance advantage of a new competitor. Positive feedback is additionally helpful for hedonic interfaces, which are used for enjoyment as well as productivity. Hedonic interfaces that offer positive feedback will create a more pleasant experience that encourages users to stay loyal even when a task can be performed better using other interfaces.

Our research used positive and negative feedback to generate psychological responses to hedonic interfaces, but we note that our research was based on prior work that manipulated psychological responses to utilitarian interfaces (Murray & Häubl, 2011). Together, the results of our study, in combination with prior research, suggest that managers can use feedback to elicit psychological responses to workplace IS systems. Additional research is required to confirm and replicate our work, but these results suggest that psychological responses, generated by feedback, could be used to discourage the use of undesirable hedonic interfaces – such as social media (Turel, 2014) – and encourage the use of new utilitarian interfaces. Since task performance has its strongest influence on interface preference when no feedback is provided, managers may be able to gradually phase out the use of feedback, to cognitively lock-in users to a new system.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by a Social Sciences and Humanities Research Council of Canada Insight Grant [#435-2015- 0100] awarded to the second author.

ORCID

Steven Bellman http://orcid.org/0000-0002-0085-2014
Kyle B. Murray http://orcid.org/0000-0003-1961-8161

References

Alba, J. W, & Hutchinson, J. W. (1987). Dimensions of consumer expertise. *Journal of Consumer Research*, *13* (4), 411–454. doi:10.1086/jcr.1987.13.issue-4

Baadte, C., & Kurenbach, F. (2017). The effects of expectancy-incongruent feedback and self-affirmation on task performance of secondary school students. *European Journal of Psychology of Education*, *32*, 113–131.

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, *16*(1), 74–94.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Becker, G. (1993). *Human capital*. Chicago, IL: University of Chicago Press.

Benyon, D., Turner, P., & Turner, S. (2005). *Designing interactive systems: People, activities, contexts, technologies*. Essex, England: Pearson Education Limited.

Brouwer, R., Dekker, T., Rolfe, J., & Windle, J. (2010). Choice certainty and consistency in repeated choice experiments. *Environmental Resource Economics*, *46*(1), 93–109.

Camp, G., Paas, F., Rikers, R., & Van Merrienboerd, J. (2001). Dynamic problem selection in air traffic control training: A comparison between performance, mental effort and mental efficiency. *Computers in Human Behavior*, *17*, 575–595.

Carter, M., Wright, R., Thatcher, J. B., & Klein, R. (2014). Understanding online customers' ties to merchants: The moderating influence of trust on the relationship between switching costs and e-loyalty. *European Journal of Information Systems*, *23*, 185–204.

Chuah, S. H.-W., Rauschnabel, P. A., Krey, N., Nguyen, B., Ramayah, T., & Lade, S. (2016). Wearable technologies: The role of usefulness and visibility in smartwatch adoption. *Computers in Human Behavior*, *65*, 276–284.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design & analysis issues for field settings*. Chicago: Rand McNally College Publishing Company.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011) From game design elements to gamefulness: Defining "gamification." In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future*

Media Environments (MindTrek'11) (A. Lugmayr, H. Franssila, C. Safran, & I. Hammouda, Eds), pp 9–15, ACM, New York.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.

Finkelstein, S. R., & Fishbach, A. (2012). Tell me what I did wrong: Experts seek and respond to negative feedback. *Journal of Consumer Research*, *39*(1), 22–38.

Finkelstein, S. R., Fishbach, A., & Tu, Y. (2017). When friends exchange negative feedback. *Motivation and Emotion*, *41*, 69–83.

Fogg, B. J., & Nass, C. (1997). Silicon sycophants: The effects of computers that flatter. *International Journal of Human-Computer Studies*, *46*(5), 551–561.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50.

Förster, J., Grant, H., Idson, L. C., & Higgins, E. T. (2001). Success/failure feedback, expectancies, and approach/avoidance motivation: How regulatory focus moderates classic relations. *Journal of Experimental Social Psychology*, *37*(3), 253–260.

Goodman, J. S., Wood, R. E., & Chen, Z. (2011). Feedback specificity, information processing, and transfer of training. *Organizational Behavior and Human Decision Processes*, *115*(2), 253–267.

Goodman, J. S., Wood, R. E., & Hendrickx, M. (2004). Feedback specificity, exploration, and learning. *Journal of Applied Psychology*, *89*(2), 248–262.

Hair, J. F., Sarstedt, M., Ringle, C. M., & Mena, J. A. (2012). An assessment of the use of partial least squares structural equation modeling in marketing research. *Journal of the Academy of Marketing Science*, *40*(3), 414–433.

Hayes, A. F. (2015). An index and test of linear moderated mediation. *Multivariate Behavioral Research*, *50*, 1–22.

Hirschman, E. C., & Holbrook, M. B. (1982). Hedonic consumption: Emerging concepts, methods and propositions. *Journal of Marketing*, *46*(3), 92–101.

Holbrook, M. B., Chestnut, R. W., Oliva, T. A., & Greenleaf, E. A. (1984). Play as consumption experience: The roles of emotions, performance, and personality in the enjoyment of games. *Journal of Consumer Research*, *11*(2), 728–739.

Jastrzembski, T. S., & Charness, N. (2007). The model human processor and the older adult: Parameter estimates and validation within a mobile phone task. *Journal of Experimental Psychology*, *13*(4), 224–248.

Johnson, E. J., Bellman, S., & Lohse, G. L. (2003). Cognitive lock-in and the power law of practice. *Journal of Marketing*, *67*(2), 62–75.

Klemperer, P. (1987). Markets with consumer switching costs. *Quarterly Journal of Economics*, *102*(2), 375–394.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*(2), 254–284.

Labrecque, J. S., Wood, W., Neal, D. T., & Harrington, N. (2017). Habit slips: When consumers unintentionally resist new products. *Journal of the Academy of Marketing Science*, *45*, 119–133.

Lakshmanan, A., & Krishnan, H. S. (2011). The aha! Experience: Insight and discontinuous learning in product usage. *Journal of Marketing*, *75*(6), 105–123.

Lakshmanan, A., Lindsey, C. D., & Krishnan, H. S. (2010). Practice makes perfect? When does massed learning improve product usage proficiency? *Journal of Consumer Research*, *37*(4), 599–613.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*(4), 492–527.

Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis*. New York: John Wiley and Sons.

Luo, L., Ratchford, B. T., & Yang, B. (2013). Why we do what we do: A model of activity consumption. *Journal of Marketing Research*, *50*(1), 24–43.

Maier, U., Wolf, N., & Randler, C. (2016). Effects of a computer-assisted formative assessment intervention based on multiple-tier diagnostic items and different feedback types. *Computers & Education*, *95*, 85–98.

McFadden, D. (1986). The choice theory approach to market research. *Marketing Science*, *5*(4), 275–298.

McFadden, D. (1981). Econometric models of probabilistic choice. In C. Manski & D. McFadden (Eds), *Structural analysis of discrete data* (pp. 198–272). Cambridge, MA: MIT Press.

Merikivi, J., Tuunainen, V., & Nguyen, D. (2017). What makes continued mobile gaming enjoyable? *Computers in Human Behavior*, *68*, 411–421.

Morales, A. C., Amir, O., & Lee, L. (2017). Keeping it real in experimental research – Understanding when, where, and how to enhance realism and measure consumer behavior. *Journal of Consumer Research*, *44*(2), 465–476.

Murray, K. B., & Bellman, S. (2011). Productive play time: The effect of practice on consumer demand for hedonic experiences. *Journal of the Academy of Marketing Science*, *39*(3), 376–391.

Murray, K. B., & Häubl, G. (2003). A human capital perspective of skill acquisition and interface loyalty. *Communications of the ACM*, *46*(12), 272–278.

Murray, K. B., & Häubl, G. (2007). Explaining cognitive lock-in: The role of skill-based habits of use in consumer choice. *Journal of Consumer Research*, *34*(1), 77–88.

Murray, K. B., & Bellman, G. (2011). Freedom of choice, ease of use, and the formation of interface preferences. *MIS Quarterly*, *35*(4), 955–976.

Murray, K. B., & Häubl, G. (2002). The fiction of no friction: A user skills approach to cognitive lock-in. In S. M. Broniarczyk & K. Nakamoto (Eds), *Advances in consumer research* (Vol. 29, pp. 11–18). Valdosta, GA: Association for Consumer Research.

Myers, R. H. (1986). *Classical and modern regression with applications*. Boston: Duxbury.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed), *Cognitive skills and their acquisition* (pp. 1–55). Hillsdale, NJ: Lawrence Erlbaum Associates.

Ouellette, J. A., & Wood, W. (1998). Habit and intention in everyday life: The multiple processes by which past behavior predicts future behavior. *Psychological Bulletin*, *124*(1), 54–74.

Pavlov, I. P. (1928). *Lectures on conditioned reflexes: Twenty-five years of objective study of the higher nervous activity (behavior) of animals (GANTT WH, trans)*. New York: Liveright.

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, *63*, 539–569.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling methods for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*(3), 879–891.

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research*, *42*(1), 185–227.

Ratchford, B. T. (2001). The economics of consumer knowledge. *Journal of Consumer Research*, *27*(4), 397–411.

Rogers, R. (2017). The motivational pull of video game feedback, rules, and social interaction: Another self-determination theory approach. *Computers in Human Behavior*, *73*, 446–450.

Schilling, M. A., Vidal, P., Ployhart, R. E., & Marangoni, A. (2003). Learning by doing something else: Variation, relatedness, and the learning curve. *Management Science*, *49*(1), 39–56.

Shapiro, C., & Varian, H. R. (1999). *Information rules*. Boston: Harvard Business School Press.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, *84*(2), 127–190.

Stigler, G. J., & Becker, G. S. (1977). De gustibus non est disputandum. *American Economic Review*, *67*(2), 76–90.

Stoll, B., Edwards, C., & Edwards, A. (2016). "Why aren't you a sassy little thing": The effects of robot-enacted guilt trips on credibility and consensus in a negotiation. *Communication Studies*, *67*(5), 530–547.

Sun, Y., Liu, D., Chen, S., Wu, X., Shen, X.-L., & Zhang, X. (2017). Understanding users' switching behavior of mobile instant messaging applications: An empirical study from the perspective of push-pull-mooring framework. *Computers in Human Behavior*, *75*, 727–738.

Takahashi, D. (2014) Mobile gaming could drive entire video game industry to $100B in revenue by 2017. *Venture Beat*,Retrieved September 20, 2016. January 14. Retrieved from http://venturebeat.com/2014/01/14/mobile-gaming-could-drive-entire-game-industry-to-100b-in-revenue-by-2017/

Te'eni, D. (1992). Analysis and design of process feedback in information systems: Old and new wine in new bottles. *Accounting, Management and Information Technologies*, *2*(1), 1–18.

Turel, O. (2014). Quitting the use of a habitual hedonic information system: A theoretical model and empirical examination of Facebook users. *European Journal of Information Systems*, *24*(4), 431–446.

Van Der Heijden, H. (2004). User acceptance of hedonic information systems. *MIS Quarterly*, *28*(4), 695–704.

Van Reekum, C. M., Johnstone, T., Banse, R., Etter, A., Wehrle, T., & Scherer, K. R. (2004). Psychophysiological responses to appraisal dimensions in a computer game. *Cognition and Emotion*, *18*(5), 663–688.

Vancouver, J. B., & Tischner, E. C. (2004). The effect of feedback sign on task performance depends on self-concept discrepencies. *Journal of Applied Psychology*, *89*(6), 1092–1098.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, *27*(3), 425–478.

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly*, *36*(1), 157–178.

Wernerfelt, B. (1985). Brand loyalty and user skills. *Journal of Economic Behavior and Organization*, *6*(4), 381–385.

Wood, W., & Rünger, D. (2016). Psychology of habit. *Annual Review of Psychology*, *67*, 289–314.

Zingoni, M., & Byron, K. (2017). How beliefs about the self influence perceptions of negative feedback and subsequent effort and learning. *Organizational Behavior and Human Decision Processes*, *139*, 50–62.

Appendix 1. Self-report measures

Prior experience (3 items)

To what extent are you someone who….?

Hates video games 1 2 3 4 5 6 7 Loves video games

Never plays video games 1 2 3 4 5 6 7 Plays video games daily

How experienced are you at playing video games like the one that you just played?

1 = "no experience at all (I have never played games like this before)";

2 = "very little experience (I've played games like this at least once before)";

3 = "some experience (I've played games like this several times)";

4 = "a moderate amount of experience (I play games like this about as much as the average person)";

5 = "more experience than most people (I play games like this at least once a month)";

6 = "very experienced (I play games like this at least once a week)";

7 = "a very great amount of experience (I play games like this at least once a day)"

Appendix 2. Smart PLS results

	Dependent variable	
Independent variables	Task performance[a]	Interface preference[b]
Interface usage[c] (*U*)	0.14***	0.18*
Prior experience (*E* [mean centred])	–	0.08
Positive feedback (*POSITIVE*)	–	0.07
Neutral feedback (*NEUTRAL*)	–	0.06
Negative feedback (*NEGATIVE*)	–	0.11
$U \times POSITIVE$	–	−0.17[d]
$U \times NEUTRAL$	–	−0.01
$U \times NEGATIVE$	–	−0.11
Task Performance	–	0.15**
R^2	.02	.04

Note. Results based on 1000 bootstrap samples. SRMR = 0.03.
[a]Measured by relative score per second (*SPS* for competitor minus *SPS* for incumbent). Zero *RSPS* means the person is indifferent between the competitor and the incumbent IS interfaces. Positive *RSPS* means that the competitor IS interface is more efficient to use than the incumbent IS interface.
[b]Coded 0 = chose incumbent interface, 1 = chose competitor interface.
[c]Coded 0 = 1 competitor interface trial, 1 = 3 competitor interface trials.
[d]p = .06, *p < .05, **p < .01, ***p < .001.