10-2013

The Effects of the Temporal Placement of Feedback on Performance and Skill Acquisition of a Medical Data Entry Task

Nathan Bechtel
Western Michigan University, nathan.t.bechtel@gmail.com

Follow this and additional works at: http://scholarworks.wmich.edu/masters_theses
🔗 Part of the Experimental Analysis of Behavior Commons, and the Industrial and Organizational Psychology Commons

Recommended Citation
http://scholarworks.wmich.edu/masters_theses/423

This Masters Thesis—Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Master’s Theses by an authorized administrator of ScholarWorks at WMU. For more information, please contact maira.bundza@wmich.edu.
THE EFFECTS OF THE TEMPORAL PLACEMENT OF FEEDBACK ON PERFORMANCE AND SKILL ACQUISITION OF A MEDICAL DATA ENTRY TASK

by

Nathan Bechtel

A thesis submitted to the Graduate College in partial fulfillment of the requirements for the degree of Master of Arts
Department of Psychology
Western Michigan University
October 2013

Thesis Committee:
Heather McGee, Ph.D., Chair
Alyce Dickinson, Ph.D.
Bradley Huitema, Ph.D.
This study was conducted in a laboratory, and compared the effects of the
temporal placement of feedback on task performance and skill acquisition. Temporal
placement of feedback refers to the timing of feedback delivery relative to performance.
Two temporal placements were examined: feedback immediately after performance and
feedback immediately prior to performance. The experimental design utilized was a form
of Latin square design that included a combination of one between-group and two within-
subjects factors. Participants were randomly assigned to one of three groups, which
differed only in the phase order utilized. Participants performed a data entry task intended
to replicate the job of a medical data entry professional.

The primary dependent variable was the number of correctly completed patient
record slides per experimental session. During feedback phases, participants were
provided with individual, graphic feedback. Participants in the baseline phase received no
feedback.

The results of this study indicate no significant differences in performance
associated with the experimental conditions, and no significant differences in the speed of
skill acquisition associated with the experimental conditions.
ACKNOWLEDGMENTS

I would like to begin by acknowledging Dr. Alyce Dickinson of Western Michigan University, without whom I may never have discovered my passion for behavior analysis or organizational behavior management. Her devotion to OBM and tireless teaching efforts inspired me to pursue my graduate school career and helped to guide me during my pursuit of my doctorate.

Secondly, I would like to thank my graduate advisor, Dr. Heather McGee for taking me into the graduate program and mentoring me for the past three years. Heather is always around to help on projects, provide helpful and meaningful feedback, and is never too busy to give her students a helping hand. Her efforts (and rigorous editing) have helped me through the arduous task of completing this paper and I could not have asked for a better guide.

I would also like to thank Dr. Bradley Huitema for his efforts on the statistical analysis portion of this study. I’m quite sure I took up more of his office hours and time than any of his students has a right to take up, and I am greatly appreciative of all the help he offered me.

Lastly I would like to thank all my fellow graduate students (particularly my lab mates) for supporting me throughout the research process. Whenever the task seemed to daunting they were always around to buoy my spirits.

Nathan Bechtel
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vii</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Feedback Overview</td>
<td>1</td>
</tr>
<tr>
<td>Feedback Definitions</td>
<td>2</td>
</tr>
<tr>
<td>Feedback Dimensions</td>
<td>3</td>
</tr>
<tr>
<td>Temporal Placement</td>
<td>9</td>
</tr>
<tr>
<td>Underlying Behavioral Principles</td>
<td>13</td>
</tr>
<tr>
<td>METHOD</td>
<td>15</td>
</tr>
<tr>
<td>Participants</td>
<td>15</td>
</tr>
<tr>
<td>Setting/Apparatus</td>
<td>17</td>
</tr>
<tr>
<td>Experimental Task</td>
<td>17</td>
</tr>
<tr>
<td>Experimental Question</td>
<td>18</td>
</tr>
<tr>
<td>Experimental Design</td>
<td>18</td>
</tr>
<tr>
<td>Dependent Variables</td>
<td>19</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>20</td>
</tr>
<tr>
<td>Independent Variable</td>
<td>21</td>
</tr>
<tr>
<td>Feedback immediately prior</td>
<td>21</td>
</tr>
<tr>
<td>Feedback immediately after</td>
<td>21</td>
</tr>
</tbody>
</table>

iii
Table of Contents – Continued

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental Procedures</td>
<td>21</td>
</tr>
<tr>
<td>Introductory meeting</td>
<td>21</td>
</tr>
<tr>
<td>Initial session</td>
<td>22</td>
</tr>
<tr>
<td>Attrition</td>
<td>22</td>
</tr>
<tr>
<td>Experimental sessions</td>
<td>23</td>
</tr>
<tr>
<td>Debrief</td>
<td>23</td>
</tr>
<tr>
<td>RESULTS</td>
<td>25</td>
</tr>
<tr>
<td>Performance between Feedback Arrangements</td>
<td>25</td>
</tr>
<tr>
<td>Latin Square Analysis of Variance</td>
<td>27</td>
</tr>
<tr>
<td>Skill Acquisition Analysis</td>
<td>31</td>
</tr>
<tr>
<td>Questionnaire Analysis</td>
<td>33</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>38</td>
</tr>
<tr>
<td>Overall Performance</td>
<td>38</td>
</tr>
<tr>
<td>Skill Acquisition</td>
<td>40</td>
</tr>
<tr>
<td>Questionnaire</td>
<td>41</td>
</tr>
<tr>
<td>Limitations</td>
<td>42</td>
</tr>
<tr>
<td>Future Research</td>
<td>43</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>47</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>51</td>
</tr>
<tr>
<td>A. Research Recruitment Flyer</td>
<td>51</td>
</tr>
<tr>
<td>B. Participant Recruitment Script</td>
<td>53</td>
</tr>
</tbody>
</table>
Table of Contents – Continued

C. HSIRB Approval.................................................................................................................. 55
D. Inclusion Questionnaire ..................................................................................................... 57
E. Graph Comprehension Quiz .............................................................................................. 59
F. Informed Consent Form....................................................................................................... 61
G. Sample Task Screen........................................................................................................... 65
H. Dependent Variable Calculation Formulae........................................................................ 67
I. Satisfaction Questionnaire ................................................................................................. 69
J. Sample Feedback Graph ..................................................................................................... 71
K. Feedback Script.................................................................................................................. 73
L. Informed Consent and Inclusion Questionnaire Meeting Script........................................ 76
M. Compensation Receipt....................................................................................................... 78
N. Debrief Script.................................................................................................................... 80
LIST OF TABLES

1. Latin Square ANOVA Table................................................................. 28
2. Fisher-Hayter Results (groups)............................................................ 29
3. Fisher-Hayter Results (phases)............................................................ 29
4. Average Group Accuracy by Condition.............................................. 30
5. Average Group Rate of Completion by Condition .............................. 30
6. Average Group Time On-Task by Condition (in Minutes).................... 31
7. Group Averages of Feedback Satisfaction Survey ................................. 34
8. Between-groups Two-sample t-test Results for Feedback Satisfaction Survey ................................................................. 35
9. Within- and Between-groups Paired t-test Results for Feedback Preference. 37
LIST OF FIGURES

1. Average number of correctly completed slides for each group, categorized by experimental condition ................................................................. 25

2. Average number of correctly completed slides for each phase, categorized by group .................................................................................................................. 26

3. Average number of correctly completed slides for each experimental condition, categorized by group ................................................................. 26

4. Average rate of completion in the first session (prior to any feedback) for each group .................................................................................................................. 27

5. Simple linear regression lines for each group, indicating the rate of skill acquisition ........................................................................................................... 32

6. Boxplot of feedback preference questionnaire results on a 5-point Likert scale ........................................................................................................... 36
INTRODUCTION

Feedback Overview

The use of performance feedback is seen in a variety of organizational applications to improve the performance of employees. Reviews of the associated literature within the field of organizational behavior management (OBM) show the extent to which feedback is used (Balcazar, Shupert, Daniels, Mawhinney, & Hopkins, 1989; Nolan, Jarema, & Austin, 1999). Balcazar et al., in a review of the first ten years of the Journal of Organizational Behavior Management (JOBM), determined that performance feedback was used in 50% of the articles. Furthermore, Nolan et al. found that the use of performance feedback increased to 71% over the following ten years.

The importance of performance feedback is also visible in the related feedback literature reviews (Alvero, Bucklin, & Austin, 2001; Balcazar, Hopkins, & Suarez, 1985-1986). These literature reviews traverse four journals (i.e., the Journal of Organizational Behavior Management [JOBM], Journal of Applied Psychology [JAP], Journal of Applied Behavior Analysis [JABA], and Academy of Management Journal [AMJ]) and twenty-five years of application between them. Balcazar et al. found that these four journals, over eleven years (1974-1984), yielded 126 applications of feedback and Alvero et al. found 68 applications within the same journals over fourteen years (1985-1998). Feedback is, by a wide margin, the most prevalent intervention in the field of OBM, due primarily to its ease of use, low cost, flexibility, and programmatic simplicity (Prue & Fairbank, 1981); however little is known about the behavioral principles affecting its use.
Feedback Definitions

Due to its prevalence, many definitions of performance feedback have been posited over the years; however, there is no consensus in the field regarding its exact definition. Prue and Fairbank (1981) define performance feedback as information provided to individuals about the quantity or quality of their past performance. A similar definition is provided by Daniels and Daniels (2004), who define feedback as “…information about performance that allows a person to change his/her behavior” (p. 171). However, Daniels’ and Daniels’ definition includes a provision which states:

Feedback must be combined with a consequence if change is to take place. To be considered performance feedback, information must serve at least two functions. First, it must tell you where you stand relative to some target or goal. Second, it should tell you what to do to improve. (p. 171)

Caveats regarding how the performance feedback is used are not unique to this definition. Urschel (2011) defined feedback as “…information about past performance that may be used by the performer to adjust future performance” (p. 8). If an overall consensus regarding the definition of feedback can be stated, it is that feedback provides information regarding past performance, not simply information regarding the task being performed.

Due to the deluge of definitions for feedback, practitioners and academics alike have used the term to describe a capacious amount of interventions. Feedback has been used as an umbrella term to describe anything from supervisory praise to corrective instruction (Peterson, 1982), thereby making it difficult to ascertain the intended definition when people use the term ‘feedback’ in research. Peterson (1982) suggested
that this ambiguity may be suppressed if behavior analysts abandoned the term feedback altogether.

Feedback Dimensions

While feedback is used extensively to improve the performance of employees, it must be noted that feedback can vary along numerous dimensions, including (a) the source that provides the feedback (e.g., supervisor, researcher), (b) the feedback medium (e.g., written, graphic, verbal), (c) the recipients of the feedback (e.g., group, individual), (d) content of the feedback (e.g., comparison of performance with past performance, comparison of performance with group performance), (e) components of the feedback (e.g., specificity, information provided), and (f) the frequency of the feedback (e.g., daily, weekly, monthly) (Prue & Fairbank, 1981). These factors, and their effects on the effectiveness of feedback, are examined within the two major feedback literature reviews: Balcazar et al. (1985-1986); and Alvero et al. (2001).

Balcazar et al. (1985-1986) and Alvero et al. (2001) found that the most frequently used application of feedback was feedback alone (37% and 29% respectively). Despite the popularity of feedback alone, it did not result in the most consistent effects (28% and 47% respectively). Balcazar et al. found that the most consistent effects (desired changes in mean performance from baseline) occurred when feedback was combined with goal setting (53% consistent). Contrary to Balcazar et al., Alvero et al. found that the most consistent effects occurred when feedback was combined with antecedents other than goal setting (100% consistent); however, being as there were only four applications of feedback combined with antecedents other than goal setting, it is important to exercise caution when analyzing this result. Generally consistent effects
were also found for feedback combined with antecedents and behavioral consequences (67% consistent), and feedback combined with goal setting and behavioral consequences (67% consistent). It is likely that the consistency of these combinations resulted from supplemental effects of the behavioral consequences. While feedback likely caused the initial increases in performance, the increased behavioral consequences continued to increase and maintain the performance. The two literature reviews also had conflicting results regarding the highest percentage of mixed effects (desired performance changes for some, but not all, participants). Balcazar et al. found the highest percentage of mixed effects for feedback alone (57%), while Alvero et al. found the highest percentage when feedback was combined with goal setting alone (57%).

Feedback medium, also referred to as “mechanism” (Balcazar et al., 1985-1986), refers to the form of feedback provided to the participants, be it graphic, verbal, written, or some combination of the three. Alvero et al. (2001) found that the most frequently used feedback medium was written (26%). This is likely due to the fact that written feedback requires the least amount of effort on the part of the experimenters. Written feedback resulted in relatively consistent effects (53% consistent, 47% mixed), but it was not the most consistent feedback medium studied. The most consistent effects were found with combinations of at least two feedback mediums: written and graphic (86% consistent) and verbal and graphic (75% consistent). There are at least three possible explanations for these results: (a) these increased effects were due to the increased exposure of the participants to the feedback; (b) no data were provided pertaining to the temporal placement of feedback when multiple mediums were employed; therefore, the two types of feedback may have been provided at different times relative to the
performance, and this could have played some role in the increased performance; and (c) different participants were more apt to understand one form of feedback over another. In any case, it seems that providing participants with multiple forms of feedback will increase the consistency of the effects.

The term “feedback recipients”, also referred to as “aggregation level of feedback” (Goltz, Citera, Jensen, Favero, & Komaki, 1990), refers to the participants whose performance the feedback is based on, and to whom it is provided. There are three options regarding feedback recipients: (a) individual, (b) group, or (c) both. Alvero et al. (2001) found that feedback is most commonly provided on an individual basis (65% of articles), which corroborates the findings of Balcazar et al. (1985-1986). However, Balcazar et al. and Alvero et al. found that group feedback provides the most consistent results (48% and 71% respectively). In an applied research study comparing individual and group feedback, Emmert (1978) found that, when supplementing group feedback, individual feedback resulted in “…an immediate increase in productivity, and generally the data remain at a higher level” (p. 141). These findings indicate that individual feedback actually helps to enhance the effects of group feedback. Goltz et al. purported similar findings when supplementing group feedback with individual feedback. Using an ABCB design, Goltz et al. determined that individual feedback (C) increased performance substantially when added to group feedback. However, upon reversal to group feedback alone, there was no significant decrease in performance. These results indicate that “…individual feedback can enhance performance improvements attained by group feedback, but is not critical in maintaining these improvements” (Goltz et al., 1990, p. 78).
Feedback content does not strictly refer to the information provided to recipients; rather, it refers to the information to which the feedback is compared. For example, feedback content could be a comparison of an individual’s performance to his or her past performance, a comparison of an individual’s performance to a standard performance, a comparison of an individual’s performance to group performance, etc. Balcazar et al. (1985-1986) found that the most popular content involved a comparison of individual performance to his or her past performance (55% of applications). Alvero et al. (2001) partially corroborated these findings, determining that two contents of feedback were equally popular: (a) comparison of individual performance to his or her past performance, and (b) comparison of individual performance with a standard of performance (28% of applications for each). The consistency results differed across the two literature reviews. Balcazar et al. found that a comparison of individual performance to his or her past performance and standard individual performance had the most consistent effects (100%), while Alvero et al. determined three contents of feedback which were almost equally consistent: (a) group performance compared with standard group performance (75%), (b) individual performance compared with standard individual performance or past individual performance (75%), and (c) group performance compared with past group performance (71%). The common variable amongst the most consistent feedback contents is that the performance in question is being compared to a similar performance. That is, individual performance is being compared to individual performance (be it past or standard) and group performance is being compared to group performance. The results indicate that as long as the comparative performance is produced by a similar performer (individual to individual and group to group), the feedback should consistently affect
performance. Unfortunately, there are few, if any, articles comparing the effects of different contents to one another.

There is extensive evidence that employee performance is strongly affected by specific, difficult goals when implemented in their natural environment (Locke, Shaw, Saari, & Latham, 1981); however, the effects of goal setting play a relatively minute role relative to feedback. When implemented subsequent to feedback, goal setting has been shown only to provide minor (if any) performance improvements (Calpin, Edelstein, & Redmon, 1988). Locke (1982) conducted an experiment to compare the use of easy and difficult goals. Multiple sections of an introductory psychology class were given a brainstorming exercise and asked to meet a specific goal (from 2 to 28 answers depending upon the section). The results indicated a sharp increase in performance as goal difficulty increased. Lorenzi (1988) replicated Locke's study on goal setting while adding an additional variable: an indirect, implicit incentive. Students were given the same brainstorming task and instructions; however, upon completion of the task, 2-4 students were chosen at random and given $1.00-$4.00 for their "exemplary" performance. Lorenzi found that "even an implicit, indirect monetary reinforcer is a more powerful discriminative stimulus than a specific goal" (p. 68). Therefore, although goal setting can be an effective antecedent, it does not consistently enhance the effects of feedback.

Since feedback alone does not consistently change behavior, it is sometimes difficult to conduct feedback research merely using feedback. Balcazar et al. (1985-1986) and Alvero et al. (2001) found that more consistent effects are produced when feedback is combined with antecedents (e.g., training, prompts) or consequences (e.g., praise,
monetary incentives). Unfortunately, it can often be difficult to determine effects of feedback when they are combined with consequences such as performance-based incentive pay. Bucklin, McGee, and Dickinson (2003) found that, after presenting feedback combined with incentive pay, removal of the feedback did not result in a decrease in participant performance. Their data suggest that feedback leads to initial increases in performance, which are then maintained by the continued consequences; thus, removal of the feedback condition does not result in a return to baseline. If research involving a return to baseline or any other phase change is to be conducted, it may be problematic to use performance-based incentives.

According to Balcazar et al. (1985-1986) and Alvero et al. (2001), the most common source of feedback delivery in the literature is supervisors or managers (43% and 53%). These applications yielded relatively consistent effects (50% and 59%); however, in a laboratory setting it is difficult to replicate the supervisor or manager source. Feedback provided by researchers also yielded somewhat consistent results in both reviews (32% and 50%), but not nearly equivalent to those of the supervisor/manager source. In laboratory studies, it is often difficult to establish an individual as a supervisor for the purposes of feedback delivery. Tittelbach, Fields, and Alvero (2008) found little difference between the effects of feedback delivered by “supervisors” (who were actually researchers) and peers in a laboratory setting. This is likely due to the fact that researchers do not have the same control over future reinforcement of performers as real-world supervisors would have. This indicates that the source of feedback in a laboratory setting is less important than the medium, frequency, content or supplemental antecedents, goals and consequences.
The real-world work environment provides many opportunities for reinforcement other than work behaviors. It has been argued that although behavioral interventions may increase employee proficiency, its primary effect is increasing the time employees spend working (Dickinson & Gillette, 1993). In order to appropriately replicate a work environment, it is often necessary to provide opportunities for non-work behaviors (e.g., games, social interactions, etc.). Matthews and Dickinson (2000) found significant differences in time allocated to working (vs. playing computer games) when performance-based incentives were provided. While their study did not involve feedback, it showed the significance of alternative activities to creating a realistic work environment within a laboratory setting.

Temporal Placement

The factors affecting feedback applications within the relative literature reviews were feedback source, mechanism (medium), frequency, participants, privacy and content (Alvero et al., 2001; Balcazar et al., 1985-1986). The temporal placement of the feedback relative to the performance is neglected in these reviews. There are four logical reasons for this omission: (a) the research articles examined did not specify the temporal placement of feedback, (b) the majority of feedback applications provided feedback at the same time relative to the performance, (c) the temporal placement of feedback had no effect on feedback effectiveness, and/or (d) the importance of the temporal placement of feedback was not recognized by the authors. Whatever the reason(s) for this factor being omitted, the fact remains that the temporal placement of feedback is a component which needs to be examined.
Determining the most appropriate temporal placement of feedback may help in determining the behavioral principles affecting feedback. Knowledge of feedback effectiveness as it relates to temporal placement may also enable researchers and practitioners to maximize the effectiveness of feedback and increase the celerity of acquisition of new skills. Previous research in this area has not used the term “temporal placement” to describe the timing of feedback, but rather “immediacy” or “feedback timing” (Alavosius & Sulzer-Azaroff, 1990; Brewer, 1989; Roberts, 1997; So, Lee, & Oah, 2013). For the purposes of this paper, temporal placement of feedback is defined as the point in time when feedback is delivered relative to the performance the feedback is based on or the performance feedback is meant to effect. For example, feedback may be provided immediately after the performance of some task, in which case the temporal placement would be described in relation to that performance of the task (immediately after). However, if the feedback was provided immediately prior to the next performance of the same task, the temporal placement would be described in relation to the second occurrence of the task (immediately prior).

One factor related to the temporal placement of feedback has been examined in numerous research studies: immediacy of feedback. Immediacy of feedback refers to how quickly the feedback is provided in relation to the performance it follows. Daniels and Daniels (2004) state “...the general rule on feedback is this: The sooner the better” (p. 177). This is a common notion about feedback; however, research involving the immediacy of feedback has a tendency to provide more feedback (or more intensive feedback) to those in the immediate feedback group, as opposed to the latent feedback group (Alavosius & Sulzer-Azaroff, 1990; Goomas, Smith, & Ludwig, 2011). Therefore,
the heightened improvements for those receiving immediate feedback may be caused by their repeated exposure to the feedback, rather than its temporal placement.

Feedback intensity is one of the most common extraneous variables in immediacy studies. Goomas et al. (2011) conducted an applied study to determine the effects of immediate feedback combined with goal setting in a distribution center. Performance increased dramatically upon implementation of a more immediate form of feedback; however, the baseline feedback condition was uncontrolled by the experimenters, and thus flawed. Prior to their feedback implementation, daily feedback was posted each morning in the break room; there were no contingencies placed on checking the feedback that was provided. The immediate feedback, on the other hand, was displayed on a large “scoreboard” in the main work area of the distribution center, easily visible to all those working on the floor. This immediate feedback was updated automatically throughout the day, thus providing many more instances of performance feedback and opportunities for behavior change throughout the day. While it has been argued that the improved performance shows the superiority of immediate feedback, it is difficult to ascertain exactly which factor or combination of factors contributed to the improved performance.

In a study of injury prevention, Alavosius and Sulzer-Azaroff (1990) determined that immediate feedback produced improved behaviors faster than weekly feedback (two days vs. several weeks). This study has been referenced as a defense for immediate feedback; however, the same issue arises with regard to the intensity of feedback. The immediate feedback, or CRF schedule as Alavosius and Sulzer-Azaroff defined it, was provided on a much more frequent basis than was the weekly, or fixed-interval schedule, feedback. This being the case, it cannot be said that the immediacy of feedback is
causally related to these improvements; rather, the increased number of opportunities to alter performance most likely had some effect on the performance improvements. Despite its greater performance improvements, intensive feedback is not always a viable option in real-world settings; thus, it is important to determine whether temporal placement has an effect on feedback effectiveness.

Overall, the arguments for immediate feedback can be looked at as arguments for more or more intensive feedback. In contrast with many of the arguments for immediate feedback, Tosti (1987) argues that “providing immediate feedback is frequently less effective than giving it later” (p. 18). Tosti’s bi-functional theory states that feedback can function as either a reinforcer or a discriminative stimulus. Utilizing this theory, Tosti maintains that if the intent is to increase a current response then feedback should be provided immediately after performance; however, if the intent is to teach a new response, then feedback should be provided immediately prior to the next performance of said response.

A study by Brewer (1989) attempted to determine the implications of the bi-functional theory on formative feedback. Brewer found that “…providing formative feedback immediately prior to the occurrence of the targeted behavior will be more effective in cueing changes in the targeted behavior than if formative feedback is provided at other times” (p. 47). This finding is dissimilar to many other studies pertaining to immediacy of feedback. This is most likely due to the fact that (a) many of the other studies were actually focused on feedback frequency and/or (b) many of the other studies were aimed at increasing an existing response, rather than a new behavior.
Underlying Behavioral Principles

While immediate and intensive feedbacks are naturally paired, providing delayed and intensive feedbacks simultaneously is a contradictory notion; thus, it is easy to view immediate feedback as more effective than delayed feedback. However, in order to fully comprehend the behavioral principles controlling the effectiveness of feedback, semantics such as these are very important. One of the most prevalent questions regarding feedback pertains to the behavioral principles controlling its effectiveness. Peterson (1982) states:

The question about which function it (feedback) serves or even whether it serves a dual function is inappropriate. Feedback, or information about past performance, can potentially serve any of a number of behavioral functions. (p. 101)

Due to the complex nature of feedback, it has been used as a panacea for performance issues, compatible with a plethora of behavioral principles.

Feedback can potentially be described as a consequence (reinforcement or punishment) if it is provided immediately following some behavior, and also has the potential to act as a discriminative stimulus for a behavior which follows it. However, according to Agnew and Redmon (1993) “an analysis of feedback interventions with reference to the precise definitions of discriminative stimulus and reinforcing stimulus leads to a rejection of these simple classifications” (p. 70). It is often the case that more complex behavioral concepts (i.e., contingency-specifying stimuli or “rules”) are a more logical explanatory maxim (Agnew & Redmon, 1993). “Rules”, according to Skinner (1969), are verbal stimuli which describe the behavioral contingencies affecting behavior. Self-generated rules and rules provided by managers or researchers regarding feedback
could affect performance by changing the function of stimuli associated with a change in
feedback (i.e., working harder).

With regards to the temporality of feedback, Annett (1969) argued that since feedback is provided between two behaviors, the fact that it comes before one response may be as important as the fact that it comes after another response. In other words, feedback effectiveness may be equally tied to both pre- and post-performance of the behavior in question. The purpose of this study was to determine the extent to which feedback effectiveness is related to each behavior by determining the effectiveness of pre- and post-performance feedback.
METHOD

Participants

Participants were forty-five undergraduate students from a mid-western university, recruited using a combination of flyers (Appendix A) and announcements in undergraduate-level psychology classes (Appendix B). Prior to recruitment, Western Michigan University’s Human Subjects Institutional Review Board approved the study (Appendix C).

Participants were screened based upon four exclusionary criteria (Appendices D and E). First, the participants had to report availability for 15-24, 30-minute sessions. The purpose of this availability check was to reduce attrition as much as possible and to ensure that participants were aware of the commitment required of them. Second, participants had to report using at least one of the available computer activities a minimum of two hours per month. Available activities were Solitaire, Spider Solitaire, Angry Birds, Mahjong and Bejeweled. Alternative activities were included in order to replicate the work environment and provide realistic non-work activities. True performance changes due to the independent variable could be disguised if alternative activities are neglected, because participants may spend 100% of their time performing the task. To ensure that participants were not performing exemplary due to the lack of alternative activities, the participants needed to report a certain level of interest in the alternatives. Third, participants could not have participated in any previous psychology department performance management research or held any positions involving data entry. Since this study intended to determine the effects of feedback on skill acquisition, prior knowledge of the experimental task could easily confound any data related to learning the
program. Lastly, participants were required to demonstrate a comprehensive understanding of the graphs being used to display the individual feedback. In order to ensure that performance changes were the result of the feedback being provided, participants needed to fully comprehend the graphs through which the feedback was being presented. Potential participants were provided with a brief quiz (Appendix E) to determine their understanding of the graphs being used. To qualify for inclusion, participants had to obtain a score of 100% on the quiz. In order to participate in the study, participants were required to first sign an informed consent form (Appendix F), which was provided during the initial meeting with the student investigator.

Participants were paid the equivalent of $4.50 per 30-minute session in a single payment at the end of the study. Payment was provided at the end of the study to avoid confounding the “feedback immediately after performance” phase with a payment reinforcer. Were payment given in combination with this form of feedback, but not the “feedback immediately prior” form, it could have altered the effect the feedback had on performance. If a participant decided to drop out of the experiment, he or she received the full amount earned up to that point.

A minimum performance criterion of 70 slides per session was included. Participants were required to complete at least 70 slides in each session in order to receive payment for the session. This criterion was implemented to further replicate a real work environment. If a person performs below some minimum level in any job, his or her employment will most likely be terminated. This minimum performance criterion was chosen based upon previous research using the same experimental task. Seventy slides
17 per session is the equivalent of two standard deviations below the average rate of completion for a similar population.

Setting/Apparatus

Sessions were conducted in a small university laboratory (1532 Wood) containing three work areas, separated by dividers. The work areas comprised a desktop computer, mouse, gel wrist-rest, keyboard, and adjustable chair. Participants were not allowed to bring in outside devices (phones, iPods, etc.) because we were unable to monitor off-task behaviors that were not occurring on the computer itself. Off-task behaviors were only measured by the program if they occurred for at least 30 seconds, at which point a timer would keep track of the time before the next on-task behavior. If a participant were to respond to a text or check an email on their phone, these behaviors would most likely consist of less than 30 seconds and subsequently not be recorded as off-task. Multiple occurrences of such behaviors could have eventually skewed the on-task time away from where it actually lay. By limiting participants’ access to such devices, we hoped to limit extraneous variables such as these from confounding our measurements.

Each participant met with the experimenter (or a trained research assistant) individually prior to and after each session in a separate room. During this time performance feedback was provided depending on the experimental phase the participant was currently undergoing. During the post-session meeting, participants were also scheduled for their next sessions.

Experimental Task

The experimental task was a computer-based data entry task, intended to replicate the job of a medical data entry professional (See Appendix G for sample screen). The
program instructed participants to perform two tasks per patient record: (a) type a patient’s identification number into a text box and (b) determine whether the patient’s heart rate (HR) in beats per minute (BPM) is within a specified range. Typing the patient’s ID number required the participant to merely retype a number which was located on the side of the screen. Determining whether the patient’s HR was within range required two separate steps. First, the participant ascertained the gender of the patient by inspecting the box labeled gender on the side of the page. Second, the participant compared the patient’s HR to the HR range for his or her gender. The participant then clicked on a “within range” or “out of range” button, depending on the HR. Once the patient ID and HR range were entered, the participant clicked the “submit” button, at which point a new set of patient data appeared.

Upon completion of the session, the program automatically cleared the final slide. The program has the ability to provide immediate feedback upon completion, in the form of slides completed and slides completed correctly; however, this function was not used for the purposes of this experiment.

**Experimental Question**

This experiment attempted to answer two separate experimental questions: (a) “Does the temporal placement of feedback affect overall participant performance on a data entry task?” and (b) “Does the temporal placement of feedback affect the speed of skill acquisition on a data entry task?”

**Experimental Design**

The experimental design was a combination of one between-group and two within-subjects factors. Three groups were utilized, with each group employing a
different condition order. Participants were randomly assigned to one of the three experimental groups. Each group was exposed to each experimental condition in a differing order. There were three experimental conditions: (A) Baseline, during which time no feedback was provided, (B) Feedback immediately after, during which feedback was provided immediately after the performance on which it was based, and (C) Feedback immediately prior, during which feedback was provided immediately prior to the next performance of the task on which it was based. The condition order of the three groups was (1) ABC, (2) BCA, and (3) CAB. Since each experimental condition was the first condition experienced by at least one group, effects on skill acquisition could be determined.

**Dependent Variables**

The primary dependent variable was the number of correctly completed patient record slides per experimental session. Other variables which may have affected the primary dependent variable were also measured, including time on-task, rate of completion, and accuracy. Time on-task refers to the amount of time (in seconds) spent working on the experimental task during each session. Rate of completion refers to the number of patient records completed per on-task-minute during experimental sessions. Accuracy refers to the percentage of patient records per session completed correctly.

The experimental task was computerized, and the program recorded all of the necessary variables automatically. Time off-task was defined as pauses in responding lasting longer than 30 seconds. Time on-task was calculated by subtracting the total number of seconds off-task from the total session time. Rate of completion and correct/incorrect patient records were also automatically recorded by the computer.
program. Other than the number of correctly completed patient records, these variables were not required until the end of the experiment, at which point the primary investigator calculated time on-task, rate of completion, and participant accuracy (Appendix H).

Data Analysis

The number of slides completed were analyzed both visually and statistically. A form of Latin square analysis of variance was utilized for the statistical analysis. This design allows for counterbalancing while still keeping the experiment size manageable. In a complete three condition counterbalanced design, six experimental groups would be required. The Latin square counterbalanced measures design allows us to reduce the number of groups to three, while still controlling for potential carry over effects.

Skill acquisition was also analyzed both visually and statistically. A simple regression analysis was completed on correctly completed slides per minute in the first five sessions for each group separately. The resulting simple regression lines provide a visual analysis of the skill acquisition phase of the experiment. A simple ANOVA was used to analyze the homogeneity of slopes between groups. Individual regression slopes for all participants were compared between the three groups to determine if differences existed.

A brief survey (Appendix I) was provided to all participants after completion of the study. The data from three specific questions on this survey were compared using paired t-tests. The relevant questions, scored on a 5-point Likert scale (1 = Strongly Disagree, 2 = Disagree, 3 = Undecided, 4 = Agree, 5 = Strongly Agree), were as follows: (i) “I preferred to receive feedback at the beginning of the session,” (ii) “I preferred to receive feedback at the end of the session,” and (iii) “I preferred receiving no feedback at
all.” A between-group analysis of a fourth question was conducted to determine if the order of experimental conditions had any effect on responding. This question read “I worked harder after seeing my feedback graph.”

**Independent Variable**

The independent variable was the temporal placement of feedback. Two experimental conditions were utilized: Feedback immediately prior and feedback immediately after. A baseline condition was also utilized, during which no feedback was provided.

**Feedback immediately prior.** Participants in this condition received individual graphic feedback before each session, during the pre-session meeting with the experimenter or research assistant. A line graph was provided that displayed the number of correctly completed patient records from all previous sessions (Appendix J). A predetermined instructional script (Appendix K) was read by the experimenter or research assistant to the participant at this time.

**Feedback immediately after.** Participants in this condition received individual graphic feedback after each session, during the post-session meeting with the experimenter or research assistant. A line graph was provided that displayed the number of correctly completed patient records from all previous sessions and the session just completed. A predetermined instructional script (Appendix K) was read by the student investigator or research assistant to the participant at this time.

**Experimental Procedures**

**Introductory meeting.** Participants were recruited using a combination of flyers and in-class announcements in undergraduate psychology courses (Appendices A & B).
Potential participants contacted the student investigator to determine an initial meeting time. Participants recruited from in-class announcements were contacted by the student investigator, via email, and scheduled for an initial meeting. During this introductory session, the student investigator provided an explanation of the study and obtained informed consent (Appendix L). If informed consent was obtained, potential participants were asked to complete the inclusion questionnaire and graph comprehension quiz (Appendices D & E). After completing these forms, the participants were free to leave. If potential participants met the four inclusion criteria, they were contacted regarding the first session time. All participants (regardless of inclusion) were assigned a 6-digit participant number to ensure participant confidentiality. Participants who met the inclusion criteria were scheduled for their next week of experimental sessions; if they did not meet the inclusion criteria they were contacted and informed that their participation was not required.

Participants who agreed to participate in the study and met the inclusion criteria were randomly assigned to one of three experimental groups. The student investigator used Microsoft® Excel to randomly assign participants to experimental groups.

**Initial session.** During the first session, the research assistant or student investigator explained the task and other protocol including payment method and use of alternative activities. Participants were scheduled for two to five sessions per week. During the participant’s final session of each week, the next week’s sessions were scheduled.

**Attrition.** In the case of participant attrition, participants were scheduled for one more meeting. This meeting allowed the student investigator to debrief the participant
and provide them with the compensation they had earned up to the point of attrition. The
debrief script and satisfaction survey used for participants who completed all sessions
(Appendices M & I) were also used for these participants. A total of fifteen participants
dropped out over the course of the experiment. Of these, eight attended their final debrief
meeting.

**Experimental sessions.** Participants attended two to five experimental sessions
each week. Participants were in each phase until a stability criteria of +/- 10% over three
sessions occurred. If stability did not occur within eight sessions, the participant was
moved on to the next phase. Participants met with either the student investigator or a
research assistant in the initial meeting room (1512 Wood) prior to each session. The
initial meeting is outlined in Appendix K. Depending upon which experimental condition
the participant was in, this meeting may have included a feedback graph (discussed in
detail above and in Appendix K). The participant was instructed to leave his or her
personal belongings, including phones and electronics, in the initial meeting room, which
was locked during the session. The student investigator or research assistant then led the
participant to his or her work station in 1532 Wood.

Once the 30-minute session was complete, the participant returned to 1512 Wood.
Depending upon which experimental session the participant was in, feedback may have
been provided at this time. After feedback was provided, participants were reminded of
their next session time or scheduled for the next week’s sessions. After scheduling,
participants were given their belongings and could leave.

**Debrief.** Upon completion of the last experimental session, participants were
scheduled for one, 5-10 minute debriefing session. During this session, participants were
provided with a compensation receipt (Appendix M), as well as the money they had earned up to that point. Following this, participants were debriefed (Appendix N) on their participation in the study. The debriefing session included: (a) a description of the purpose of the study, (b) an explanation of the experimental phases, (c) an explanation of the use of feedback in the study, (d) a brief satisfaction survey used to determine participant preference for the different feedbacks (Appendix I), and (e) an opportunity for participants to ask any questions they may have had regarding their participation. After being debriefed, the participant’s obligation to the study was complete and they were free to leave.
RESULTS

Performance between Feedback Arrangements

The group averages by condition can be seen in Figure 1. Visual inspection of these data indicate only slight differences between feedback conditions. Feedback prior to performance yielded the highest total average of correctly completed slides per session, while feedback after performance yielded the lowest. These slight differences were not statistically significant (see Latin Square Analysis of Variance section).

![Bar chart showing average number of correctly completed slides for each group, categorized by experimental condition.](image)

**Figure 1.** Average number of correctly completed slides for each group, categorized by experimental condition.

Differences between phases can be seen in Figure 2. Phase I had the lowest average number of slides completed, while Phase II and Phase III had significantly higher averages. Visual inspection of these data indicates that participants performed better after Phase I, regardless of the experimental condition to which they were exposed. These results are supported by the statistical analysis of these data (see Latin Square Analysis of Variance section).
Figure 2. Average number of correctly completed slides for each phase, categorized by group.

Figure 3. Average number of correctly completed slides for each experimental condition, categorized by group.

The condition averages by group can be seen in Figure 3. Visual inspection of these data indicates that Group 2 performed significantly better than Group 1 or Group 3 on average. The slight difference between Group 1 and Group 3 is also apparent, but not to the same extent. It is unclear whether these group differences are due to intrinsic differences between the groups or some kind of ordering effect. In order to assess this,
the average first session scores for each group were calculated and analyzed (see Figure 4). Visual inspection of these data indicates that Group 2 was more adept at the data entry task prior to the receipt of any feedback. These results are supported by the statistical analysis of these data (see Latin Square Analysis of Variance section).

![Figure 4. Average rate of completion in the first session (prior to any feedback) for each group.](image)

**Latin Square Analysis of Variance**

The number of correctly completed slides was the primary variable used in the Latin square ANOVA. The analysis indicated that there were no significant differences between any of the three treatment conditions (feedback prior, feedback after, or baseline). An $F$-value of 1.924 was found for the conditions in the analysis (critical $F$-value of 3.10 [$\alpha = .05$]). This resulted in a $p$-value of 0.152 for the conditions. In other words, participants performed roughly the same whether they received no feedback, feedback prior to performance, or feedback after performance. See Table 1 for the detailed Latin square ANOVA table.
Table 1

*Latin Square ANOVA Table*

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F-obtained</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Groups</td>
<td>62375</td>
<td>2</td>
<td>31187.50</td>
<td>4.03*</td>
<td>0.025*</td>
</tr>
<tr>
<td>Error A</td>
<td>325096</td>
<td>42</td>
<td>7740.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conditions</td>
<td>3529</td>
<td>2</td>
<td>1764.50</td>
<td>1.92</td>
<td>0.152</td>
</tr>
<tr>
<td>Phases</td>
<td>27398</td>
<td>2</td>
<td>13699.00</td>
<td>14.94*</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Error B</td>
<td>78858</td>
<td>86</td>
<td>916.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>497258</td>
<td>134</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note:* SS = Sum of Squares, df = Degrees of Freedom, and MS = Mean Square
* = Statistically significant at the $\alpha = .05$ level

A significant difference was found between experimental groups (Groups 1, 2 and 3, corresponding to the treatment orders ABC, BCA, and CAB respectively). An F-value of 4.03 ($p = 0.025$) was found for the groups source in the analysis. A Fisher-Hayter test was then conducted to determine between which pair(s) of groups the difference laid. The only significant, pairwise difference was found between Group 2 (BCA) and Group 3 (CAB). There was no statistically significant difference between Group 1 (ABC) and Group 2 (BCA) or Group 1 (ABC) and Group 3 (CAB). In other words, Group 2 performed significantly better than Group 3, and there was no significant difference in performance between Group 1 and Group 3 or Group 1 and 2. See Table 2 for detailed Fisher-Hayter results.

Lastly, there was a statistically significant difference between the phases of the experiment. An F-value of 14.94 ($p < .001$) was found for the phases in the analysis.
A Fisher-Hayter test was then conducted to determine between which pair(s) of phases the differences laid. There was a statistically significant difference between Phases I and II, and Phases I and III. There was no statistically significant difference between Phases II and III. In other words, participants performed significantly better in the second phase and third phase of the experiment, regardless of the feedback they received in that phase, and there was no significant difference in performance between the second phase and third phase, regardless of the feedback received in that phase. See Table 3 for detailed Fisher-Hayter results.

### Table 2

<table>
<thead>
<tr>
<th>Fisher-Hayter Results (Groups)</th>
<th>Fisher-Hayter Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pairwise comparisons</td>
<td>df</td>
</tr>
<tr>
<td>Group 1 vs. Group 2</td>
<td>45</td>
</tr>
<tr>
<td>Group 1 vs. Group 3</td>
<td>45</td>
</tr>
<tr>
<td>Group 2 vs. Group 3</td>
<td>45</td>
</tr>
</tbody>
</table>

*Note:* df = Degrees of freedom, $q$ = studentized range statistic

* = Statistically significant at the $\alpha = .05$ level

### Table 3

<table>
<thead>
<tr>
<th>Fisher-Hayter Results (Phases)</th>
<th>Fisher-Hayter Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pairwise comparisons</td>
<td>df</td>
</tr>
<tr>
<td>Phase 1 vs. Phase 2</td>
<td>45</td>
</tr>
<tr>
<td>Phase 1 vs. Phase 3</td>
<td>45</td>
</tr>
<tr>
<td>Phase 2 vs. Phase 3</td>
<td>45</td>
</tr>
</tbody>
</table>

*Note:* df = Degrees of freedom, $q$ = studentized range statistic

* = Statistically significant at the $\alpha = .05$ level
The secondary dependent variables (accuracy, rate of completion and time on-task) were also visually analyzed; however, these variables yielded no new information regarding the effects of the temporal placement of feedback. All groups averaged over 95% accuracy in all conditions. This indicates that the task was simple enough that participants did not require feedback to maintain a high rate of accuracy. See Table 4 for detailed accuracy information.

Table 4

*Average Group Accuracy by Condition*

<table>
<thead>
<tr>
<th>Feedback Condition</th>
<th>Group Condition A</th>
<th>Group Condition B</th>
<th>Group Condition C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>96.36%</td>
<td>97.31%</td>
<td>97.57%</td>
</tr>
<tr>
<td>Group B</td>
<td>98.17%</td>
<td>97.37%</td>
<td>98.11%</td>
</tr>
<tr>
<td>Group C</td>
<td>96.44%</td>
<td>97.19%</td>
<td>95.60%</td>
</tr>
</tbody>
</table>

*Note:* Accuracy calculated as total number of correctly completed slides / total number of completed slides

Rate of completion data was used to analyze the skill acquisition question (see "Skill Acquisition Analysis" section). Since the primary concern of feedback effectiveness is performance rather than task proclivity, the rate of completion data were not used in the primary analysis. These data can be seen in Table 5.

Table 5

*Average Group Rate of Completion by Condition*

<table>
<thead>
<tr>
<th>Feedback Condition</th>
<th>Group Condition A</th>
<th>Group Condition B</th>
<th>Group Condition C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>6.913</td>
<td>8.442</td>
<td>8.870</td>
</tr>
<tr>
<td>Group B</td>
<td>9.527</td>
<td>7.489</td>
<td>9.306</td>
</tr>
<tr>
<td>Group C</td>
<td>7.664</td>
<td>8.165</td>
<td>6.577</td>
</tr>
</tbody>
</table>

*Note:* Rate of completion was calculated as total number of completed slides / time on task (in minutes)
Average time on-task was also calculated for all participants in the study. The average time spent on-task was approximately 26.2 minutes per session for all participants, indicating that participants did not take exceedingly long breaks from the task (approximately 87% of time spent on-task). See Table 6 for detailed time on-task information.

Table 6

<table>
<thead>
<tr>
<th>Feedback Condition</th>
<th>Group A</th>
<th>Group B</th>
<th>Group C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition A</td>
<td>27.25</td>
<td>26.50</td>
<td>25.36</td>
</tr>
<tr>
<td>Condition B</td>
<td>24.65</td>
<td>28.05</td>
<td>24.56</td>
</tr>
<tr>
<td>Condition C</td>
<td>23.38</td>
<td>27.82</td>
<td>28.02</td>
</tr>
</tbody>
</table>

*Note:* Time on-task was calculated as (session time – time off-task)

Skill Acquisition Analysis

Rate of completion was the primary variable used to analyze skill acquisition. Upon inspection of the simple linear regression lines (rate of completion vs. number of sessions) for each group, it becomes apparent that Group 2 and Group 3 acquired the skill at virtually the same rate, while Group 1 lagged slightly. Since visual inspection alone (Figure 5) does not provide an explicit answer to the skill acquisition question, further statistical analysis was necessary.

The rate of completion for the first five sessions was used to calculate the simple linear regression lines for each experimental condition. These simple linear regression lines reflect the rate of skill acquisition for each condition. The individual regression lines for all participants were then compared using an ANOVA to determine the homogeneity of slopes. The ANOVA indicated that there was no statistically significant difference
between conditions regarding the factor of skill acquisition. In other words, no specific feedback condition produced faster learning than any other feedback condition. An $F$-value of 0.36 ($p = 0.701$) was found for the groups in the analysis.

![Figure 5. Simple linear regression lines for each group, indicating the rate of skill acquisition.](image)
Questionnaire Analysis

The complete results of the Feedback Satisfaction Survey can be seen in Table 7. All questions were statistically analyzed between-groups (via two sample $t$-tests) to determine whether or not they were of statistical relevance; however, the variables of the greatest experimental relevance are the total average Likert scores. Only one statistically significant difference was found between groups via the two sample $t$-tests. The question read “I worked harder after seeing my feedback graph.” The results indicated that there was a significant difference in responding between Group 1 and Group 3 ($p = 0.015$). There was no statistically significant difference in responding between Group 1 and Group 2, or between Group 2 and Group 3 ($p = 0.076 \& 0.631$, respectively).

A detailed outline of the between-groups questionnaire analysis can be found in Table 8.
## Table 7

**Group Averages of Feedback Satisfaction Survey**

<table>
<thead>
<tr>
<th>Feedback Satisfaction Survey Questions</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>'I tried to alter my performance harder after seeing my feedback graph'</td>
<td>4.067</td>
<td>4.133</td>
<td>4.000</td>
<td>4.067</td>
</tr>
<tr>
<td>'I worked based on my seeing my feedback graph'</td>
<td>3.333</td>
<td>4.000</td>
<td>4.333</td>
<td>3.822</td>
</tr>
<tr>
<td>'I enjoyed seeing my feedback graph'</td>
<td>3.933</td>
<td>4.333</td>
<td>4.200</td>
<td>4.156</td>
</tr>
<tr>
<td>'The sessions were not long enough for me to receive my feedback at the beginning of each session'</td>
<td>3.333</td>
<td>3.800</td>
<td>3.467</td>
<td>3.489</td>
</tr>
<tr>
<td>'I preferred to receive my feedback at the break'</td>
<td>4.067</td>
<td>3.800</td>
<td>3.467</td>
<td>3.778</td>
</tr>
<tr>
<td>'I preferred receiving no feedback at all'</td>
<td>2.800</td>
<td>3.200</td>
<td>3.000</td>
<td>3.000</td>
</tr>
<tr>
<td>'I preferred feedback at the end of the session'</td>
<td>1.867</td>
<td>1.533</td>
<td>1.689</td>
<td>1.689</td>
</tr>
</tbody>
</table>

*Note: Questions can be found in Appendix 1*

*= Questions pertaining to Table 9 statistics*
### Table 8

**Between-groups Two-Sample t-test Results for Feedback Satisfaction Survey**

<table>
<thead>
<tr>
<th>Questions</th>
<th>95% CI</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 1 vs. Group 2</td>
<td>(-0.829, 0.696)</td>
<td>-0.18</td>
<td>0.858</td>
</tr>
<tr>
<td>Group 1 vs. Group 3</td>
<td>(-0.446, 0.579)</td>
<td>0.27</td>
<td>0.791</td>
</tr>
<tr>
<td>Group 2 vs. Group 3</td>
<td>(-0.570, 0.837)</td>
<td>0.40</td>
<td>0.696</td>
</tr>
<tr>
<td>Question 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 1 vs. Group 2</td>
<td>(-1.407, 0.074)</td>
<td>-1.85</td>
<td>0.076</td>
</tr>
<tr>
<td>Group 1 vs. Group 3</td>
<td>(-1.429, -0.171)</td>
<td>-2.65</td>
<td>0.015*</td>
</tr>
<tr>
<td>Group 2 vs. Group 3</td>
<td>(-0.703, 0.436)</td>
<td>-0.49</td>
<td>0.631</td>
</tr>
<tr>
<td>Question 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 1 vs. Group 2</td>
<td>(-0.936, 0.136)</td>
<td>-1.53</td>
<td>0.137</td>
</tr>
<tr>
<td>Group 1 vs. Group 3</td>
<td>(-0.889, 0.356)</td>
<td>-0.88</td>
<td>0.387</td>
</tr>
<tr>
<td>Group 2 vs. Group 3</td>
<td>(-0.430, 0.697)</td>
<td>0.49</td>
<td>0.630</td>
</tr>
<tr>
<td>Question 4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 1 vs. Group 2</td>
<td>(-1.264, 0.331)</td>
<td>-1.20</td>
<td>0.240</td>
</tr>
<tr>
<td>Group 1 vs. Group 3</td>
<td>(-0.783, 0.783)</td>
<td>0.00</td>
<td>1.000</td>
</tr>
<tr>
<td>Group 2 vs. Group 3</td>
<td>(-0.331, 1.264)</td>
<td>1.20</td>
<td>0.240</td>
</tr>
</tbody>
</table>

* = Statistically at the α = 0.05 level

A cross-question analysis was conducted on Questions 5-7 to determine whether there was a statistically significant difference in feedback preference. The results of the cross-question analysis indicate participants preferred receiving feedback at any time over receiving no feedback at all, and that participants reported a significantly stronger
preference for feedback prior to performance over feedback after performance. A visual representation of these data can be seen in Figure 6.

![Boxplot of feedback preference questionnaire results on a 5-point Likert scale. The boxplot whiskers indicate the highest and lowest response for each question. The boxes indicate the first quartile (top 25%), median, and third quartile (bottom 25%).](image)

Figure 6. Boxplot of feedback preference questionnaire results on a 5-point Likert scale. The boxplot whiskers indicate the highest and lowest response for each question. The boxes indicate the first quartile (top 25%), median, and third quartile (bottom 25%).

In order to augment this visual display, a paired t-test was conducted to compare each question. Statistically significant differences were found between all three questions. Feedback at the beginning of the session was strongly preferred to feedback at the end of the session with a mean difference in Likert scores of 0.778 ($p = 0.012$; Cohen's $d = 0.692$), and was also preferred to no feedback at all, with mean difference in Likert scores of 2.089 ($p < 0.001$; Cohen's $d = 2.157$). Feedback at the end of the session was strongly preferred to no feedback at all, yielding a mean difference in Likert scores of 1.311 ($p < 0.001$; Cohen's $d = 1.476$). Within-group analyses were also conducted on all three questions for each group. The results reflect those of the between-groups analysis, in that they show that feedback of either kind was strongly preferred to no feedback in all three
groups; however, there was no statistically significant difference in preference between feedback prior to performance and feedback after performance in Group 2 or Group 3. See Table 9 for complete t-test statistics.

Table 9

*Within- and Between-groups Paired t-test Results for Feedback Preference*

<table>
<thead>
<tr>
<th>Groups</th>
<th>95% CI</th>
<th>t-value</th>
<th>p-value</th>
<th>Cohen's D</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All Groups (between)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FB after vs. FB prior</td>
<td>(0.181, 1.375)</td>
<td>2.63</td>
<td>0.012*</td>
<td>0.692</td>
</tr>
<tr>
<td>FB after vs. No FB</td>
<td>(0.917, 1.705)</td>
<td>6.71</td>
<td>0.000*</td>
<td>1.476**</td>
</tr>
<tr>
<td>FB prior vs. No FB</td>
<td>(1.628, 2.550)</td>
<td>9.13</td>
<td>0.000*</td>
<td>2.157**</td>
</tr>
<tr>
<td><strong>Group 1 (within)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FB after vs. FB prior</td>
<td>(0.170, 2.364)</td>
<td>2.48</td>
<td>0.027*</td>
<td>1.071**</td>
</tr>
<tr>
<td>FB after vs. No FB</td>
<td>(0.034, 1.833)</td>
<td>2.23</td>
<td>0.043*</td>
<td>0.889**</td>
</tr>
<tr>
<td>FB prior vs. No FB</td>
<td>(1.332, 3.068)</td>
<td>5.44</td>
<td>0.000*</td>
<td>2.275**</td>
</tr>
<tr>
<td><strong>Group 2 (within)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FB after vs. FB prior</td>
<td>(-.0523, 1.723)</td>
<td>1.15</td>
<td>0.271</td>
<td>0.537</td>
</tr>
<tr>
<td>FB after vs. No FB</td>
<td>(1.087, 2.246)</td>
<td>6.17</td>
<td>0.000*</td>
<td>2.419**</td>
</tr>
<tr>
<td>FB prior vs. No FB</td>
<td>(1.392, 3.141)</td>
<td>5.56</td>
<td>0.000*</td>
<td>2.400**</td>
</tr>
<tr>
<td><strong>Group 3 (within)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FB after vs. FB prior</td>
<td>(-.638, 1.572)</td>
<td>0.91</td>
<td>0.380</td>
<td>0.440</td>
</tr>
<tr>
<td>FB after vs. No FB</td>
<td>(0.683, 1.984)</td>
<td>4.39</td>
<td>0.001*</td>
<td>1.468**</td>
</tr>
<tr>
<td>FB prior vs. No FB</td>
<td>(.958, 2.642)</td>
<td>4.58</td>
<td>0.000*</td>
<td>1.855**</td>
</tr>
</tbody>
</table>

* = Statistically significant at the $\alpha = .05$ level
** = Large effect size
DISCUSSION

**Overall Performance**

The primary purpose of this study was to determine whether the temporal placement of feedback altered the effectiveness of feedback on performance. Visual inspection of the three experimental conditions indicated only minor performance differences, with the highest performance levels occurring under the feedback provided prior to performance condition. However, these small differences were not statistically significant. When analyzed in detail, there were statistically significant differences in performance, but these differences appear to be due to performance improvements throughout the study (phase to phase increases) and intrinsic group differences, as opposed to the feedback type or the order in which the conditions were presented.

There are several possible reasons for this outcome. First, it is possible that there were no innate differences in the effects of these forms of feedback on performance. The results indicate that feedback provided prior to performance and feedback provided after performance were equally ineffective since performance increases were comparable to those of baseline. However, it is possible that these results were due to the fact that feedback was provided without any supplementary reinforcement (i.e., no praise, evaluation or incentives). These results complement those of Johnson, Dickinson and Huitema (2008), who found that objective feedback (absent of any evaluative statements) did not improve performance on a check-processing task.

Another possible reason for the lack of differences between experimental conditions is the experimental task which was utilized. The experimental task was relatively simple, and once participants became proficient at completing it they could do
so with little effort. This is supported by the fact that all groups saw significant increases in performance after the first phase with smaller improvements for two of the groups between the second and third phases, indicating an increase in fluency throughout the study. The simplicity of the task might have resulted in reinforcement for the behavior of engaging in the task, as very little response effort was required and the participants might have received natural feedback regarding the speed of typing.

The experimental task may have also provided other natural consequences. Upon typing in the patient's ID number, participants were able to see whether or not it matched the display and adjust it accordingly. Also, while the program did not tell participants the number of slides they had completed correctly, it is feasible that they were able to count the number of slides they completed. Since participants were only required to complete 70 slides in order to earn their payment for each session, it is reasonable to suspect that some participants counted until they completed this number and then engaged in alternative activities for the remainder of the session. Two participants whose data indicated such behavior were questioned by the experimenter during the debriefing session and admitted to engaging in this counting behavior. While most participants' data did not indicate such behavior, it is still a factor which bears consideration.

Lastly, the length of the sessions may have had an effect on participant performance levels. Sessions were only a half hour long, and it is likely that participants did not require a break during sessions, as indicated by the time on-task data which shows that participants spent an average of approximately 26 minutes per session on task. This implies that participants were working at full effort throughout most of each session regardless of the availability of alternative activities.
Skill Acquisition

The secondary purpose of this study was to determine if the temporal placement of feedback had an effect on skill acquisition. Visual inspection of the increase in rate of completion (correctly completed slides per minute) over the first five sessions indicated only a slight difference between feedback (prior and after) and no feedback. There were no clear differences between the two types of feedback. A statistical analysis comparing these increases confirmed that these differences were not statistically significant; thus, the temporal placement of feedback appeared to have no effect on the speed of skill acquisition for this task.

There are several possible reasons for this outcome. First, it is possible that the temporal placement of feedback has no effect on the speed of skill acquisition. While both forms of feedback increased the speed of skill acquisition slightly over the baseline condition, there was almost no difference between the effects of the two forms. Similar to the effects on overall performance, it is possible that these results were due to the fact that feedback was provided without any supplementary reinforcement or evaluative statements.

The experimental task which was utilized may have also contributed to the lack of differences in skill acquisition between experimental conditions. The experimental task was very easy to learn, and participants were capable of seeing correct typing performance as they completed the task. Therefore, the participants might not have necessarily required feedback to determine how well they were performing and subsequently increase their completion speed. The task also lent itself to counting behaviors as discussed above.
Questionnaire

A brief questionnaire was provided after the last experimental session, but prior to the debrief session. All participants completed the questionnaire before receiving payment for the sessions they completed. The main results of this questionnaire indicate that participants preferred to receive feedback prior to performance as opposed to feedback after performance or no feedback at all. Since the two feedback types did not differ in their effects on performance or skill acquisition, it is important to note these social validity results. Although performance was not significantly affected, participants’ feedback preference may have affected their overall satisfaction with the task or the experiment as a whole.

It is possible that participants preferred receiving feedback prior to performance due to the increased session time associated with receiving feedback after performance. When participants were subjected to the feedback after performance condition, they were required to sit in the meeting room after each session for roughly 1-2 extra minutes while the experimenter printed out a new feedback graph. This extra time could have acted as a punisher, which was repeatedly paired with the feedback after performance condition and subsequently decreased participant preference for this condition. However, the baseline condition was not associated with this post-session waiting period and resulted in a significantly lower participant preference, indicating that it is unlikely that the post-session waiting period had a significant effect on feedback preference.

It is also important to note the real world implications of these results. Even if the post-session waiting time acted as a punisher, the same waiting period would be associated with feedback in a real world work environment. Presenting feedback prior to
the next performance of some task (i.e. the next day) would not only negate any post-
session waiting period, but would also allow those providing the feedback more time to
prepare any graphs or supplementary reinforcement. Thus, when implementing a
feedback contingency in the workplace, it appears that feedback prior to performance is
the more socially valid method of delivery.

The questionnaire results also indicated that participants enjoyed seeing the
feedback graphs. This corroborates the social validity of providing feedback to
performers, indicating that they enjoy seeing how well they are performing on a task.
Participants also reported that they tried to alter their performance based on their
feedback graphs, and that they worked harder after seeing their feedback graphs.
Although this was not reflected in the performance data, a more difficult task or different
work environment may have produced different results.

Limitations

There were four primary limitations to this study. The first limitation is the task
that was utilized. As discussed above, the task may have been too simple to adequately
illuminate performance improvements. Participants received some natural feedback from
their typing behavior due to the fact that they could see what they were typing and
compare it to the data presented by the program. A good counter-measure for this
limitation would be to refine the program so that asterisks appear on the screen when
typing (similar to a password bar on a website). This alteration would reduce the amount
of natural feedback participants receive, making them more reliant on experimenter-
provided feedback graphs. The task also required a minimal amount of comparison
(patient heart rate was compared to the acceptable range). It may be appropriate to add another variable for comparison to increase the difficulty of the task.

The second limitation is the length of experimental sessions. As discussed in the “Results” section, participants averaged approximately 26 minutes on task per session (87% of time spent on-task). The sessions could easily be increased by 15-30 minutes in order to nullify this issue. However, increasing the session times might alter the skill acquisition results because participants would have a longer period of time in which to learn how to perform the task.

The third limitation is the lack of supplementary reinforcement to augment the feedback provided. Objective feedback was provided in order to avoid confounding the results with reinforcers. However, it is likely that the lack of performance differences between feedback (of either type) and baseline was the result of the use of objective feedback.

Lastly, the design which was utilized may have affected the results. The skill acquisition question required three groups, each beginning with a different experimental condition. By utilizing this design, we were able to determine the effects of the temporal placement of feedback on skill acquisition; however, the primary question of effects on performance may have been more easily answered with a multiple-baseline design.

Future Research

The current study lends itself to many possible future research and replication options. The lack of significant differences between feedback prior to performance and feedback after performance may be attributable to many factors of the study. Future research on the effects of the temporal placement of feedback on performance should
utilize a more difficult experimental task. A task which provides little to no natural feedback, allows for longer experimental sessions, and reduces the possibility of counting behaviors as much as possible would be ideal. This type of task would allow the feedback conditions to exert greater control over performance by limiting other contingencies. Increased task difficulty is also essential to future research in this area. One possibility is a task which requires participants to complete multiple tasks concurrently. A task of this type was utilized by Bucklin, et al, (2003). The task (SYNWORK) utilized four tasks simultaneously: memory, arithmetic, visual and auditory monitoring.

Another option for future research would be to alter the design which was utilized. A combination between- and within-groups design was used in order to answer the secondary question of skill acquisition. Ignoring the skill acquisition question would allow for more options in experimental design which would in turn allow for easier analysis of the performance question. A multiple-baseline or altering treatments design could be implemented to analyze the effects of both types of feedback on performance. The performance question could also be ignored, and a design which focuses on the skill acquisition question could be utilized. This option would still require a between-groups design since skill acquisition effects are permanent. Three groups (baseline, feedback prior, and feedback after) would be required to make this a viable option. This option would allow for more participants due to the fact that only five sessions would be required per participant, as opposed to 15-24. As many as 135 participants could be recruited for the same cost as the 45 participants who were recruited for the current study, resulting in increased power in the statistical analysis.
Another area of concentration for future research would involve adding supplementary reinforcement to the feedback contingency. While this would admittedly reduce the amount of control feedback exerts over behavior, it is likely to increase the effectiveness of conditions other than baseline. Feedback in the workplace is often related to some form of praise or incentive, so it is reasonable to study these contingencies. A pay for performance contingency would be an acceptable addition to the experiment, as long as participants still received their compensation at the end of the study as opposed to receiving it after each session. The feedback graphs could be altered to represent the amount of money earned, as well as the number of correctly completed slides. This method would hopefully reduce any direct effects of compensation on performance and cause feedback to result in improved performance.

Lastly, future research could be conducted in an applied setting as opposed to a laboratory setting. While this option would limit experimental control, it has numerous benefits which may outweigh the potential limitations. First, an applied setting would provide a more realistic (and likely more difficult) task for participants to complete. While experimental tasks are generally developed to replicate real-world work environments, perfect replication is often impossible. Second, an applied setting would help determine the logistics of each type of feedback. While feedback delivered immediately after performance is often the preferred method, it is difficult to immediately provide any graphic or comparative feedback after performance. Were feedback to be delivered prior to the next performance of some task, it would give the person delivering the feedback more time to create a visually stimulating graphic. This factor in itself could make a feedback prior condition more appealing to managers and supervisors. Lastly, the
social validity question could be revisited in an applied setting. The results of the current study indicate that participants preferred receiving feedback prior to performance, and an applied replication could help determine if these results hold true across tasks and populations.
REFERENCES


Appendix A

Research Recruitment Flyer
Research Participants Needed!

I am looking for individuals to participate in a study to examine productivity on a medical data entry task over time when performers are given individualized performance feedback. Participation will involve a computerized medical data entry task.

In order to be eligible for participation in this study, you must be available to attend 15-24, 30-minute sessions (in Wood Hall) over a 12-week period in the Fall 2012 semester. Additionally, you cannot have previously participated in other performance management studies conducted here in the Psychology department at Western Michigan University or held any positions involving data entry.

If you choose to participate in this study, you will receive monetary compensation that will equal $4.50 per half-hour session.

All participant information will remain completely confidential.

If you are interested in learning more about this study, please contact Nathan Bechtel at nathan.t.bechtel@wmich.edu or (269) 873-4610. Be sure to provide your name, e-mail address or telephone number, and the times you can be reached. Please remember that you must be available for 15-24, 30-minute sessions over a 12-week period during the Fall 2012 semester.

Contact Nathan Bechtel
Psychology Department
nathan.t.bechtel@wmich.edu
(269) 873-4610
Appendix B

Participant Recruitment Script
To be read aloud by the student investigator at undergraduate classes:

"Hi, my name is Nathan Bechtel. I am a graduate student in the Psychology Department and I am getting ready to start my master’s thesis research. I am visiting your class today to recruit participants for my study. To be a participant, you must be available to attend 15-24, 30-minute sessions spread across 12 weeks during the Fall 2012 semester. Additionally, you cannot have previously participated in other performance management studies conducted here in the Psychology department at Western Michigan University or held any positions involving data entry.

Participation will involve a medical transcription data entry task in which you will be asked to enter data on a computer for 15-24 sessions. Sessions will last approximately 30 minutes and you will be required to attend all sessions within a 12 week period. You will be able to take breaks during the session whenever you want. If you choose to participate in my study, you will receive monetary compensation that will equal $4.50 per half-hour session.

Your participation is completely voluntary and you may withdraw from the study at any time. If you do withdraw, you will be paid the money you have earned up to that point. Your willingness to participate in the study or your withdrawal from the study will not affect your grade in any course or your affiliation with Western Michigan University, and your participation will remain confidential.

If you are interested in learning more about my study, please list your contact information on the individual participant recruitment slips, which I will collect in a few minutes. You can also contact me at nathan.t.bechtel@wmich.edu or (269) 873-4610. Please remember that you must be available for 15-24, 30-minute during the Fall 2012 semester. I will contact you within the week to talk more about your potential participation. Thank you for your time."
Appendix C

HSIRB Approval
Date: August 17, 2012

To: Heather McGee, Principal Investigator
   Nathan Bechtel, Student Investigator for thesis

From: Amy Naugle, Ph.D., Chair

Re: HSIRB Project Number 12-07-16

This letter will serve as confirmation that your research project titled “The Effects of the Temporal Placement of Feedback on Performance and Skill Acquisition of a Medical Data Entry Task” has been approved under the expedited category of review by the Human Subjects Institutional Review Board. The conditions and duration of this approval are specified in the Policies of Western Michigan University. You may now begin to implement the research as described in the application.

Please note: This research may only be conducted exactly in the form it was approved. You must seek specific board approval for any changes in this project (e.g., you must request a post approval change to enroll subjects beyond the number stated in your application under “Number of subjects you want to complete the study”). Failure to obtain approval for changes will result in a protocol deviation. In addition, if there are any unanticipated adverse reactions or unanticipated events associated with the conduct of this research, you should immediately suspend the project and contact the Chair of the HSIRB for consultation.

Reapproval of the project is required if it extends beyond the termination date stated below.

The Board wishes you success in the pursuit of your research goals.

Approval Termination: August 17, 2013
Appendix D

Inclusion Questionnaire
Participant Number: ___________

Date: ________________________

1. Please check one:

Gender Male: _____ Female: _____

2. What is your age? ______

3. Have you participated in any performance management research within Western Michigan University’s psychology department in the past?  
   Yes: _____ No: _____

4. Have you held any position in the past which involved data entry?  
   Yes: _____ No: _____

4. Do you play any of the following computer games? (Circle one for each game)
   a. Solitaire Yes No
   b. Spider Solitaire Yes No
   c. Angry Birds Yes No
   d. Mahjong Yes No
   e. Bejeweled Yes No

5. On average, how many hours per month do you spend playing the aforementioned games? (Circle one)
   < 1  1  2  3  4  5  6  7  8  9  10 or more

6. Are you available to attend 15-24, 30-minute sessions (approximately 10 hours) over the Fall 2012 semester? (Circle one)
   Yes No Don’t know
Appendix E

Graph Comprehension Quiz
Answer the following questions based on the graph provided.

1. During which session did the participant get the highest number of correct responses? (Circle one)
   1 2 3 4 5 6

2. During which session did the participant get 28 correct responses? (Circle one)
   1 2 3 4 5 6

3. How many correct responses did the participant get in session 5? (Circle one)
   22 42 16 31 37 25
Appendix F

Informed Consent Form
Western Michigan University  
Psychology Department

**Principal Investigator:** Dr. Heather McGee, Ph.D.  
**Student Investigator:** Nathan T Bechtel  
**Title of Study:** The Effects of the Temporal Placement of Feedback on Performance and Skill Acquisition of a Medical Data Entry Task

You have been invited to participate in a research project titled *"The Effects of the Temporal Placement of Feedback on Performance and Skill Acquisition of a Medical Data Entry Task."* This project will serve as Nathan Bechtel’s thesis for the requirements of the I/O psychology master’s degree, under the supervision of Dr. Heather McGee, Ph.D.

This consent document will explain the purpose of this research project and will go over all of the time commitments, the procedures used in the study, and the risks and benefits of participating in this research project. Please read this consent form carefully and completely and please ask any questions if you need more clarification.

**What are we trying to find out in this study?**  
The purpose of this study is to examine the effects of feedback on performance levels and skill acquisition on a medical data entry task.

**Who can participate in this study?**  
Four exclusionary criteria will be used in determining appropriate participants. First, you must be available to attend 15-24, 30-minute sessions during the Fall 2012 semester. Second, you must play one of five computer games for at least 2 hours per month. Third, you cannot have participated in any past psychology department performance management studies or held any data entry jobs. Lastly, you must demonstrate a full comprehension of line graphs, via a brief quiz.

**Where will this study take place?**  
The study will be conducted in suite 1504, Wood Hall, rooms 1512 and 1532.

**What is the time commitment for participating in this study?**  
You must be available for a total of 15-24, 30-minute sessions during the Fall 2012 semester for a total of approximately 8-12 hours.

**What will you be asked to do if you choose to participate in this study?**  
You will be asked to perform a computer-based medical transcription data entry task designed to simulate the job of a medical data entry clerk. The computer program will
provide you with data corresponding to patients. You will first type the patient’s ID number into a box labeled “Patient ID,” and then, based on the information provided by the program, indicate whether the patient’s heart rate is inside or outside the normal range by clicking on the appropriate button. After you click the “Submit” button, information about another patient will be presented. Also, you will be asked not to talk to anyone about the features of this study.

What information is being measured during the study?
The computer program will automatically measure your performance on the medical data entry task.

What are the risks of participating in this study and how will these risks be minimized?
The nature of this task is one that requires little physical effort, and should not require any exertion greater than what you experience in your everyday activities. During sessions you may become tired or experience minor physical discomfort or stress. To minimize these risks, you may take breaks whenever you like during each session. During these breaks you may play one of several games available on the computer or just relax.

What are the benefits of participating in this study?
Data collected during this study may benefit the general scientific community by providing information on the effects of performance feedback. This research will add to our understanding of human performance and the working conditions that affect performance. The findings from this study may be applied to real-world work environments.

Are there any costs associated with participating in this study?
Aside from the time commitment of approximately 8-12 hours, there are no costs associated with participation in this study.

Is there any compensation for participating in this study?
For each 30-minute experimental session you will be compensated $4.50 for a total of $67.50 – $108.00 if you complete the entire study. You will receive the full amount in cash during the final debriefing session. If you decide to withdraw from the study, you will be paid the full amount earned up to the point of withdrawal.

Who will have access to the information collected during this study?
The principal investigator, the student investigator, and the research assistants will have access to the information collected during this study. At the beginning of the study you will be assigned a participant identification number so that your performance data can be tracked throughout the study, while your personal information remains confidential. Your identity will remain completely confidential.
What if you want to stop participating in this study?
You can choose to stop participating in this study at any time, for any reason, without penalty. You will not suffer any prejudice or penalty by your decision to stop your participation. You will receive no academic or personal consequences should you choose to withdraw from the study. The investigator can also decide to stop your participation in the study without your consent.

Should you have any questions prior to or during the study, you can contact the primary investigator, Dr. Heather McGee, Ph.D. at heather.mcgee@wmich.edu, or the student investigator at nathan.t.bechtel@wmich.edu. You may also contact the Chair, Human Subjects Institutional Review Board at 269-387-8293 or the Vice President for Research at 269-387-8298 if questions arise during the course of the study.

This consent document has been approved for use for one year by the Human Subjects Institutional Review Board (HSIRB) as indicated by the stamped date and signature of the board chair in the upper right corner. Do not participate in this study if the stamped date is older than one year.

I have read this informed consent document. The risks and benefits have been explained to me. I agree to take part in this study.

Please Print Your Name

Participant’s signature    Date
Appendix G

Sample Task Screen
Appendix H

Dependent Variable Calculation Formulae
Time on-task: Total Session Time (in seconds) – Time off-task (in seconds)

Rate of completion: Number of patient records completed per session

\[ \frac{\text{Number of patient records completed per session}}{\text{Time on-task (in minutes)}} \]

Accuracy (%): Number of correctly completed patient records per session

\[ \frac{\text{Number of correctly completed patient records per session}}{\text{Number of completed patient records per session}} \times 100 \]
Appendix I

Satisfaction Questionnaire
Feedback Satisfaction Survey

Participant Number: ______________________________
Date: ______________________________

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>strongly disagree</td>
<td>disagree</td>
<td>undecided</td>
<td>agree</td>
<td>strongly agree</td>
</tr>
</tbody>
</table>

1. I tried to alter my performance based on my feedback graphs
   - 1
   - 2
   - 3
   - 4
   - 5

2. I worked harder after seeing my feedback graph
   - 1
   - 2
   - 3
   - 4
   - 5

3. I enjoyed seeing my performance graph each session
   - 1
   - 2
   - 3
   - 4
   - 5

4. The sessions were not long enough for me to need a break
   - 1
   - 2
   - 3
   - 4
   - 5

5. I preferred to receive feedback at the beginning of the session
   - 1
   - 2
   - 3
   - 4
   - 5

6. I preferred to receive feedback at the end of the session
   - 1
   - 2
   - 3
   - 4
   - 5

7. I preferred receiving no feedback at all
   - 1
   - 2
   - 3
   - 4
   - 5
Appendix J

Sample Feedback Graph
Appendix K

Feedback Script
Pre-session Instructions:

FIP participants only:

Before the participant arrives, the research assistant or student investigator will print an individual feedback graph of the participant’s performance up to that point. When the participant arrives in suite 1504 Wood the research assistant or student investigator will greet him or her, direct the participant into room 1512 Wood and close the door for privacy purposes.

The research assistant or student investigator will then read the following aloud:

“Here is a graph of your performance for each session on the task thus far. Please take a moment to look it over. The graph will be updated each time you complete a session.”

The research assistant or student investigator will then give the participant a moment to look over the graph. Once the participant has had sufficient time to view the feedback, the research assistant or student investigator will read the following aloud:

All participants:

“Before we start today’s session, remember to leave your personal belongings, cell phone, iPod, etc., in this room and turn the cell phone to silent. I will lock the door when we leave to ensure no harm comes to them. Your task for today’s session will be a computerized data entry task. We are looking at your performance level throughout the study, so be sure to complete as many patient records as you can in the time allotted. There is a minimum performance of at least 70 completed slides in order for you to earn your payment for the session. If you get tired or feel stressed, feel free to take a break at any time. You can play the games available on the computer or just relax. You will not be penalized for taking a break. Please stay at your own workstation for the full session, and do not talk with the other participants. If you have any questions, I will be in the room across the hall throughout the session, so don’t hesitate to ask. Once the 30 minutes are up, I will take you back to get your belongings.”

The research assistant or student investigator will take the participant to their work station and prompt them to begin the task. The session time will begin when the “START” button is clicked. The computer will only allow 30 minutes of data entry, after which no responses will be allowed.

Post-session Instructions:

When the 30-minute session is over, the research assistant or student investigator will take the participant back to the initial meeting room (1512 Wood). The research assistant or student investigator will then follow one of two scripts, depending upon the experimental phase the participant is currently undergoing:
**FIA:**

The research assistant or student investigator will read the following script to the participant.

“Here is a graph of your performance for each session on the task thus far. Please take a moment to look it over. The graph will be updated each time you complete a session. This graph includes your performance from the session which you just completed.”

The research assistant or student investigator will then give the participant a moment to look over the graph. When the participant is finished, the research assistant or student investigator will let the participant know when his or her next 30-minute session is scheduled and retrieve his or her personal items. The participant is then free to leave.

**FIP & Baseline:**

The research assistant or student investigator will let the participant know when his or her next 30-minute session is scheduled and retrieve his or her personal items. The participant is then free to leave.

*Note – All participants who receive individual feedback will be allowed to keep the graph.*
Appendix L

Informed Consent and Inclusion Questionnaire Meeting Script
Informed Consent / Inclusion Questionnaire Meeting Script

Participants will meet with the research assistant or student investigator in 1512 Wood after being recruited. The research assistant or student investigator will give the potential participant an informed consent form and read the following script:

“This is the informed consent form for this study. It explains the research, risks and benefits, the necessary time commitment, compensation and your rights as a participant. Please listen carefully as I read through this document, and if you still wish to participate in the study, fill out the bottom portion.”

If the potential participant does not consent, they are free to leave. If the participant consents, the research assistant or student investigator will collect the completed informed consent form and give the student a copy of the inclusion questionnaire and graph comprehension quiz, and read the following script:

“Now we need to collect some information from you. This information will remain completely confidential. Please fill out this form (inclusion questionnaire) to the best of your ability. When you have completed that form, please complete this three-question quiz (graph comprehension quiz). When you have finished both of those you are free to leave, and we will contact you regarding the first session schedule within the next two weeks.”

Completed participant forms will be stored in a locked filing cabinet in 1532 Wood. Those participants who meet the inclusion criteria will be contacted within two weeks of this meeting and scheduled for their first two experimental sessions.
Appendix M

Compensation Receipt
Compensation for Participation

Date: ________________

Participant Name: ____________________

_______ Sessions X $4.50 per session = ________________
Appendix N

Debrief Script
Debrief Session Instructions

The following script is to be read aloud by the student investigator or research assistant to all participants during the final debriefing session:

"Thank you for participating in our study. The purpose of this last meeting is to give us the opportunity to explain the purpose of the study you have just participated in. You will be given the opportunity to ask any questions you may have after the explanation.

The purpose of this study was to evaluate the effects of feedback provided at different times relative to performance. All participants were subject to the same three conditions, in which feedback was provided before performance, after performance, or not at all. The only difference among participants was the order in which they were placed under each condition. All data related to performance was accurate and the graphs provided were correct depictions of your performance on the data entry task.

The secondary purpose of this study was to determine whether the timing of feedback delivery had an effect on the speed of skill acquisition. In order to determine this effect, three different phase orders were used. Since skill acquisition occurs fairly quickly on a task like this, only the first 5 sessions were used in answering this question.

The computer games were available for two reasons. First, we wanted to see if time allocated to the task vs. computer games would increase based on the feedback timing. Second, we wanted the environment to replicate that of a real data entry professional.

The research assistant or student investigator will now provide the participant with the satisfaction questionnaire.

We would like you to complete this brief questionnaire regarding your satisfaction with the study and the different forms of feedback provided during the sessions. Take as much time as you need to complete the questionnaire and put it in this folder when you are done.

After the participant has completed the satisfaction questionnaire, the research assistant or student investigator will provide the participant with the compensation receipt and the money they earned.

"You will now be compensated for your participation in this study. You completed (x number of) sessions throughout the study, each of which earned you $4.50. The total amount of money you have earned is printed on the compensation receipt.

Do you have any questions or concerns about this study or your participation at this time?

Please do not discuss this study with anyone else because we are still in the process of debriefing other participants. You are now free to leave.