Feedback After Good Trials Enhances Learning

Suzete Chiviacowsky and Gabriele Wulf

Recent studies (Chiviacowsky & Wulf, 2002, 2005) have shown that learners prefer to receive feedback after they believe they had a “good” rather than “poor” trial. The present study followed up on this finding and examined whether learning would benefit if individuals received feedback after good relative to poor trials. Participants practiced a task that required them to throw beanbags at a target with their nondominant arm. Vision was prevented during and after the throws. All participants received knowledge of results (KR) on three trials in each 6-trial block. While one group (KR good) received KR for the three most effective trials in each block, another (KR poor) received feedback for the three least effective trials in each block. There were no group differences in practice. However, the KR good group showed learning advantages on a delayed retention test (without KR). These results demonstrated that learning is facilitated if feedback is provided after good rather than poor trials. The findings are interpreted as evidence for a motivational function of feedback.
where KR, presented for individual trials or as an average, respectively, is delayed until a set of trials has been completed (e.g., Schmidt, Young, Swinnen, & Shapiro, 1989; Wulf & Schmidt, 1996; Young & Schmidt, 1992).

Yet, there are also findings inconsistent with the guidance view. For example, frequent feedback does not consistently lead to more effective performance during practice than less frequent feedback (e.g., Nicholson & Schmidt, 1991; Winston & Schmidt, 1990; Wulf, Lee, & Schmidt, 1994; Wulf, Schmidt, & Deubel, 1993). Furthermore, learning complex skills does not necessarily suffer from frequent feedback (e.g., Wulf, Shea, & Matschiner, 1998; Swinnen, Lee, Verschueren, Serrien, & Bogaerts, 1997), and the guidance hypothesis cannot explain interactions of feedback frequency and the type of attentional focus (Wulf, McConnel, Gärtner, & Schwarz, 2002). In their recent review of the feedback literature, Wulf and Shea (2004) concluded that, although the guidance hypothesis contributed to a better understanding of how feedback affects performance and learning, future research needs to examine how feedback interacts with other factors (e.g., task complexity, level of expertise, focus of attention) to influence learning.

Recent studies investigating the effects of self-controlled feedback suggest that another factor might have to be considered when determining the effectiveness of augmented feedback, namely the accuracy of the movements for which feedback is provided (Chiviacowsky & Wulf, 2002, 2005). While giving learners the opportunity to decide when to receive feedback (i.e., self-controlled feedback) has generally enhanced their learning compared to not having this opportunity (e.g., yoked condition; e.g., Chiviacowsky & Wulf, 2002, 2005; Janelle, Barba, Frehlich, Tennant, & Gaurgauh, 1997; Janelle, Kim, & Singer, 1995), the study by Chiviacowsky and Wulf (2002) revealed another interesting finding: learners preferred to receive feedback after they thought they had a relatively successful trial but not when they thought their performance was relatively poor. This was evident from postexperimential interviews of self-controlled learners. Furthermore, interviews of yoked learners showed that they would also have preferred feedback after good trials but not after poor trials. Of course, for them feedback was distributed randomly (i.e., provided independently of their performance on the respective trial). In contrast, self-controlled learners had, on average, smaller errors on those trials when they requested feedback relative to when they did not ask for feedback (Chiviacowsky & Wulf, 2002, 2005). That is, they asked for feedback predominantly after good trials. This suggests that self-controlled feedback might be more effective, because it is more in accordance with the performer’s needs, or preferences, than externally controlled feedback (yoked condition); it might also suggest feedback is more effective if presented after good trials (independent of whether feedback is self-controlled).

This hypothesis seems contrasts with the guidance view (e.g., Salmoni, Schmidt, & Walter, 1984; Schmidt, 1991), though. According to that view, feedback would be expected to be particularly important after poor trials, or relatively large errors, when it is assumed to guide the learner to the correct movement. After good trials, or small errors, feedback is viewed as being less important. In fact, procedures such as bandwidth feedback (e.g., Lai & Shea, 1999; Lee & Carnahan, 1990; Sherwood, 1988) are designed to provide learners with feedback when errors exceed a certain bandwidth of “tolerable” error. Conversely, no quantitative feedback is provided on relatively good trials, that is, when errors fall within the specified bandwidth (thereby indicating the movement was essentially correct). The finding that learners preferred to receive feedback after good trials (Chiviacowsky & Wulf, 2002, 2005) seems to be at odds with the view that feedback is more effective after large errors. Yet, if learners are able to differentiate between “good” and “bad” trials, as shown in previous studies (Chiviacowsky & Wulf, 2002, 2005), feedback informing them of poor performance might, in fact, be more or less redundant. In contrast, feedback after a good trial could confirm that the movement is correct or help fine tune the movement. This information might be equally or more important than error feedback. In addition, it might be more motivating for learners to receive “positive” rather than “negative” feedback, which, in turn, could lead to more effective learning.

The purpose of the present study was, therefore, to determine whether feedback is more effective when provided after relatively good or relatively poor trials. If learning benefits more from feedback after successful trials, this would pose additional difficulties for the guidance view of feedback. In the present study, participants practiced a motor task (tossing beanbags to a target) and received feedback on three of the six trials after completing each six-trial block (i.e., 50% feedback). While one group received KR about the accuracy of the three best throws in each block, another received KR about the three poorest throws. The effectiveness of feedback after good versus poor practice trials was assessed in a retention test without KR 1 day after practice.

Method

Participants

Twenty-four undergraduate students (6 men, 18 women) with a mean age of 21.1 years participated in this experiment. All participants provided informed consent. They had no prior experience with the experimental task and were not aware of our specific study purpose.
Apparatus and Task

The task required participants to toss beanbags (100 g) at a target placed on the floor, using the nondominant arm. The target was circular, had a radius of 10 cm, and was placed 3 m from the participant. Concentric circles with radii of 20, 30, 40, 50, 60, 70, 80, 90, and 100 cm were drawn around the target. These served as zones to assess the accuracy of the throws. If the beanbag landed on the target, 100 points were awarded. If it landed in one of the other zones, or outside the circles, 90, 80, 70, 60, 50, 40, 30, 20, 10, or 0 points, respectively, were recorded.

Procedure

Participants were quasirandomly assigned to the “KR good” and “KR poor” groups, with 9 women and 3 men in each group. After each block of six trials, participants in the KR good group received KR on the three best (i.e., most accurate) tosses in that block, whereas those in the KR poor group received KR on the three poorest tosses. Participants in both groups were informed that, at the end of each block of six trials, they would receive KR on three of those trials. However, they did not know for which trials they would receive KR. Participants were allowed to look at the target before each set of six trials. During those six trials, participants wore opaque swimming goggles to prevent them from viewing the results. A digital chronometer was used to control the timing of the trials and KR presentation. Participants had 6 s to complete a trial. KR was written on a board and presented for 15 s. It consisted of the trial number and respective score, as well as directional information. That is, if the participant overshot the target, a plus sign was added to the accuracy score (e.g., +60), if the participant undershot the target, a minus sign was added (e.g., -90). Thus, KR not only provided information about the extent of the deviation from the target but also about whether the toss was short or long. All participants performed 60 trials during the practice phase, and 1 day after practice they performed a retention test consisting of 10 trials without KR.

Data Analysis

Accuracy scores were analyzed in a 2 (group) x 10 (blocks of 6 trials) analysis of variance (ANOVA), with repeated measures on the last factor for the practice phase, and in a one-way ANOVA for the retention test. In addition, to determine whether scores on KR trials were actually higher for the KR good group, compared to the KR poor group, we analyzed the scores separately for KR and no-KR trials in a 2 (groups: KR good vs. KR poor) x 2 (trial type: KR vs. no KR) x 10 (blocks of 6 practice trials). Furthermore, we wanted to see whether participants were able to differentiate between good and poor trials, based on their intrinsic feedback. Therefore, we calculated the change in the accuracy score from trial to trial. If individuals have a feel for how they performed on a given trial, one would expect to see less change after good trials (i.e., the three best trials in each block), compared to poor trials (i.e., the three worst trials in each block), in both groups. The average change scores for the first and second half of practice were analyzed in a 2 (group) x 2 (trial type: good vs. poor) x 2 (phase: first vs. second half) ANOVA, with repeated measures on the last two factors. Finally, we wanted to ensure that the distribution of KR trials within each block was comparable between groups and that group differences, if any, could not be attributed, for example, to the fact that one group received KR on later trials for which it might be easier to associate the intrinsic feedback with KR. Therefore, we determined the trials for which each participant would receive KR within each six-trial practice block. We then calculated for each group the three trials in each block for which most participants received or didn’t receive KR, respectively.

Results

Practice

Accuracy Scores. The KR poor group tended to have somewhat lower scores than the KR good group early in practice, but both groups increased their scores and showed similar performances toward the end of practice (see Figure 1). Neither the main effect of group, $F(1, 22) < 1$, nor the Group x Block interaction, $F(9, 198) = 1.02$, $p > .05$, were significant. Only the main effect of block was significant, with $F(9, 198) = 18.76$, $p < .001$, $\eta^2 = .46$.

Accuracy Scores on KR Versus No-KR Trials. An analysis of the accuracy scores on practice trials with KR versus trials without KR revealed that scores on KR trials were clearly higher for the KR good group than the KR poor group, while the opposite was true for the no-KR trials (see Figure 2). The interaction of group and trial type was significant, $F(1, 22) = 1137.69$, $p < .001$. This confirmed that higher scores were reported to the KR good group than the KR poor group.

Change Scores. Both groups showed a greater change in the accuracy score from trial to trial after poor trials, compared to good trials (see Figure 3). Also, the changes were generally smaller in the second half than in the first half of the practice phase. The main effects of trial
type, $F(1, 22) = 17.20, p < .001, \eta^2 = .44$, and phase, $F(1, 22) = 4.45, p < .05, \eta^2 = .17$, were significant. There were no significant differences between groups, $F(1, 22) < 1$, and no interaction effects.

**KR Trials Within Blocks.** Finally, we determined on which trials, within the six-trials blocks, the KR good and KR poor groups received KR. This was done to ascertain whether the delay between the trials for which KR was provided and the actual KR was greater for one group than another. During the first half of the practice phase, the KR good group received more KR on Trials 1, 2, and 3 in each six-trial block than on the remaining trials, while the KR bad group received more KR on the Trials 2, 4, and 5 (see Table 1). In the second half of practice, the KR good group received more KR on Trials 1, 2, and 6, while the KR poor group received more KR on Trials 2, 3, and 6, compared to the remaining trials. Thus, whereas the KR good group had more trials between the trials selected for KR and the KR itself early in practice, compared to the KR poor group, this difference was attenuated in the second half of practice.

**Retention**

On the retention test without KR, performed 1 day after the practice phase, the KR good group clearly had higher accuracy scores than the KR poor group (see Figure 1). This group difference was significant, $F(1,$
Thus, providing KR after the most effective trials during practice resulted in superior learning.

Discussion

The present study examined whether augmented feedback would be more effective for learning if given after relatively good or after relatively poor trials. According to the guidance hypothesis (e.g., Salmoni et al., 1984; Schmidt, 1991), feedback should be more beneficial if presented after larger rather than smaller errors. Yet, learners appeared to prefer and, if given the opportunity, even select feedback mainly after good trials rather than poor trials (Chiviacowsky & Wulf, 2002, 2005), suggesting that feedback after successful trials might actually be more advantageous for learning.

The present results indeed showed a learning advantage if feedback was presented after trials with relatively small errors, or high accuracy scores (KR good group), compared to trials with relatively large errors, or low accuracy scores (KR poor group). That is, although both groups received KR on 50% of the practice trials, retention performance was enhanced if they received feedback on the more accurate 50%. This finding is in line with the results of Chiviacowsky and Wulf (2002, 2005). Their results showed that learners who controlled the feedback schedule not only preferred feedback after good trials but actually requested it more often after good trials than poor trials. Importantly, learning was also enhanced compared to yoked participants. An interesting question, therefore, is: Why does feedback after good trials benefit learning?

It does not appear that the position of KR trials within the six-trial practice blocks played a significant role in the differential learning effects of the KR good and KR poor conditions. The position of KR trials did not differ much between groups. In fact, the KR poor group tended to receive KR on trials that occurred relatively later in each block, compared to the KR good group. Thus, one might argue the KR poor group had the advantage that the KR trials were relatively fresh in their

![Figure 3. Changes in accuracy scores after good versus poor trials for the KR good (receiving feedback for the three most effective trials in each block) and KR poor (receiving feedback for the three least effective trials in each block) groups during the first and second half of practice.](image)

<table>
<thead>
<tr>
<th>Group</th>
<th>First half Trials</th>
<th>Second half Trials</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>KR good</td>
<td>49.3</td>
<td>45.7</td>
</tr>
<tr>
<td>KR poor</td>
<td>43.5</td>
<td>28.0</td>
</tr>
</tbody>
</table>

*Note. KR good = knowledge of results provided for the three most effective trials in each block; KR poor = group receiving feedback for the three least effective trials in each block; numbers in boldface type represent trials, for which, on average, more KR was provided.*
memory when KR was eventually provided. Yet, the KR good group outperformed this group in retention.

One might assume feedback after relatively successful trials would encourage learners to repeat a (successful) movement rather than change the movement pattern to correct for errors. In fact, “maladaptive short-term corrections” (e.g., Schmidt, 1991) are seen as a negative effect of frequent feedback. That is, a performer’s attempts to correct even small response errors (that may simply represent an inherent variability in the motor system) are viewed as resulting in unproductive response variability and preventing learners from developing a stable movement representation. Yet, our analysis of the trial-to-trial changes did not support the view that the learning benefits of the KR good condition were due to reduced response variability. Both groups showed changes from trial-to-trial, and both demonstrated, on average, less change after good trials than poor trials. As KR was only provided after every six trials, participants must have made trial-to-trial changes (or no changes) based on their intrinsic feedback. That is, learners in both groups seemed to have a feel for how well they performed on a given trial (similar to Chiviacowsky and Wulf, 2002, 2005).

A study by Cauraugh, Chen, and Radlo (1993) using bandwidth-KR manipulations also showed that quantitative KR about relatively “good” or “poor” trials did not affect the amount of change in performance from trial to trial. Bandwidth KR (quantitative KR outside the bandwidth) and reversed-bandwidth (quantitative KR inside the bandwidth) conditions resulted in greater performance changes outside the bandwidth, (i.e., when participants knew their performance was relatively poor). Thus, trial-to-trial changes in performance appear to be less a function of whether KR is given after relatively good or poor trials; rather, they seem to depend more on the performer’s perception of his or her performance. Interestingly, a recent study showed that a brain region along the cingulate sulcus demonstrates a shift in activation as a function of learning, such that early in the learning process it is activated by extrinsic feedback, while later in learning it responds to intrinsic feedback (Mars et al., 2005). Thus, this region, which appears to be involved in error detection and correction, is activated by the earliest source of error information available.

It appears participants in the present study developed an error-detection-and-correction capability, as evidenced by the smaller changes after good trials and reduced changes during the second half of practice (when performance was more accurate). But this capability was developed similarly under KR good and KR poor conditions.

The most viable explanation for the benefits of receiving KR after good trials (while poor trials are essentially ignored) might be that it creates a greater success experience for learners than KR after relatively poor trials (with good trials being ignored). This success experience might be more motivating for learners and, in turn, enhance the learning process. Aside from its informational role, KR has long been assumed to have motivational properties as well (e.g., Thorndike, 1927). More recent studies have confirmed this. A study by West, Bagwell, and Dark-Freudeman (2005), for example, showed greater performance gains when participants set goals and received positive feedback, compared to a control condition. Positive feedback has been shown to encourage participants to raise their goals (Flies & Judge, 2005) and expectations for future performance (Singer & McCaughan, 1978). Interestingly, certain brain areas, including the rostral anterior cingulate cortex, posterior cingulate cortex, right superior frontal gyrus, and striatum, have been found to respond more strongly to positive feedback (Nieuwenhuis, Slater, Alting von Geusau, Heslenfeld, & Holroyd, 2005). Participants performing a time estimation task showed greater activation in those brain areas when they received random positive feedback. Although learners in the present study were not informed on which trials they would receive feedback, the KR good group clearly had smaller errors than the KR poor group when they received feedback. Because learners appear to be relatively sensitive to how well they perform—as shown by the trial-to-trial change scores in the present study and the fact that learners chose KR after relatively good trials in an earlier study (Chiviacowsky & Wulf, 2002)—it is possible learners noticed a relationship between their performance and feedback. Postexperimental interviews in future studies could determine if, and to what extent, learners become aware of such a relationship. Given the behavioral and neurophysiological changes observed for positive relative to negative (or neutral) feedback, it appears the learning advantages of the KR good condition in the present study might be largely due to motivational factors.

In future experiments, it might be interesting to examine the generalizability of the present findings to different tasks. For example, would benefits of KR after good trials also be found for tasks in which intrinsic feedback is less easily interpreted than in the present task? Although participants performed the beanbag-toss task in the present study using the nondominant arm, adult participants can presumably use a wealth of previous experiences with similar tasks to judge the movement outcome. This might be different for other tasks for which participants lack experience. Along the same line, does the “location” of the responses relative to the task goal play a role for the effectiveness of KR provided after good or poor trials? This might be an issue, for instance, if all responses showed large deviations from the target (e.g., early in learning a difficult task), or if all performances were very close to the target (e.g., late in learning a simple task).
The learning advantages of feedback after good relative to poor trials do not seem to be in line with the guidance view of feedback (e.g., Salmoni et al., 1984; Schmidt, 1991). One might argue they contradict a strict interpretation of the view that error feedback is beneficial because it guides the learner to the correct response. If this were the case, feedback after larger errors should be more beneficial than feedback after smaller errors. The guidance view clearly focuses on the informational properties of feedback, and the motivational effects of feedback seem to have been downplayed somewhat in recent years. For example, Schmidt and Lee (2005) stated, with respect to the motivational effects of feedback, such as keeping learners alert, causing them to set higher goals, or making practice more enjoyable: "Most of these effects are probably performance phenomena, which can be expected to subside when the feedback is withdrawn after training." They also acknowledged there might be indirect learning effects, such as encouraging individuals to practice more often or longer. Yet, newer findings—including the learning benefits of positive feedback seen in self-controlled feedback studies as well as in the present study, and neurophysiological effects of positive feedback (Nieuwenhuis et al., 2005)—suggested the motivational properties of feedback have a direct effect on learning. It might be a fruitful endeavor for future studies to examine the motivational role of feedback more directly (e.g., by using questionnaires), so that we come to a more complete understanding of the various roles feedback plays in the learning process.

References


Notes

1. An analysis of the first practice trial indicated that there was no significant difference between the KR good (55.0) and KR poor (51.7) groups, $F(1, 22) < 1$. (The accuracy scores on the first trial were relatively high, compared to the average accuracy for the first block. In fact, accuracy generally decreased across trial blocks. The reason is presumably that participants were allowed to look at the target before the first trial of each block, but were not allowed to view it for the remainder of that block.)

2. We thank David Wright and an anonymous reviewer for these suggestions.

Authors’ Note

Please address all correspondence concerning this article to Suzete Chiviacowsky, Escola Superior de Educação Física, Universidade Federal de Pelotas, Rua Luís de Camões, 625 - CEP 96055-630, Pelotas - RS – Brazil.

E-mail: schivi@terra.com.br