

Information Presentation in Health Apps and Devices: The Effect of Color, Distance to Goal, Weight Perception, and Interest on Users' Self-Efficacy for Accomplishing Goals

Elizabeth Eikey, Penn State University
Erika Poole, Penn State University
Madhu Reddy, Penn State University

Abstract

Colors elicit different emotions and can impact behavior. However, few studies have examined the effects of color in health apps and devices. While these devices are becoming increasingly popular, little research has examined how the presentation of feedback within these technologies impacts users' beliefs, such as self-efficacy. Yet understanding how information presentation affects users' self-efficacy is important as self-efficacy is linked to successful behavior change. This study explores how manipulating the color of an exercise progress bar within a simulated health device influences users' self-efficacy for completing today's exercise goal and future exercise goals. This preliminary study aims to better understand how color choice, distance to goal, users' weight perception, and interest in completing an exercise goal affect users' self-efficacy by simulating health app exercise progress bars. By understanding what influences self-efficacy, we can design better health apps and devices to increase the likelihood that users will reach their goals.

Keywords: information presentation; health; color; self-efficacy; design

Citation: Eikey, E., Poole, E., Reddy, M. (2015). Information Presentation in Health Apps and Devices: The Effect of Color, Distance to Goal, Weight Perception, and Interest on Users' Self-Efficacy for Accomplishing Goals. In *iConference 2015 Proceedings*.

Copyright: Copyright is held by the authors.

Contact: exe145@ist.psu.edu, epoole@ist.psu.edu, mreddy@ist.psu.edu

1 Introduction

The importance of color choice is all around us. We use color to elicit all sorts of emotions and impact behavior. However, little research has examined the effects of color in the context of health apps and devices. This research explores how the presentation of information, specifically color choice for exercise progress bars, in health apps and devices influences users' beliefs in their abilities to reach their health-related goals. Use of health and fitness apps and ubiquitous health devices are becoming increasingly popular. Often these health apps and devices help people reach their health-related goals, such as losing weight, eating healthier, and exercising more. In order to allow users to gauge their progress, these apps and devices often show feedback, in the form of graphs and other visualizations for instance, that help users see how close they are to reaching their milestones. However, few studies have looked at how the presentation of this information affects users. Understanding how color influences users' beliefs is a key to promoting successful health behaviors, such as exercise.

How does color choice impact users' beliefs in their abilities to accomplish goals? Are there ways we can manipulate colors in order to promote users' belief in their abilities to reach their goals? What other things influence users' self-efficacy? In order to explore these questions and using the work of Choe, Lee, Munson, Pratt, & Kientz (2013) as a model, we conducted an online experiment and survey to explore how manipulating color of an exercise progress bar and distance to goal influences users' self-efficacy for completing 10,000 steps today and in the future. In this preliminary study, we aim to better understand how color choice, distance to goal, weight perception, and interest in completing an exercise goal affect users' self-efficacy by simulating health app exercise progress bars. The findings of this study can be leveraged to design health apps and devices that promote users' self-efficacy and thus increase their chances of successfully achieving their health and fitness goals.

2 Related Work

2.1 Self-Efficacy Theory

Users' self-efficacy is important to consider as it is related to their actual behaviors (Bandura, 1977). The more likely they are to believe in their ability to reach their health goals, the more likely they are to attempt tasks and persist in those efforts longer (Schunk, 1990). For many years, researchers have used self-

efficacy theory to understand smoking cessation, alcohol reduction, weight control, and exercise (Strecher, Devellis, Becker, & Rosenstock, 1986). Self-efficacy is a key factor in successful weight loss and maintenance (Edell, Edington, Herd, Brien, & Witkin, 1987; Elfhag & Rössner, 2005; Strecher et al., 1986). Studies show a positive relationship between self-efficacy and successful health behavior change (Anderson-Bill, Winett, & Wojcik, 2011; Strecher et al., 1986). For instance, Manley et al. (2014) found some positive correlations between self-efficacy and body mass index in middle school children. Additionally, (Gao, Xiang, Lochbaum, & Guan, 2013) found that self-efficacy mediated the effects of fitness performance also among middle school students.

Despite this, little research has explored self-efficacy's role in health technologies. (Anderson-Bill et al., 2011) found self-efficacy (as well as social support) is linked with successful internet exercise and nutrition interventions. However, their focus was not on design. We must pay close attention to how we may impact users' self-efficacy through our design choices. Choe et al. (2013) (the research we used to model this study) considered the effects of design through framing. They found using achievement framing positively impacted individuals' perceptions of performance capabilities. Thus, it is important to understand the effects of design in order to promote self-efficacy and encourage healthy behaviors.

2.2 Color Research

Previous studies have looked at the effects of color on emotions, perceptions, performance, and behaviors both in physical environments and online. Different color choices can elicit different emotions. For example, Valdez & Mehrabian (1994) found that blue, blue-green, green, red-purple, purple, and purple-blue elicited pleasure and green-yellow, blue-green, and green elicited arousal. Similar responses have been seen when designing webpages. In good web design, blue colors are often seen as the most aesthetically pleasing hues (Hall & Hanna, 2004; Kondratova & Goldfarb, 2007).

The emotions linked with these colors can also impact behavior and performance. For instance, aesthetically pleasing websites (i.e. using blue hues) are correlated with increased likelihood to purchase products (Moore, Stammerjohan, & Coulter, 2005). Even in store design itself, using certain colors can impact consumer behavior. When comparing blue store design to red store design, Bellizzi & Crowley (1983) found that blue store design was associated with more positive retail outcomes, whereas red store design was associated with more negative retail outcomes. Similarly, other researchers have explored the effects of color on achievement and performance, such as in an academic context. For example, Elliot, Maier, Moller, Friedman, & Meinhardt (2007) ran a series of experiments to test the effects of red on test performance. They found using red (compared to green, black, grey, and white) caused avoidance motivation and negatively impacted scores on IQ and anagram tests.

Studying these effects in relation to exercise and fitness goals within health apps and devices is crucial since the impact of color choice can vary from context to context (Gnamb, Appel, & Batinic, 2010). For instance, while red can negatively affect performance in some instances (Elliot et al., 2007), red can have positive effects in the appropriate context, such as increasing appetite in restaurants (Singh, 2006). Additionally, sports teams who wear red tend to outperform their competitors (Gnamb et al., 2010). In spite of the numerous studies examining the impact of color, little research has explored these effects in the context of health behaviors.

3 Methods

3.1 Experiment Design

The goal was to explore how presenting information differently affects individuals' self-efficacy. We chose to use a between-subjects design. Since participants have to report their self-efficacy ratings, we felt that having the same participants test each condition could make them sensitive to the true aims of the study. Also, we were concerned about order effects of presenting the different progress bars and how that might influence their self-efficacy ratings. We looked at the effects of *valence* by altering the color of an exercise progress bar and the *distance to goal/number of steps completed* (2500, 5000, 7500 steps out of a goal of 10,000 steps). *Valence* refers to the emotion categorization of the progress bar colors (neutral, positive, negative). For the positive valence condition, the exercise progress bar was colored green, to convey negative valence, it was colored red, and for the neutral condition, it was colored black. This experiment was a 3 (valence: neutral, positive, negative) x 3 (steps completed: 2500, 5000, 7500) design to total nine conditions. The neutral valence + 2500, 5000, or 7500 acted as the control condition(s). In addition to asking users' interest in completing a health goal like in the work of Choe et al., (2013), we also asked questions related to their views on their own weight. Users' perceptions of their weight may influence how easily attainable they believe their goal is and thus may impact their self-efficacy ratings.

We chose to examine the effects of red and green. The color red is often related to mistakes (i.e. school teacher's corrections) or to danger or stop (i.e. red traffic lights and signs) (Gnambs et al., 2010), and green is often associated with go (i.e. green traffic lights and signs) in western societies. Additionally, these colors are commonly used to show progress in health apps, and these colors were also used in the work of Choe et al. (2013). Therefore, we felt these were appropriate choices to represent "negative" and "positive" associations with exercise progress.

3.2 Hypotheses

H1: Effect of Valence on Self-Efficacy

- H1a: Those presented with positive valence feedback (green progress bar) will be more likely to report higher ratings of self-efficacy for completing today's step goal than those in the neutral valence condition (black progress bar).
- H1b: Those presented with positive valence feedback (green progress bar) will be more likely to report higher ratings of self-efficacy for completing tomorrow's step goal than those in the neutral valence condition (black progress bar).
- H1c: Those presented with negative valence feedback (red progress bar) will be more likely to report lower ratings of self-efficacy for completing today's step goal than those in the neutral valence condition (black progress bar).
- H1d: Those presented with negative valence feedback (red progress bar) will be more likely to report lower ratings of self-efficacy for completing tomorrow's step goal than those in the neutral valence condition (black progress bar).

H2: Effect of Distance to Goal/Steps Completed on Self-Efficacy

- H2a: Those who are closer to their goal (7500 steps completed) will be more likely to report higher ratings of self-efficacy for completing today's step goal than those who farther from their goal (2500 or 5000 steps completed).
- H2b: Those who are closer to their goal (7500 steps completed) will be more likely to report higher ratings of self-efficacy for completing tomorrow's step goal than those who farther from their goal (2500 or 5000 steps completed).
- H2c: Those who are far from their goal (2500 steps completed) will be more likely to report lower ratings of self-efficacy for completing today's step goal than those who closer to their goal (5000 or 7500 steps completed).
- H2d: Those who are far from their goal (2500 steps completed) will be more likely to report lower ratings of self-efficacy for completing tomorrow's step goal than those who closer to their goal (5000 or 7500 steps completed).

H3: Effect of Weight Perception on Self-Efficacy

- H3a: Those who are more likely to report a healthy weight perception will be more likely to report high levels of self-efficacy for reaching today's step goal.
- H3b: Those who are more likely to report a healthy weight perception will be more likely to report high levels of self-efficacy for reaching tomorrow's step goal.
- H3c: Those who are more likely to report an overweight perception will be more likely to report low levels of self-efficacy for reaching today's step goal.
- H3d: Those who are more likely to report an overweight perception will be more likely to report low levels of self-efficacy for reaching tomorrow's step goal.
- H3e: Those who are more likely to report an underweight perception will be more likely to report low levels of self-efficacy for reaching today's step goal.
- H3f: Those who are more likely to report an underweight perception will be more likely to report low levels of self-efficacy for reaching tomorrow's step goal.

H4: Effect of Interest on Self-Efficacy

- H4a: Those with a higher interest in completing a daily step goal will be more likely to report higher levels of self-efficacy for reaching today's step goal than those with low interest.
- H4b: Those with a higher interest in completing a daily step goal will be more likely to report higher levels of self-efficacy for reaching tomorrow's step goal than those with low interest.




Valence	Steps Completed	# of Participants Included in the Analysis	Example Feedback
Neutral	2500	30	
	5000	30	
	7500	32	
Positive	2500	31	
	5000	30	
	7500	35	
Negative	2500	31	
	5000	26	
	7500	28	
Total:		273	

Table 1. Feedback manipulation, example feedback, and number of participants

3.3 Survey Design and Study Conditions

Through Qualtrics, an online survey software that allows the design and distribution of online questionnaires, we created an online experiment and questionnaire with nine conditions. Each questionnaire consisted of five portions: (1) familiarity with pedometers and health devices, (2) interest in completing 10,000 steps per day, (3) self-efficacy measures, (4) check/filter questions, and (5) demographic information for a total of 19 questions per participant. As part of the demographic information, we asked participants to rate the following statement on a 7-point Likert-like scale (1-Strongly disagree to 7-Strongly agree): “I believe I am... underweight, at a healthy weight, and overweight.”

At the beginning of the survey, we first asked questions related to pedometer and health device usage, such as “How often do you use a pedometer (step counter) or other device to track the number of steps you take?” In order to allow participants to better gauge how much effort is required to complete 10,000 steps, we also asked, “Approximately how far do you think 10,000 steps is?” and then provided the answer (5 miles) on the next page. Then we asked, “How interested are you in completing 10,000 steps per day to maintain a desirable level of physical activity for health?” and asked them to rate it on a 7-point Likert-like scale (1-Not at all interested to 7-Very interested) (similar to Choe et al., 2013). Then we provided participants with a short scenario from the work of Choe et al. (2013):

“Studies have suggested taking 10,000 steps per day to maintain a desirable level of physical activity for health. Suppose you purchase a pedometer (step counter) to monitor your step count, and you set a daily goal of 10,000 steps. You need to wear it every day in your pocket or on your waist, and it gives you real-time feedback of your step count.”

Afterward, we presented one of the nine step count feedback. We manipulated the feedback by changing the:

- *Valence:* We presented feedback neutrally with a black progress bar (control condition), positively with a green progress bar, or negatively with a red progress bar.
- *Distance to goal/number of steps completed:* We presented the number of steps completed as 2500, 5000, or 7500.

Examples of feedback we showed participants are in Table 1. After presenting the feedback, we measured participants’ self-efficacy by asking the following two questions adapted from Bandura (1977) as cited in the work of Choe et al. (2013): (1) “Rate how confident you are that you can achieve your daily goal as of now (assume it is 4:30pm, weekday)” and (2) “Rate how confident you are that you can walk 10,000 steps tomorrow (assume it is a weekday).” We measured self-efficacy on a 7-point Likert-like scale (1-Certain I cannot meet my goal to 7-Certain I can meet my goal).

We included two filtering questions at the end of the survey. Since it was important that participants could adequately understand the feedback, we presented each participant with feedback that showed they had completed just over 2500 steps and asked them to correctly describe how much of their

goal remains by selecting one of three multiple choice answers: “Less than 50% of my goal remains,” “More than 50% of my goal remains,” and “None of the above.” In order to filter out participants who did not read the survey questions closely, we also asked them to answer a multiple choice reading question: “Approximately how far do you think 10,000 steps is?”

Once we collected all of our data, we ran a one-way analysis of variance (ANOVA) to test the effects of *valence* and *distance to goal* on self-efficacy. Then we ran linear regressions to explore how *weight perception* and *interest* impact self-efficacy. In order to control for *interest* and *weight perception*, we ran an analysis of covariance (ANCOVA).

3.4 Participants and Recruitment

In total, we recruited 379 participants from the U.S. through Amazon Mechanical Turk. Researchers have found that using Mechanical Turk often allows for a more demographically diverse population than other internet and college samples and the data obtained is at least as reliable as other methods (Buhrmester, Kwang, & Gosling, 2011). Each participant was compensated \$0.12 for approximately four minutes of his/her time. Participants were randomly assigned to one of the nine conditions. We excluded a total of 106 participants based on the following exclusion criteria:

- Feedback question wrong: 2
- Reading question wrong: 81
- Both feedback and reading questions wrong: 3
- Did not complete experiment: 20

Of the 273 participants used in the analyses, 42% were male and 58% were female. The majority (42%) were ages 25-34 years, followed by 24% 18-24 years, 19% 35-44 years, 8% 45-54 years, and 7% 55+ years. Eighty five percent (85%) reported they had used a pedometer or similar device in the past, and 25% currently use one. Participants reported various education levels and marital statuses. Participants' self-reported race and ethnicity were consistent with U.S. census data.

4 Results

		Self-Efficacy Today	Self-Efficacy Tomorrow
H1: Effect of <i>valence</i>		Not significant ($p > .05$)	Not significant ($p > .05$)
H2: Effect of <i>distance to goal/steps completed</i>		Significant ($p < .000$)	Not significant ($p > .05$)
H3: Effect of <i>weight perception</i>	Healthy weight	Significant ($p = .023$)	Significant ($p = .008$)
	Overweight	Not significant ($p > .05$)	Significant ($p = .026$)
	Underweight	Not significant ($p > .05$)	Not significant ($p > .05$)
H4: Effect of <i>interest</i>		Significant ($p < .000$)	Significant ($p < .000$)

Table 2. Summary of findings for hypotheses 1-4

4.1 H1: Effect of Valence on Self-Efficacy

For the effect of *valence* on users' self-efficacy rating for reaching 10,000 steps today, there was no statistically significant difference between groups as determined by one-way analysis of variance (ANOVA) ($F(2, 270) = 1.100, p = .334$) (see Table 2). We found no support for H1a or H1c. Similarly, there was no statistically significant difference between groups for self-efficacy ratings for completing tomorrow's goal as determined by one-way ANOVA ($F(2, 270) = .557, p = .573$). As shown in Table 2, found no support for H1b or H1d. Therefore, the color of the progress bar (red, green, black) was not a significant predictor of self-efficacy ratings.

4.2 H2: Effect of Distance to Goal/Steps Completed on Self-Efficacy

For the effect of *steps completed* on users' self-efficacy rating for reaching 10,000 steps today, there was a statistically significant difference between groups as determined by one-way ANOVA ($F(2, 270) = 20.038, p < .000$) (see Table 2). A Tukey post-hoc test revealed that each condition was significantly different from the other. The closer users were to their 10,000 step goal, the more likely they were to report higher self-efficacy and thus, H2a and H2c are supported. However, this relationship was not seen when users were asked to rate their self-efficacy for completing tomorrow's step goal. As shown in Table 2, a one-way ANOVA determined there was no statistically significant difference between groups for the

self-efficacy ratings for achieving a 10,000 step goal tomorrow ($F(2, 270) = 1.337, p = .264$). Therefore, we found no support for H2b or H2d.

4.3 H3: Effect of Weight Perception on Self-Efficacy

For the effect of *perception of weight* on users' self-efficacy rating for reaching 10,000 steps today, a linear regression established that *healthy weight perception* statistically significantly predicts self-efficacy ratings for today, $F(1, 270) = 5.238, p = .023$ and *healthy weight perception* accounted for 1.5% of the explained variability in self-efficacy for today (see Table 2). The more likely users were to rate themselves as at a healthy weight, the more likely they were to report higher self-efficacy ratings for completing their goal today, which supports H3a. As Table 2 shows, *overweight perception* was not a significant predictor of self-efficacy for today as determined by a linear regression ($F(1, 270) = 3.012, p = .084$). Therefore, no support was found for H3c.

However, for self-efficacy ratings for achieving 10,000 steps tomorrow, both *healthy weight perception* ($F(1, 270) = 7.078, p = .008$) and *overweight perception* ($F(1, 270) = 5.012, p = .026$) were significant predictors as determined by a linear regression, as shown in Table 2. *Healthy weight perception* accounted for 2.6% of the explained variability in self-efficacy for tomorrow, and *overweight perception* accounted for 1.5% of the explained variability in self-efficacy for tomorrow. In this case, users that were more likely to rate their weight as healthy or overweight were more likely to report higher self-efficacy ratings for completing their goal tomorrow. Therefore, we found support for H3b. However, H3d was rejected. We found no support for either H3e or H3f; *underweight perception* was not a statistically significant predictor of either self-efficacy ratings for today ($F(1, 270) = 2.183, p = .141$) or tomorrow ($F(1, 270) = 1.315, p = .253$) (see Table 2).

4.4 H4: Effect of Interest on Self-Efficacy

For the effect of *interest* on users' self-efficacy rating for reaching 10,000 steps today, a linear regression established that *interest* statistically significantly predicts self-efficacy ratings for today, $F(1, 271) = 12.597, p < .000$ and for tomorrow, $F(1, 271) = 65.783, p < .000$ (see Table 2). *Interest* accounted for 4.1% of the explained variability in self-efficacy for today and 19.2% for tomorrow. Like the work of Choe et al. (2013), the more likely users were to be interested in completing 10,000 steps per day, the more likely they were to report higher ratings of self-efficacy for today and tomorrow. Therefore, H4a and H4b are supported.

4.5 Self-Efficacy Findings Controlling for Variables



Figure 1. Self-efficacy rating for today based on condition controlling for interest and healthy weight perception

	Self-Efficacy Today	Self-Efficacy Tomorrow
<i>Valence</i>	No main effect ($p > .05$)	No main effect ($p > .05$)
<i>Distance to goal/steps completed</i>	Significant main effect ($p = .046$)	No main effect ($p > .05$)
<i>Valence*distance to goal/steps completed</i>	No interaction effect ($p > .05$)	Significant interaction effect ($p = .006$)
<i>Valence*Interest</i>	No interaction effect ($p > .05$)	Significant interaction effect ($p = .007$)
<i>Distance to goal/steps completed*Interest</i>	Significant interaction effect ($p = .036$)	No interaction effect ($p > .05$)
<i>Distance to goal/steps completed*Healthy weight perception</i>	Significant interaction effect ($p = .048$)	No interaction effect ($p > .05$)
<i>Distance to goal/steps completed*Interest*Healthy weight perception</i>	Significant interaction effect ($p = .040$)	No interaction effect ($p > .05$)

Table 3. Summary of main and interaction effects when controlling for *interest* and *weight perception*

Since *interest* and *perception of weight* (specifically, *healthy weight perception* and *overweight perception*) both impacted users' self-efficacy ratings, we ran a one-way analysis of covariance (ANCOVA) to control for those variables, as shown in Table 3. For self-efficacy ratings for today (see Figure 1), we found a main effect for *steps completed*, $F(1, 270) = 3.109$, $p = .046$, but no main effect for *valence*, $F(1, 270) = .299$, $p = .742$.

There was a significant interaction effect between *steps completed* and *interest*, $F(2, 270) = 3.367$, $p = .036$, an interaction effect between *steps completed* and *healthy weight perception*, $F(2, 270) = 3.078$, $p = .048$, and an interaction effect among *steps completed*, *interest*, and *healthy weight perception*, $F(2, 270) = 3.251$, $p = .040$. This indicates that a difference in *interest* had different effects on the self-efficacy rating for today at different *steps completed* and that a difference in *healthy weight perception* had different effects on the self-efficacy rating for today at different *steps completed*.

For self-efficacy ratings for tomorrow (see Figure 2 on the following page), we found no main effect for *steps completed*, $F(2, 270) = 1.000$, $p = .370$, or *valence*, $F(2, 270) = 2.242$, $p = .091$. However, there was an interaction effect between *valence* and *interest*, $F(2, 270) = 5.013$, $p = .007$ and an interaction effect between *valence* and *steps completed*, $F(4, 270) = 3.727$, $p = .006$. This indicates that a difference in color (*valence*) had different effects on the self-efficacy rating for tomorrow at different *interest* levels and that a difference in color (*valence*) had different effects on the self-efficacy rating for tomorrow at different distances (*steps completed*) to the goal.

To break this interaction down, we ran simple contrast (which is a more focused test to compare the difference between two means in factorial designs with more than two levels of one or more of the independent variables), which there was a marginally significant difference ($p = .052$) between revealed low *steps completed* (2500) and high *steps completed* (7500). However, the midpoint (5000) was not significantly different than the low *steps completed* ($p = .149$) or the high *steps completed* ($p = .646$). The simple contrast found no significant differences for *valence*.

When controlling for *interest* and *perception of weight*, trends emerge that can further research in this area. The closer users are to their goal, the more likely they are to report higher levels of self-efficacy for achieving today's goal, which is consistent with the work of Choe et al. (2013). This was observed for each condition (although the results were not statistically significant). Those users who had the black progress bar (neutral condition) reported higher self-efficacy ratings for today at each steps completed condition as compared to those who had red (negative) and green (positive). At 2500 steps completed, users with the red progress bar reported higher self-efficacy ratings for today than those with the green progress bar. At 5000 and 7500 steps completed, on the other hand, users with the green progress bar reported higher ratings of self-efficacy for today than those with red.

These trends differed for the self-efficacy ratings for completing tomorrow's goal. For instance, there were noticeable differences in trends of the neutral condition (black progress bar) for self-efficacy ratings for today and tomorrow. When users were asked how confident they were that they could achieve a 10,000 step goal tomorrow based on feedback from today's goal, users who had the black progress bar reported lower self-efficacy ratings for tomorrow the closer they were to achieving their goal today.

The opposite trend was seen for the users with both the red and green progress bars; they reported higher self-efficacy ratings for tomorrow the closer they were to achieving their goal today. At 2500 steps completed, users who had the red progress bar reported higher self-efficacy ratings for tomorrow than those with green but lower than those with black. At 5000 and 7500 steps completed,

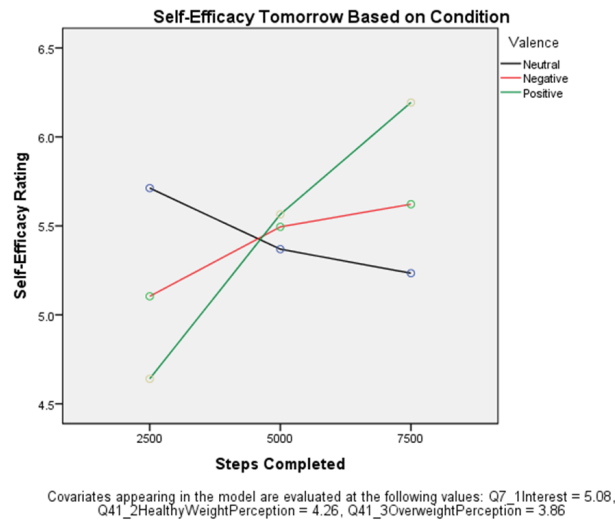


Figure 2. Self-efficacy rating for tomorrow based on condition controlling for interest and healthy and overweight perception

those with the red progress bar reported lower self-efficacy ratings for tomorrow than those with green but higher than those with black.

5 Discussion

5.1 Implications

These findings have a number of implications. While distance to goal may be related to self-efficacy because individuals feel as though they do not have enough time to reach their goal (i.e. if it's 4:30PM and I've only completed 25% of my goal, then it is less likely I will reach my goal), there may be ways we can still promote their self-efficacy. For immediate goals, we may be able to increase users' self-efficacy by making their goal appear more reachable through either different visualizations or more information about how to reach those goals (for instance, tell users, "You can walk another 1,200 steps in only 10 minutes or the time it takes to listen to less than 3 songs!"). On the other hand, for long-term goals, we may want to be careful to frame today's feedback as it can impact users' beliefs about tomorrow.

As consistent with the work of Choe et al. (2013), *interest* played a role in self-efficacy ratings for today. It is important to note that interest played a more significant role in users' self-efficacy ratings for achieving tomorrow's goal, which is crucial in long-term successful behavior change. This finding may help us understand motivation (Deci, 1992), as interest may be one reason some are motivated to exercise. Interest is related to individuals' intrinsic motivation, or motivation that results from internal factors rather than external ones (Deci, 1992). For instance, a person may be motivated to exercise simply because they find the activity interesting and enjoyable. While this finding does not impact design per se, we can find ways to interest people in healthy activities by making health systems and apps more entertaining. One possible way is through gamification, which has been shown to produce positive effects in the areas of health and wellness (Hamari, Koivisto, & Sarsa, 2014).

Unlike Choe et al. (2013), we examined the effects of *weight perception* on self-efficacy. Interestingly, for accomplishing today's health goal, users' who are more likely to rate themselves as healthier are more likely to report higher self-efficacy ratings. For reaching tomorrow's goal, those users who are more likely to rate themselves as overweight or at a healthy weight are more likely to report higher ratings of self-efficacy. It is not surprising that users who believe they are healthy have higher confidence in accomplishing health goals both today and tomorrow. However, it is interesting that users who believe they are overweight have higher confidence in achieving tomorrow's goals (but not for today's goals). Those users who reported being overweight may be interested in losing weight or becoming healthier and thus, believe that even if they cannot accomplish their goal today, they may be

able to tomorrow. Understanding how weight and body perception influence users may help us cater design to specific audiences.

While we found no significant results for *valence* alone, we speculate that an increased sample size may show significant results where we saw trends in the data. When controlling for *interest* and *weight perception*, we found significant interaction effects for *valence* and *interest* and for *valence* and *distance to goal/steps completed* but only for tomorrow's self-efficacy ratings. This may suggest that varying color choices may impact users' beliefs in their ability to accomplish future goals at different distances to their goals for today. If we wish to promote confidence in achieving future goals, it seems best to frame the information more neutrally (with a black progress bar) when users are farther from reaching their goal today. When users are closer to reaching their goal today, it may be best to frame the feedback positively (with a green progress bar). For both immediate and future goals, we have to be aware of unintended consequences of trying to frame feedback positively using color (in this case, using green) when users are in early stages of their goal and negatively (in this case, using red) in later stages of their goal. For instance, users who were given a green progress bar but were only at 2500 steps completed may have felt the feedback was not accurate based on their goal and thus, may have felt less confident in their ability to reach their goal today. If users do not feel the feedback is sincere, we may unintentionally lower their belief in their abilities to reach their goal.

5.2 Limitations and Future Research

This research is not without its limitations. While the sample allowed for a wide range of participant types, sample size was an issue due to the number of participants excluded. Although this number is consistent with the work of Choe et al. (2013), removing this many participants left us with roughly 30 participants per condition. A larger sample size may help show significant results where there were trends in the data. While using a between-subject design makes sense in this case (due to potential sensitization and carry-over effects of having the same participants test each condition), we cannot account for random error/noise due to individual differences in the same way we could if we used a within-subjects design.

Many apps are introducing other colors, such as yellow, to show progress and progress bars may also incorporate more than one color. It is important to also consider colors such as blue, as blue hues have been highly studied in other contexts and are shown to be positively impact people's behaviors (Bellizzi & Crowley, 1983). More research about using the color black and other "neutral" colors is also needed in relation to promoting short-term and long-term goals. Another important note about color is its impact on different people. Studies have shown that color affects males and females differently (Gnambis et al., 2010). While red and green are commonly used colors to denote negative and positive emotions respectively in western societies, these colors may elicit different emotions in other areas of the world (Gnambis et al., 2010). Therefore, future research will benefit from testing the impact of many different colors on males and females in different cultures and how it relates to both psychological concepts linked to health behavior change and actual behavior.

While understanding how to present feedback in fitness devices is useful, fitness apps only represent a portion of the health apps available. We also need to study feedback in other technologies, such as calorie counting apps, weight loss apps, etc. Additionally, we have to conduct more studies that examine the impact of weight and body perception as well as interest as they play a role in users' self-efficacy. Future research may help us uncover ways to customize apps and devices based on these factors.

Another limitation is not working with users' actual data. This may have impacted their self-efficacy ratings as the data provided were simulated. Additionally, while it is important to examine whether or not the presentation of feedback impacts self-efficacy as it is linked to successful health behavior change (Strecher et al., 1986), self-efficacy is only one psychological concept that can impact health behavior change. Future studies need to examine other concepts, such as motivation. Although interest can help explain intrinsic motivation, adults are less often motivated purely by interest (Deci, 1992). Therefore, we need to further explore other factors (including extrinsic ones) that motivate individuals to exercise (such as to lose weight, look better, become healthier).

Since our goal is to change behavior, we also need to measure actual behaviors and outcomes. However, given the limited knowledge in this area, this study represents an important first step to understanding the effect of colors, distance to goal, weight perception, and interest within health apps and devices. Future studies should also consider the idea of tailoring health information as individual differences may play a role in how information presentation affects users. In future research, we plan on

exploring not only how design impacts self-efficacy but also actual behaviors over time by running an experiment using an actual app as opposed to a simulated one.

6 Conclusion

This preliminary study examined the effect of color choice on users' self-efficacy ratings for accomplishing an exercise goal today and tomorrow. While no significant results were found specifically for color choice (*valence*), there are trends in the data that suggest using green and red can be used to positively and negatively (respectively) frame feedback. Both interest and perception of weight play an important role in users' confidence in their abilities to reach their goals. Distance to goal (*steps completed*) influences users' self-efficacy for today but not for tomorrow. These findings have numerous implications for design. Through design, we can promote self-efficacy and positively impact health outcomes. However, more research is needed to examine the effects of color choice in health apps and devices on user self-efficacy and behavior.

7 References

- Anderson-Bill, E. S., Winett, R. a, & Wojcik, J. R. (2011). Social cognitive determinants of nutrition and physical activity among web-health users enrolling in an online intervention: The influence of social support, self-efficacy, outcome expectations, and self-regulation. *Journal of Medical Internet Research*, 13(1), e28. doi:10.2196/jmir.1551
- Bandura, a. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191–215. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/847061>
- Bellizzi, J. A., & Crowley, A. E. (1983). The effects of color in store design. *Journal of Retailing*, 59(1), 21–45.
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, 6(1), 3–5. doi:10.1177/1745691610393980
- Choe, E. K., Lee, B., Munson, S., Pratt, W., & Kientz, J. A. (2013). Persuasive performance feedback: The effect of framing on self-efficacy. In *AMIA* (pp. 1–9). Washington, D.C.
- Deci, E. L. (1992). The relation of interest to the motivation of behavior: A self-determination theory perspective. In *The role of interest in learning and development* (Renninger,., pp. 43–70). Hillsdale, NJ.
- Edell, B. H., Edington, S., Herd, B., Brien, R. M. O., & Witkin, G. (1987). Brief Report: Self-efficacy and self-motivation as predictors of weight loss. *Addictive Behaviors*, 12, 63–66.
- Elfhag, K., & Rössner, S. (2005). Who succeeds in maintaining weight loss? A conceptual review of factors associated with weight loss maintenance and weight regain. *Obesity Reviews: An Official Journal of the International Association for the Study of Obesity*, 6(1), 67–85. doi:10.1111/j.1467-789X.2005.00170.x
- Elliot, A. J., Maier, M. a, Moller, A. C., Friedman, R., & Meinhardt, J. (2007). Color and psychological functioning: The effect of red on performance attainment. *Journal of Experimental Psychology: General*, 136(1), 154–168. doi:10.1037/0096-3445.136.1.154
- Gao, Z., Xiang, P., Lochbaum, M., & Guan, J. (2013). The impact of achievement goals on cardiorespiratory fitness: Does self-efficacy make a difference? *Research Quarterly for Exercise and Sport*, 84(3), 313–22. doi:10.1080/02701367.2013.814908

- Gnambs, T., Appel, M., & Batinic, B. (2010). Color red in web-based knowledge testing. *Computers in Human Behavior*, 26(6), 1625–1631. doi:10.1016/j.chb.2010.06.010
- Hall, R. H., & Hanna, P. (2004). The impact of web page text-background color combinations on readability, retention, aesthetics, and behavioral intention. *Behaviour & Information Technology*, 23(3), 183–195.
- Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work? - A literature review of empirical studies on gamification. *2014 47th Hawaii International Conference on System Sciences*, 3025–3034. doi:10.1109/HICSS.2014.377
- Kondratova, I., & Goldfarb, I. (2007). Color Your Website : Use of Colors on the Web, 123–132.
- Manley, D., Cowan, P., Graff, C., Perlow, M., Rice, P., Richey, P., & Sanchez, Z. (2014). Self-efficacy, physical activity, and aerobic fitness in middle school children: examination of a pedometer intervention program. *Journal of Pediatric Nursing*, 29, 228–37. doi:10.1016/j.pedn.2013.10.011
- Moore, R. S., Stammerjohan, C. A., & Coulter, R. a. (2005). Banner advertiser-Web site context congruity and color effects on attention and attitudes. *Journal of Advertising*, 34(2), 71–84. doi:10.1080/00913367.2005.10639189
- Schunk, D. (1990). Goal setting and self-efficacy during self-regulated learning. *Educational Psychologist*, 25(1), 71–86.
- Singh, S. (2006). Impact of color on marketing. *Management Decision*, 44(6), 783–789. doi:10.1108/00251740610673332
- Strecher, V. J., Devellis, B. M., Becker, M. H., & Rosenstock, I. M. (1986). The role of self-efficacy in achieving behavior change health. *Health Education Quarterly*, 13(1), 73–92.
- Valdez, P., & Mehrabian, A. (1994). Effects of color on emotions. *Journal of Experimental Psychology. General*, 123(4), 394–409. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/7996122>