Designing Closeness to Increase Gamers' Performance

Lucas Colusso, Gary Hsieh, Sean A. Munson
Human Centered Design and Engineering
University of Washington
{colusso, garyhs, smunson}@uw.edu

ABSTRACT
Designers often make use of social comparisons to motivate people to perform better. In this paper, we present the concept of *closeness to comparison* to improve the efficacy of social comparison feedback. Specifically, we test two design strategies related to *closeness*: (1) comparing users to a target described as a similarly experienced player and (2) adjusting the visual representation of performance so player scores appear closer to the comparison target. We evaluate the effects of these strategies for social comparison on player performance in an online game. In a controlled experiment with 425 participants, both feedback techniques improved game performance, but only for experienced players. We conclude with design implications for helping designers create social comparisons that motivate higher game performance.

Author Keywords
Games and Play; Social comparison.

ACM Classification Keywords
H.5.m. Information Systems: Information interfaces and presentation.

INTRODUCTION

Social comparison is one of the most popular strategies used by designers to influence user behavior [16]. For example, social comparison information is often shown in games, health applications, financial tools, and other systems to motivate users [15,17,13,16]. In games specifically, designs often focus on comparisons to players with higher, or the highest, scores. Leaderboards, for example, are extremely popular in a variety of games, from the 'Hall of Fame' screen on arcade video games to multiplayer rankings and typically show the highest scoring players and sometimes the players ranked immediately above and below the current player [7]. Other applications also use social comparison features to motivate participation. Fitbit, a fitness-tracking app, shows a leaderboard with the user's friends ranked by 7-day step totals in their gamified dashboard.

Comparisons to people who perform better (upward comparison) can encourage players to improve their ranking, which prompts feelings of increased self-efficacy [15] and often leads people to improve their performance [8]. However, despite these benefits, observers can judge themselves incapable of reaching them and decide not to engage in the tasks if the comparison is too extreme [10,11,21]. While comparisons to higher performers can motivate performance, they can also be daunting and undermine performance, leading to jealousy, low self-esteem, or general dislike [10,11,21]. As designers strive to keep their users active and engaged [16], designing comparisons to higher performers poses a challenge.

Drawing on research on social comparison [10,11,21] we propose a class of designs for social comparison that may preserve the benefits of comparisons to higher performers, but mitigate some of the aforementioned negative consequences. These strategies rely on the idea of bringing users *closer* to the comparison target, which can make them evaluate their own chances of success more positively. People tend to be motivated by comparisons where others are relatively close with respect to opinions or performance [10,11]. Positive evaluation of comparisons leads to engagement and triggers action [10,11]. The first strategy manipulates the *comparison target group*: comparing a user to the top performer within a group of players with similar expertise rather than the top player across all players. Prior research has shown that comparing people to those who are similar can be more effective than comparisons to people who are not [10,21]. The second strategy manipulates the *visual representation* of the performance feedback: displaying performance feedback that has been visually skewed so that the user appears closer to the comparison target, rather than visually accurate performance feedback.

Through an online game, we evaluate whether the *closeness to comparison* concept, and its implementation in the proposed strategies, can affect player performance. We find that comparing players to others labeled as "similar", and visually skewing player initial performance up, led experienced players to achieve higher scores. We provide game designers with insights on how to design features using social comparisons that make comparisons to higher performances seem less daunting, supporting an increase in user performance.

CLOSENESS AS A COMPARISON STRATEGY

According to social comparison theory, people are propelled by a unidirectional drive to improve their performance and to minimize discrepancies between their and others' performance [10]. These factors generate competitiveness, a manifestation of the social comparison process [10,11].

CHI'16, May 07 - 12, 2016, San Jose, CA, USA.

ACM 978-1-4503-3362-7/16/05 $15.00
DOI: http://dx.doi.org/10.1145/2858036.2858206

In prior work, comparison strategies vary between comparison to close and distant social groups and the distance to them regarding the attribute in question [11,14,19]. This prior research offers a framework to better understand social comparisons and serves as a design space to explore and evaluate. In particular, research describes three factors that increase comparison effects: *closeness to comparison target*, *similarity to the target's performance*, and *relevance of performance* [10,11]. For scope, we explored the use of the first two factors in this work. We discuss opportunities for using the third factor in discussion.

The factor of *closeness to comparison target* posits that the effects of social comparisons are generally stronger when the target is interpersonally close (e.g., a friend) or seen as similar (e.g., someone with the same background) [10,11]. Additionally, seeing others' successful experiences supports people's self-evaluations and beliefs they can reach the same results [10,15]. Creating textual feedback using the closeness to comparison target factor motivates our first hypothesis:

H1. *Describing a comparison target as similar will motivate players to achieve higher performance than when they are compared to a target described as the leader.*

The second factor is *similarity to the target's performance*. People assign greater attention to comparison information when it is similar to theirs, rather than comparisons to higher performers [10,11]. Subtle visual feedback can affect behaviors, even when people do not consciously perceive it, [1,2] and self-efficacy, a common performance mediator [3]. Since theory posits that when people consider a certain performance to be reachable, they will be more motivated to engage in a task [5,10], we form a second hypothesis about adjusting the visual representation of comparison targets so that players appear closer to them (Figure 1):

H2. *Adjusted bar graphs so a player appears closer to the comparison target will motivate players to achieve higher performance than proportional comparison bar graphs.*

METHODS

To evaluate the effects of these social comparison strategies in design, we conducted a between-subjects experiment. Participants were invited to play a casual online game called Flappy Bird (Figure 2). In this game, players have to keep the bird from falling to the ground by pressing the spacebar to make it "fly", while avoiding hitting the pipes (obstacles). The objective is to fly past as many pipes as possible. Their score in the game is the total number of pipes they flew pass without "dying" (falling to the ground or hitting the pipes) in a single round. For our study, participants had to complete at least one round of the game and could then play as many rounds as they liked. After each round, the player would receive feedback, which persisted throughout the following round (Figure 2, left). When a player beats the "leader's" score, their score becomes the "leader" score to mimic leaderboards, which show the a player's score as the leader's score after they beat that score. This still presents

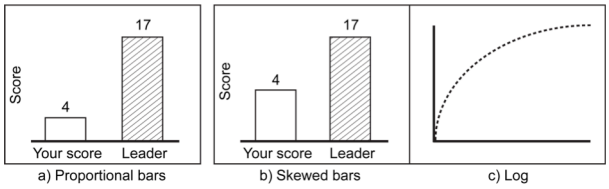

Figure 1. Differentiation between the a) Proportional feedback and b) Skewed feedback.

an implied social comparison that the player would then be top ranked (or top ranked among similar players).

Participants were randomly assigned into one of three conditions for the feedback shown after each round. The baseline condition compared players to the *leader* of the game (biggest number of pipes crossed in one round) using a *proportional* comparison bar graph (Figure 1a). This represents the type of social comparison you would often seen in casual games. Second, for *target group closeness* (H1), participants were told before starting the game they would be be compared to players with a similar gaming experience. The actual feedback displayed "Similar player" instead of "Leader". We controlled the scores shown so they were the same across conditions. In *performance visual representation closeness* (H2), we adjusted the visual representation of feedback using a logarithmic scale (Figure 1b). This adjusted lower scores up, so that participant performance appeared closer to their target's performance though the scores remained unchanged (Figure 1).

As previously shown, people put more effort into difficult goals, so long as they are attainable [6]. To test different levels of difficulty, we conducted a pilot with 400 players. We then selected three scores that we used in the experiment – high (196), medium (17) and low difficulty (9) – corresponding to the highest, third and second quartiles of pilot players high scores, respectively.

The post-survey included questions relating to Videogame Self-efficacy [6], Social Comparison (SCO) [12], Self-esteem [18], and an Intrinsic Motivation Inventory [9]. We also asked open-ended questions about their goals for playing and why they left the game (see Appendix).

Data and code available at: github.com/lucascolusso/floppybird.

Participants

For the experiment, we recruited 430 participants through Amazon Mechanical Turk. We only included Turkers living in the US, with more than 90% HIT success rate, and more than 100 hits approved. Each participant received a fixed $1 US rate to complete the task. We removed data from five

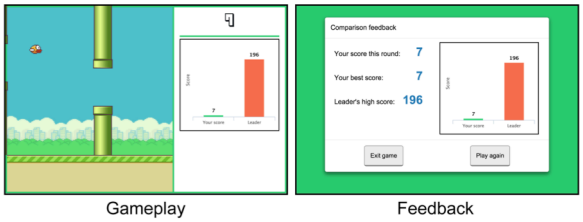

Figure 2. Flappy bird game main screens.

participants who failed attention check questions. The remaining 425 participants were balanced in gender and mean age was 34 years. 143 participants were assigned to a closer *target group*, 153 to a closer *performance visual representation*, and 129 assigned to baseline.

To assess player prior gaming experience, we asked about their gaming experience with a 5-point scale before the game and how frequently they play video games in a post survey that was presented when they left the game (Appendix). Both measures correlate ($r(423) = 0.80, p < 0.001$). The distribution of self-reported experience was concentrated on the top point scale options (n = 340, 80%). We then categorized participants as experienced gamers if they reported 3 or above in the experience scale.

RESULTS

Participants' highest scores averaged 14.4 points (SD = 19.8). Outcomes were independent of players beating comparison scores and the comparison levels. They played 9.5 rounds on average (SD = 12.2). Based on the distribution of player high scores, we selected Poisson regression to model the effects of the manipulations on player performance (Table 1, Model 1). On initial inspection, we observed that both comparison to a similar target (mean = 16.5, SD = 22.5) and visually skewed performance graphs (mean = 15.2, SD = 22.0), increased players performance ($p < 0.001$) relative to the baseline condition (mean = 11.2, SD = 12.5), where players saw a comparison to the leader through an accurate graph.

On further inspection we observed what appeared to be an interaction effect between player gaming experience and the effects of the manipulations (Figure 3). Poisson regression (Table 1, Model 2) confirms this result: both manipulations appear to have affected only experienced gamers.

To understand differences in effect on experienced and inexperienced gamers, we coded the open-ended questions. Inexperienced and experienced gamers reported different goals in playing the game. Inexperienced gamers (n = 85) had discovery-like goals, such as "Just get some points"; "To know what the game is about"; "unambitiously trying

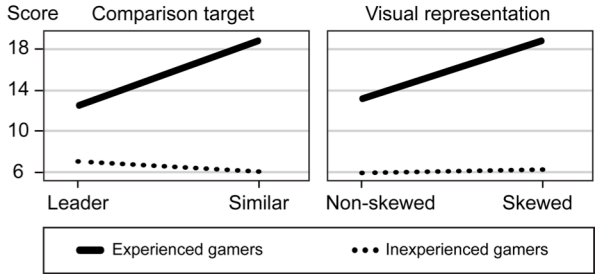

Figure 3. Interaction between scores, the IV's (Comparison target, left and Visual representation, right) with participant's level of experience (experienced or inexperienced gamers).

the game" (n = 46) or non-competitive goals: "learning game mechanics"; "test game and how I perform" (n = 29). In contrast, experienced gamers (n = 340) had competitive goals: "beating the score" or "scoring as high as possible" (n = 181). When experienced gamers with the five highest scores were asked why they stopped playing (each played less than three rounds, below the sample's average of nine), they said they had already beaten their comparison scores and there was nothing left to accomplish. The comparison goals are more important to experienced gamers (e.g., "I got pass the goal and made it to 201", "I had a high score and the game was getting dull", "I beat the score", "I stopped because I was 4x the leaders score.") than inexperienced gamers. We speculate that experienced gamers attended to the comparisons more than inexperienced gamers, and so only experienced gamers were influenced by the comparisons.

We should also note that higher social comparison orientation (SCO) predicted a higher score. This is expected since all conditions used social comparison. However, when we examine the interaction effect between SCO and closeness strategies, we found that gamers with low SCO, skewed had a positive effect on score, but for gamers with high SCO, skewed had a negative effect on scores. We discuss this interaction in the next session.

DISCUSSION AND LIMITATIONS

This study explores the use of two *closeness to comparison* strategies to improve social comparison feedback for game performance. The first strategy is to tell participants they are being compared to similarly experienced players (H1 –

	Model 1: Main Effects			Model 2: Main and Interaction Effects		
	Estimate	Std. Error	*p*	Estimate	Std. Error	*p*
Baseline Performance (intercept)	1.4533	0.0664	<0.001 ***	1.2591	0.1127	<0.001 ***
Comparison Score	0.0008	0.0001	<0.001 ***	0.0009	0.0001	<0.001 ***
Experience (coded)	0.8330	0.0435	<0.001 ***	0.5056	0.0732	<0.001 ***
Social Comparison Orientation (SCO)	0.0634	0.0142	<0.001 ***	0.2083	0.0282	<0.001 ***
Similar Target	0.3764	0.0335	<0.001 ***	-0.0199	0.1545	0.8975
Similar Target * Experience (coded)	-	-	-	0.6516	0.1085	<0.001 ***
Similar Target * SCO	-	-	-	-0.0537	0.0368	0.1447
Visual Skewing	0.2643	0.0335	<0.001 ***	1.0283	0.1497	<0.001 ***
Visual Skewing * Experience (coded)	-	-	-	0.2658	0.1047	0.0111 *
Visual Skewing * SCO	-	-	-	-0.3123	0.0359	<0.001 ***

*$p < .05$, **$p < .01$, ***$p < .001$.

Table 1. We observed main effects for labeling the target as a similar player and for visual skewing (model 1). On closer inspection, however, these effects occurred only for experienced players (model 2).

target group closeness). The second strategy is to make player performance appear closer to the comparison score (H2 – performance visual closeness). Both strategies led to higher performance for experienced players, which suggest they may enhance the benefits of social comparisons or avoid its possible downsides. Our work also points out an additional challenge and opportunity when using these strategies to improve game performance: they seem to only improve performance for experienced gamers. A possible explanation may be that these strategies affect experienced gamers because only they care about the game enough. If someone cares about an activity, it is natural that they are more likely to be affected by information about it [10,11]. However, it is important to point out that inexperienced gamers played about the same number of rounds (M = 11.6, SD = 12.24) as the experienced gamers (M = 9.04, SD = 12.26). This raises doubts that intrinsic interest in the game is the primary reason why inexperienced gamers do not seem to be affected by the proposed design strategies.

Our qualitative data suggest an alternative explanation. Experienced gamers wanted to beat the comparison scores, while inexperienced players were interested in exploration, playing in discovery mode. This indicates that experienced and inexperienced gamers may represent two different types of players. Using Bartle's Taxonomy of player types [4] to understand the qualitative data, we find that inexperienced gamers behaved like explorers while experienced gamers as achievers – score oriented, enjoy beating challenges. It is not that either group is more or less interested in the game, but that they have different play goals and foci. More research is needed to grasp the differences between inexperienced and experienced gamers and to better support their game play using social comparisons.

Interestingly, we found an interaction effect between SCO and the visual closeness (*skewed*) manipulation. The skewed visualization had a positive effect for players who had low SCO, but a negative effect for those players who had high SCO (those exposed to skewed scored less). This suggests a potential drawback in the visual closeness strategy: while visual closeness can improve the attainability of a score, it might be less motivating for those who are driven by comparisons (high in SCO). Additional research is needed to explore this potential tradeoff.

Implications for design

Our study shows that experienced players perform better when they think they are *closer* to their comparison. On a high level, this means that drawing experienced players *closer* to comparisons to higher performers is an effective way for game designers to support an increase in player's performance. While this strategy worked only for experienced players, it did not harm the performance of inexperienced players. Here, we describe design strategies to achieve comparison closeness.

Target group closeness. Our results suggest that designers could frame or deceive players into thinking they are being compared to similar others to increase their performance. One approach is to simply inform the player that they are being compared to a similar user group, through the idea of benevolent deception [2]. A more practical solution is to highlight the similarities between players to create the sense of closeness. Alternatively, designers may also narrow the comparison group to a subset of similar others instead of providing comparisons to all other players.

Performance visual closeness. Subtly skewing performance feedback may seem like an unusual technique for supporting performance increase, but it is an example of benevolent deception [1,2]: one that benefits both the system designer and the end-user. In our implementation, designers do not need to fabricate performance results, but merely adjust the presentation of graphical feedback. In designing performance feedback to enhance experienced gamer performance, we propose skewing visualizations to present favorable comparisons. Designers can make user performance appear closer to their comparisons.

In this study, we tested visual cues that made a gamer's performance appear similar to their target's. Researchers and designers may want to develop further techniques using benevolent deception, following the work of [1,2,7]. In addition, more research is needed to explore how *closeness* interacts with other attributes, such as SCO, background, age groups, and physical location. Research would also benefit from understanding different use cases, such as newcomers to a group or people who are not intrinsically interested in a given task.

Game design has often informed design in other domains, such as gamification in personal informatics and self-tracking. Comparison closeness might be an effective strategy for enhancing the effects of feedback in the context of exercise (e.g., Fitbit), workplace performance systems, energy consumption, education, or safety.

CONCLUSION

We conducted an online experiment with 425 participants who played a casual game with performance feedback on their score. The concept of closeness to comparison was tested with two different designs: target group closeness and performance visual closeness. A closer target group comparison and skewed performance feedback helped increasing gamer performance, but only for experienced gamers, who were motivated by comparisons that deemed them closer to the leader's performance or similar to the other player. Using the concept of closeness, we provide game designers with insights on how to design features using social comparisons that make comparisons to higher performers seem less daunting. These strategies can support an increase in user performance.

ACKNOWLEDGEMENTS

We thank NSF #1348543 and CAPES Brazil #13544/13-6 for funding, members of the Prosocial Computing Lab, Daniel Epstein, Ahmer Arif, and Leandro Oliveira.

REFERENCES

1. Alexander T. Adams, Jean Costa, Malte F. Jung, and Tanzeem Choudhury. 2015. Mindless computing: designing technologies to subtly influence behavior. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (UbiComp '15). ACM, New York, NY, USA, 719-730. DOI=http://dx.doi.org/10.1145/2750858.2805843
2. Eytan Adar, Desney S. Tan, and Jaime Teevan. 2013. Benevolent deception in human computer interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*(CHI '13). ACM, New York, NY, USA, 1863-1872. DOI=http://dx.doi.org/10.1145/2470654.2466246
3. Albert Bandura and David C. McClelland. 1977. Social learning theory. Prentice-Hall Englewood Cliffs, NJ. DOI=http://dx.doi.org/10.2307/2065952
4. Richard Bartle. 1996. Hearts, clubs, diamonds, spades: Players who suit MUDs. Journal of MUD research 1, 19.
5. David A. Bergin. 1999. Influences on classroom interest. In Educational psychologist 34, 2, 87-98. Taylor & Francis. DOI=http://dx.doi.org/10.1207/s15326985ep3402_2
6. Lucas Blair. 2011. The use of video game achievements to enhance player performance, self-efficacy, and motivation. University of Central Florida Orlando, Florida.
7. Jason T. Bowey, Max V. Birk, and Regan L. Mandryk. 2015. Manipulating Leaderboards to Induce Player Experience. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play* (CHI PLAY '15). ACM, New York, NY, USA, 115-120. DOI=http://dx.doi.org/10.1145/2793107.2793138
8. Sunny Consolvo, Predrag Klasnja, David W. McDonald, and James A. Landay. 2009. Goal-setting considerations for persuasive technologies that encourage physical activity. In *Proceedings of the 4th International Conference on Persuasive Technology* (Persuasive '09). ACM, New York, NY, USA, Article 8, 8 pages. DOI=http://dx.doi.org/10.1145/1541948.1541960
9. E. L. Deci, R. M. Ryan. 2005. Intrinsic motivation inventory (IMI). Retrieved July 23, 2006.
10. Leon Festinger. 1954. A theory of social comparison processes. In Human relations 7, 2, 117-140. DOI=http://dx.doi.org/10.1177/001872675400700202
11. Stephen Garcia, Avishalom Tor, Tyrone Schiff. 2013. The Psychology of Competition: A Social Comparison Perspective. In *Perspectives on Psychological Science* 8, 6, 634-650. Sage Publications. DOI=http://dx.doi.org/10.1177/1745691613504114
12. Frederick X. Gibbons and Bram P. Buunk. 1999. Individual differences in social comparison: development of a scale of social comparison orientation. In Journal of personality and social psychology 76, 1, 129. American Psychological Association. DOI=http://dx.doi.org/10.1037/0022-3514.76.1.129
13. F. Maxwell Harper, Joseph Konstan and Sherry Xin Li. 2010. Social comparisons and contributions to online communities: A field experiment on movielens. In *The American economic review* 1358-1398. JSTOR. DOI=http://dx.doi.org/10.1257/aer.100.4.1358
14. Sasha Issenberg. October 31, 2012. The "Voter Report Card" MoveOn Hopes Will Shame Slackers. In Slate.com. Accessed July 23, 2015. Retrieved from http://www.slate.com/blogs/victory_lab/2012/10/31/moveon_can_its_voter_report_card_shame_slackers_into_turning_out.html
15. Daniel King, Paul Delfabbro and Mark Griffiths. 2010. Video game structural characteristics: A new psychological taxonomy. In *International Journal of Mental Health and Addiction* 8, 1, 90-106. Springer. DOI=http://dx.doi.org/10.1007/s11469-009-9206-4
16. Robert E. Kraut, Paul Resnick, Sara Kiesler, Moira Burke, Yan Chen, Niki Kittur, Joseph Konstan, Yuqing Ren and John Riedl. 2012. *Building successful online communities: Evidence-based social design*. MIT Press.
17. Fitbit. Retrieved from https://www.fitbit.com/
18. Richard W. Robins, Holly M. Hendin and Kali H. Trzesniewski. 2001. Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. In *Personality and social psychology bulletin* 27, 2, 151-161. Sage Publications. DOI=http://dx.doi.org/10.1177/0146167201272002
19. Jen Shang and Rachel Croson. 2009. A field experiment in charitable contribution: The impact of social information on the voluntary provision of public goods. *The Economic Journal* 119, 540, 1422-1439. Wiley Online Library. DOI=http://dx.doi.org/10.1111/j.1468-0297.2009.02267.x
20. Edward R. Tufte. The visual display of quantitative information. 2, 9, Graphics press Cheshire, CT. DOI=http://dx.doi.org/10.1016/0195-6663(91)90030-v
21. Ladd Wheeler, René Martin and Jerry Suls. 1997. The proxy model of social comparison for self-assessment of ability. In *Personality and Social Psychology Review* 1, 1, 54-61. Sage Publications. DOI=http://dx.doi.org/10.1207/s15327957pspr0101_4