STOCKUS, CHRISTOPHER A. Ph.D. Experiences and Forecasts of the Big-Fish-Little-Pond Effect. (2024)
Directed by Dr. Ethan Zell. 133 pp.

Research on the big-fish-little-pond effect (BFLPE) indicates that students evaluate themselves more favorably when they have high rank in low rank schools than low rank in high rank schools. In this integrated dissertation, I provide three empirical papers examining experiences and forecasts of this effect. BFLPE experiences refer to the impact of having either high rank in a low rank group or low rank in a high rank group on one's own self-evaluations and performance-related affect. Forecasts of the BFLPE refer to the predicted impact of these same social comparison conditions on another person's self-evaluations and performance-related affect. In empirical paper one, we examined whether experienced BFLPEs occur in different regional groups (i.e., at the school, state, and country level). In empirical paper two, we tested whether growth mindsets, that is, the belief that one's abilities are malleable and capable of change, ameliorate the negative effects of having low rank in a high rank group. In empirical paper three, we examined whether the BFLPE occurs in empathic forecasting. In sum, this program of research contributes to our understanding of the conditions and contexts in which the BFLPE is likely to emerge, provides additional support for the robustness of the BFLPE, further establishes the causal nature of the effect, and indicates that intergroup comparison neglect is a key mechanism underlying the BFLPE.

EXPERIENCES AND FORECASTS OF THE BIG-FISH-LITTLE-POND EFFECT

by

Christopher A. Stockus

A Dissertation
Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Greensboro

2024

Approved by

Dr. Ethan Zell
Committee Chair

APPROVAL PAGE

This dissertation written by Christopher A. Stockus has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair

Dr. Ethan Zell

Committee Members

Dr. Brittany Cassidy

Dr. Douglas Levine

Dr. Paul Silvia

March 1, 2024

Date of Acceptance by Committee

March 1, 2024

Date of Final Oral Examination

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I: INTEGRATED INTRODUCTION

More than 70 years of social comparison research suggests that people evaluate themselves and form self-concepts based on comparisons made between the self and similar others (Festinger, 1954; Wheeler & Suls, 2020). Social comparisons generally result in more favorable self-evaluations of ability, self-esteem, and affect following downward comparisons (i.e., comparisons between the self and inferior others) than upward comparisons (i.e., comparisons between the self and superior others; Alicke et al., 2013; Gerber et al., 2018). The big-fish-little-pond effect (BFLPE) is an application of social comparison theory and is perhaps the most studied social comparison phenomenon. Specifically, the BFLPE is the tendency for students with above average rank in below average schools to evaluate themselves more favorably and feel better than their peers who have below average rank in above average schools (Marsh, 1987; Marsh et al., 2015). The BFLPE occurs because people focus on their rank relative to others in their group, but pay less attention to the rank of their group relative to other groups, during self-assessment (i.e., local dominance; Zell & Alicke, 2010, 2020).

There is strong support for the occurrence of the BFLPE in educational settings. More specifically, school-average performance is negatively associated with student academic self-concepts after controlling for individual student performance, such that students evaluate themselves more favorably when they have high rank in low rank schools than low rank in high rank schools (Marsh & Parker, 1984; Marsh et al., 2021; von Keyserlingk et al., 2019). The BFLPE obtains across a variety of geographic regions, including Western, Asian, and Middle Eastern countries and across a variety of age groups, including elementary, middle, high school, and college students (Fang et al., 2018; Marsh et al., 2015, 2019). Furthermore, prospective studies suggest that the BFLPE is stable two to four years after graduation from high school

(Marsh et al., 2007). Research suggests that the BFLPE is associated with students' academic self-concepts (Koivuhovi et al., 2022; Televantou et al., 2021; von Keyserlingk et al., 2019) and other important outcomes, including achievement emotions (e.g., enjoyment, pride, anxiety, shame; Basarkod et al., 2023; Pekrun et al., 2019), career aspirations (Nagengast & Marsh, 2012), and motivation (Marsh et al., 2021; Zell et al., 2017). Most work examining the BFLPE is non-experimental, but experimental studies that manipulate student and school-average performance provide causal support (Alicke et al., 2010; Stockus & Zell, 2023b; Zell & Lesick, 2021).

In this dissertation, I provide two papers assessing experiences of the BFLPE and one paper assessing forecasts of the BFLPE. Experienced BFLPEs refer to the effect of having either high rank in a low rank group or low rank in a high rank group on *one's own* self-evaluations and performance-related affect (Zell & Lesick, 2021). In other words, experienced BFLPEs refer to the effect of social comparison conditions on the self. Forecasted BFLPEs refer to the predicted effect of having either high rank in a low rank group or low rank in a high rank group on *another person's* self-evaluations and performance-related affect (Stockus & Zell, 2024a). That is, forecasted BFLPEs involve predicting how social comparison conditions affect another person. Past work on the BFLPE has extensively studied experiences of the effect and its consequences for academic self-concepts (Koivuhovi et al., 2022), self-evaluations (Zell & Lesick, 2021), and emotion (Pekrun et al., 2019). However, past work has not examined forecasts of the BFLPE.

In the present dissertation, I describe three papers testing the occurrence of experienced BFLPEs in different regional contexts (Stockus & Zell, 2023b), whether growth mindsets reduce these experienced BFLPEs (Stockus & Zell, 2023a), and whether the BFLPE occurs in empathic forecasting (Stockus & Zell, 2024a). The first paper examines experiences of the BFLPE in

different regional contexts (i.e., at the school, state, and country level) and how those experiences impact self-evaluations and performance-related affect (Stockus & Zell, 2023b). For example, we examined whether people evaluate themselves more favorably when they have high rank in a low rank country than low rank in a high rank country. The second paper also examines experienced BFLPEs and whether growth versus fixed mindsets, that is, endorsement of the belief that abilities are malleable rather than stable, reduces the size of the effect (Stockus & Zell, 2023a). Specifically, we test whether the detrimental effects of being a little fish in a big pond are reduced following activation of a growth mindset. The third paper examines forecasts of how the BFLPE impacts other people's emotional experience, whether group rank neglect underlies the BFLPE in empathic forecasting, and whether extremity of group ranks amplifies the BFLPE in empathic forecasting (Stockus & Zell, 2024a).

Social Comparison

Social comparison theories argue that people think about themselves in relation to others and rely on others to form evaluations of their abilities and opinions (Garcia et al., 2013; Krizan, 2018). For example, a student may compare their academic abilities to another student in their class or school. Social comparisons made between the self and others may have important implications for self-evaluations and self-concept formation (Festinger, 1954; Wheeler & Suls, 2020). These comparisons may also have pronounced consequences for a variety of other important outcomes, including mood, motivation, and self-esteem (Gerber et al., 2018; Zell & Strickhouser, 2020). People compare themselves to others on different dimensions, including physical attractiveness (Bocage-Barthélémy et al., 2018), verbal ability (Zell & Lesick, 2021), and work performance (Breidenthal et al., 2020).

One early social comparison study indicates that people compare themselves to relevant others and that these comparisons impact self-esteem (Morse & Gergen, 1970). Male college students who were ostensibly candidates for a job were exposed to either a highly desirable or undesirable stimulus person. The desirable stimulus person was well-dressed, composed, and carrying a briefcase. The undesirable stimulus person was poorly dressed and disheveled. Exposure to a desirable stimulus produced a decrease in self-esteem, whereas exposure to an undesirable stimulus increased self-esteem. This study demonstrates that social comparisons can have contrastive effects. That is, comparisons between the self and superior targets deflate self-esteem, whereas comparisons between the self and inferior targets enhance self-esteem.

More generally, social comparisons may result in either assimilation or contrast (Markman & McMullen, 2003; Mussweiler, 2003, 2020). Connective comparisons made between the self and similar others impress upon people that they are similar to the comparison target, resulting in assimilation effects. Non-connective comparisons made between the self and dissimilar others impress upon people that they are different from the comparison target, resulting in contrast effects. As in the study described above, contemporary reviews on social comparison indicate that contrastive effects are the norm (Gerber et al., 2018; Krizan, 2018; Wheeler & Suls, 2020). Social comparison contrast results from two main types of social comparison: upward and downward (Alicke et al., 2013; Gerber et al., 2018; Zell & Strickhouser, 2020). Upward comparisons typically result in more negative self-evaluations and downward comparisons typically result in more positive self-evaluations. For example, a student who compares their poor academic performance with a superior performer (upward comparison) may feel dissatisfied and envious of the referent's performance. Alternatively, a student who

compares their excellent academic performance with that of an inferior performer (downward comparison) may experience feelings of pride and gratitude.

Although most social comparisons result in contrast, social comparison assimilation does occur. Identification of the self with the social comparison referent may influence whether comparisons result in assimilation or contrast, such that people who strongly identify with the referent are more likely to experience assimilation, whereas people who do not identify with the referent are more likely to experience contrast (Markman & McMullen, 2003). For example, an underperforming student who identifies with a strong performer in their class may be motivated to improve their performance at school, resulting in an upward assimilation effect. Conversely, a similar comparison where there is not a strong identification between the target and referent may result in a contrast effect and produce feelings of dissatisfaction, envy, and dejection.

Research suggests that social comparisons occur via both explicit and implicit routes (Bocage-Barthélémy et al., 2018). Some research indicates that social comparisons are often spontaneous, effortless, and unintentional reactions to others' performances. For example, one paper found that subliminal exposure to thin models increased appearance anxiety among women, demonstrating that social comparisons impact self-evaluations via an implicit route (Chatard et al., 2017). Another paper indicates that people compare themselves to others even when they are aware that these comparisons are non-diagnostic and will mentally undo those comparisons only if they have sufficient cognitive resources (Gilbert et al., 1995). In sum, research suggests that social comparisons can proceed either through controlled or automatic routes. Further, automatic comparisons may be more frequent than those that occur via controlled processes. That is, although some comparisons are actively selected by people, many if not most are foisted upon them by the environment (Wood, 1989; Zell & Alicke, 2020).

Finally, research indicates that social comparisons are pervasive in daily life (Summerville & Roese, 2008). In this experience sampling study, four types of comparative thought were examined: social, counterfactual (imagining what might have been), past-temporal (comparison to past circumstances), and future-temporal (comparison to future circumstances). The four types of comparison were found to occur with similar frequency: social (24.2%), counterfactual (25.5%), past-temporal (19.4%), and future-temporal (30.9%). Thus, social comparisons are likely as frequent as other prominent forms of comparative thought. Other research indicates that the frequency of social comparisons may be influenced by factors such as mood and self-esteem (Wheeler & Miyake, 1992). More specifically, pre-comparison negative mood may lead to more upward comparisons than downward comparisons. Moreover, people with high self-esteem may engage in self-enhancing comparisons (e.g., downward comparisons) more often than people with low self-esteem.

Big-Fish-Little-Pond Effect

One application of social comparison is the BFLPE, where students evaluate themselves more favorably when they have high rank in low rank schools than low rank in high rank schools (Marsh et al., 2015; Marsh & Parker, 1984). Being a big fish in a little pond (BFLP) involves a downward comparison at the student level and an upward comparison at the school level. Conversely, being a little fish in a big pond (LFBP) involves an upward comparison at the student level and a downward comparison at the school level. Early research on the BFLPE collected data from 10th grade male students in public high schools in Australia and the United States (Marsh, 1987). Participants evaluated themselves in relation to their peers on dimensions including academic ability, reading ability, and intelligence. Equally able students reported more favorable academic self-concepts in lower ability schools. In other words, when student rank was

controlled for, students felt better about their academic performance at schools with low average ability (where they had high rank) than at schools with high average ability (where they had low rank). The BFLPE has positive effects for students who have high rank in low rank schools and negative effects for students who have low rank in high rank schools, particularly in terms of their self-evaluations and performance-related affect (Zell & Lesick, 2021).

In applied research, the BFLPE is operationalized as the tendency for average ability students to report more favorable academic self-concepts in lower rank schools where they are surrounded by lower performing students (Marsh & Seaton, 2015). The BFLPE is the result of two types of social comparison (i.e., rank information; Marsh et al., 2014; Zell & Alicke, 2010). First is the positive relationship between student rank and academic self-concepts, where student academic self-concepts increase as student rank increases. Second is the negative relationship between school rank and academic self-concepts, where student academic self-concepts decrease as school rank increases (see Figure 1). The BFLPE is the product of these two paths, specifically, the negative relationship between school rank and academic self-concepts controlling for the positive relationship between student rank and academic self-concepts (Marsh, 1987; Marsh et al., 2015).

In applied studies on the BFLPE, multi-level modelling is used to account for the fact that students are nested within classes or schools. Multi-level modelling allows for the analysis of hierarchical or nested data where responses may be situated within larger units and responses within those units may not be independent. For example, a sample may be composed of 5,000 students who attend 40 different schools. Scores from students attending the same school are not necessarily independent. These students may be exposed to similar factors (e.g., teachers, curriculum) that are different in other schools. In BFLPE analyses, individual student

achievement has a positive association with academic self-concepts and school-average achievement has a negative association with academic self-concepts (Marsh & Seaton, 2015). The multi-level nature of the BFLPE model resolves the non-independence of student and school effects by decomposing any relationship between these variables into within- and between-school components.

Whereas applied BFLPE studies measure student and school ranks and assess their association with academic self-concepts (Marsh & Parker, 1984; Marsh et al., 2019), experimental studies of the BFLPE manipulate student and school ranks and assess their impact on self-evaluations and performance-related affect (Zell & Alicke, 2009; Zell & Lesick, 2021). These studies indicate that having high rank in a low rank group has positive effects whereas having low rank in a high rank group has negative effects. In everyday life, people likely consider individual and group rank when making academic and vocational choices, such as which school or job to pursue (Frank, 1985; Gladwell, 2013). Along these lines, one empirical paper suggests that people from individualistic cultures more often choose to be a big fish in a little pond than people from collectivistic cultures (Wu et al., 2018), perhaps because people in collectivist cultures are more interested in affiliating with prestigious groups.

Generalizability and Robustness

There is strong support for the BFLPE in educational settings (Koivuhovi et al., 2022; Marsh et al., 2019, 2021; von Keyserlingk et al., 2019). The BFLPE obtains across a variety of age groups, including elementary, middle, high school, and college students and appears to increase somewhat as students age (Fang et al., 2018; Marsh et al., 2015). The effect also generalizes across geographic regions, including Western, Asian, and Middle Eastern countries, although it appears to be somewhat variable across countries (Fang et al., 2018; Marsh et al.,

2015, 2019), with preliminary evidence that the effect may be somewhat stronger in Western and Asian countries than in Islamic countries (Marsh et al., 2015). Furthermore, the BFLPE demonstrates stability over time and has been shown to remain for young adults two to four years following graduation from high school (Marsh et al., 2007). Other research supports the robustness of the BFLPE by indicating that 13 student-level moderators do not substantially reduce the magnitude of the effect (e.g., achievement motivation; Marsh et al., 2021).

Although the majority of BFLPE studies indicate that it is robust and generalizable, one review synthesizes evidence suggesting that the BFLPE is moderated by select student and teacher factors (Stockus & Zell, 2024b). For example, this review found evidence that the BFLPE is stronger for people who are female, in a higher grade level, high in neuroticism, low in narcissism, and who live in Western societies (e.g., the United States and United Kingdom). Additionally, the effect is weaker for students who have positive relationships with their teachers. Patterns of moderation for student ability, achievement goals, teacher grading practices, and teacher frame of reference were inconclusive. More research is needed to evaluate whether and when moderating factors reduce harmful impacts of the BFLPE (i.e., being an LFBP).

BFLPE Outcomes

Academic self-concepts are the most studied outcome in the BFLPE literature (Televantou et al., 2021), but the effect does extend to other outcomes. Students with above average rank in below average schools report more positive emotions and less negative emotions than their peers with below average rank in above average schools (Zell & Lesick, 2021). In a cross-sectional study, for example, German students in grades 5-10 completed a math achievement test and a math achievement emotions questionnaire (Pekrun et al., 2019, Study 1). The BFLPE was positively associated with positive achievement emotions (i.e., enjoyment) and

negatively associated with negative achievement emotions (i.e., anger, anxiety, and hopelessness). Subsequent studies found that the BFLPE is associated with more positive (e.g., pride) and less negative achievement emotions (e.g., shame) when examined longitudinally across an entire school year (Pekrun et al., 2019, Studies 2-3).

Other research suggests that the BFLPE is associated with more enjoyment in academic domains including math, science, and reading (Basarkod et al., 2023). Further, research suggests that students with high rank in low rank schools are more motivated to succeed in school and have more lofty career aspirations than their peers who have low rank in high rank schools. That is, school performance is negatively associated with motivation and aspirations after adjusting for individual student performance (Marsh et al., 2014; Nagengast & Marsh, 2012). Finally, experimental studies suggest that students' rank in their school has a causal impact on intrinsic motivation and further that student rank has an indirect effect on persistence via intrinsic motivation (Zell et al., 2017). Altogether, both applied and experimental studies suggest that the BFLPE has implications not just for self-evaluations, but also for emotions, motivation, and aspirations.

Experimental Evidence

Most work examining the BFLPE is non-experimental, preventing causal conclusions. However, research manipulating student rank and school rank provides causal support (Zell & Alicke, 2009). For example, in one study, college students completed a verbal reasoning test and received manipulated feedback about their performance indicating that they had either high rank in a low rank school or low rank in a high rank school (Zell & Lesick, 2021, Study 1). Specifically, participants received feedback indicating that they ranked better than 65% (35%) of about 200 students at their school and that their school ranked better than 35% (65%) of about 40

other schools. Participants then completed measures of verbal self-evaluations and performance-related affect. Consistent with the BFLPE, students told that they had high rank in a low rank school evaluated themselves more favorably, and reported more favorable performance-related affect, than students told that they had low rank in a high rank school.

Another experiment testing the BFLPE indicates that it occurs even when participants are randomly assigned to minimal groups, suggesting that mere identification with a group is sufficient to produce the effect (Alicke et al., 2010). In each session, 10 participants were sub-divided into two groups of 5 participants by drawing letters out of a hat. Participants then completed a lie detection test and received feedback indicating that they ranked fifth or sixth relative to all ten people who participated at the same time as them. Further, some participants were told that they ranked fifth overall but worst in their five-person group (i.e., everyone in their group ranked better than everyone in the other group) or sixth overall but best in their five-person group (i.e., everyone in their group ranked worse than everyone in the other group). Consistent with the BFLPE, ranking highest in an inferior group produced more favorable self-evaluations than ranking lowest in a superior group. This paper suggests that even trivial groups are sufficient to produce a BFLPE. Nonetheless, it remains possible that the BFLPE is even larger in more meaningful groups, such as those that people strongly identify with.

Intergroup Comparison Neglect

Other studies suggest that intergroup comparison neglect is a mechanism underlying the BFLPE (Zell & Alicke, 2010, 2020). That is, people tend to neglect the rank of their group and instead focus on their own rank within the group during self-evaluation. In one study, for example, participants completed a verbal reasoning test and were told that they outperformed either 80% or 40% of test-takers at their university (intragroup comparison) and that their

university outperformed either 80% or 40% of other universities (intergroup comparison; Zell & Alicke, 2009, Study 1). Participants with high rank at their school evaluated themselves favorably, and participants with low rank at their school evaluated themselves unfavorably, regardless of their school's rank relative to other schools. A subsequent study using more extreme ranks provides further evidence in support of group rank neglect as a mechanism underlying the BFLPE (Zell & Lesick, 2021, Study 2). Specifically, participants who ranked better than 85% of students at their university evaluated themselves very favorably, and participants who ranked better than 15% of students at their university evaluated themselves very unfavorably, regardless of whether their university ranked better than 85% or 15% of other universities.

One non-experimental study also suggests that intergroup comparison information is neglected during self-assessment, resulting in the BFLPE (Huguet et al., 2009). In this study, French 6th grade students completed an academic self-concept questionnaire and indicated their perceived rank relative to other students in their class. The BFLPE was eliminated when students' perceived rank in their class was entered into the model. This finding suggests that the BFLPE is driven by student rank (i.e., one's rank relative to other students in one's class), not school rank (i.e., the rank of one's school relative to other schools). Thus, having a below average rank at an above average school may be detrimental to student academic self-concepts, regardless of the rank of their school relative to other schools.

Aims of the Present Research

This dissertation presents three papers that contribute to the BFLPE literature by examining experiences and forecasts of the effect. BFLPE experiences refer to the impact of having either high rank in a low rank group or low rank in a high rank group on *one's own* self-

evaluations and performance-related affect. Forecasts of the BFLPE refer to the predicted impact of these same social comparison conditions on *another person's* self-evaluations and performance-related affect. In papers one and two, we describe research examining experiences of the BFLPE, and in paper three, we describe research examining experiences and forecasts of the BFLPE. These papers fill several gaps in the BFLPE literature, specifically by testing the effect experimentally, examining experienced BFLPEs in novel regional groups, testing whether experienced BFLPEs are reduced by growth mindsets, examining whether BFLPEs occur in empathic forecasting, and testing whether experienced and forecasted BFLPEs are governed by the same mechanism (intergroup comparison neglect).

Paper one experimentally tested the emergence of experienced BFLPEs in novel regional contexts (Stockus & Zell, 2023b). More specifically, we tested this experienced BFLPE at the school, state, and country level and investigated whether it is accentuated in more local reference groups. In Study A1 we examined the BFLPE at the country level, in Study A2 we examined the BFLPE at the state and country level, and in Study A3 we examined the BFLPE at the school and country level. We found that a strong BFLPE emerged in each reference group, but consistent with a non-experimental study (Marsh et al., 2019), the regional BFLPE was not more pronounced at the school (state) level than at the country level. Altogether, this paper suggests that the BFLPE occurs not only in educational contexts, but also in regional contexts, such as in one's state or country.

Paper two tested whether growth mindsets moderate experienced BFLPEs (Stockus & Zell, 2023a). Specifically, we tested whether activation of a growth mindset reduces harmful effects of being an LFBP. Growth mindsets refer to the belief that one's abilities are malleable and capable of change, which is associated with greater resilience in educational settings

(Burnette et al., 2023b; Yeager & Dweck, 2020; Zhao et al., 2021). Participants in two studies read an article suggesting that intellectual abilities are malleable or fixed prior to a manipulation of the BFLPE. Results yielded a highly robust BFLPE on self-evaluations and affect. However, the BFLPE was not substantially altered by growth mindsets in either study. Consistent with an earlier non-experimental study which found that the BFLPE was robust to individual differences in growth mindsets (Marsh et al., 2021), this paper provides experimental evidence suggesting that experienced BFLPEs are robust to growth mindset interventions.

Paper three examined experiences and forecasts of the BFLPE (Stockus & Zell, 2024a). In Study C1, participants were given verbal reasoning performance feedback about themselves or another person and reported either their own or the other person's affect. Consistent with earlier studies which found that empathic forecasting biases are pronounced for negative emotions (Boucher et al., 2015; Green et al., 2013), forecasters anticipated that the BFLPE impacted negative emotions significantly more than it actually did. Study C2 found that empathic forecasts were strongly influenced by another person's rank in their group, but only weakly influenced by group rank, resulting in a large BFLPE. Finally, Study C3 found that forecasts were more favorable when the target was a BFLP than an LFBP and that this effect was significantly more pronounced in extreme than moderate feedback conditions. These studies suggest that forecasters anticipate that the BFLPE has a stronger effect on other people's negative emotions than people actually experience, that intergroup comparison neglect underlies the BFLPE in empathic forecasting, and that extremity of group ranks exacerbates the BFLPE in empathic forecasting.

CHAPTER II: THE REGIONAL BIG-FISH-LITTLE-POND EFFECT: EVIDENCE FROM NATIONAL AND SUBNATIONAL COMPARISONS

Stockus, C. A., & Zell, E. (2023b). The regional big-fish-little-pond effect: Evidence from national and subnational comparisons. *British Journal of Social Psychology*, *62*(3), 1158–1176. https://doi.org/10.1111/bjso.12623

According to the big-fish-little-pond effect (BFLPE), students evaluate themselves more favorably when they have high rank in a low rank school than low rank in a high rank school. We examined whether the BFLPE impacts self-evaluations in regional settings, where the reference group is one's nation or subnation. In Study A1, participants told that they ranked above average in a below average nation evaluated themselves more favorably than participants told that they ranked below average in an above average nation. Study A2 demonstrated that this regional BFLPE occurs both when the reference group is one's nation (USA) and one's subnation (e.g., California, Florida). Finally, Study A3 found that the BFLPE occurs and is similar in size when the reference group is one's nation versus one's school. In sum, these experiments provide novel support for the BFLPE in regional contexts and suggest that social comparisons that involve regional groups substantially impact self-evaluations.

Introduction

According to social comparison theory, people evaluate their attributes, abilities, and opinions based on comparisons with similar others (Festinger, 1954). These comparisons are a primary contributor to the self-concept (Gerber et al., 2018; Zell & Strickhouser, 2020) and affect self-evaluations in various domains, including academic abilities, vocational skills, health, and wellness (Križan & Gibbons, 2014; Suls et al., 2020). One of the most robust and oft-studied demonstrations of social comparison is the big-fish-little-pond effect (BFLPE), which is the

tendency for students to evaluate themselves more favorably when they have high rank in a low rank school than when they have low rank in a high rank school, after adjusting for student's objective ability (i.e., grades or standardized test scores; Marsh, 1987; Marsh et al., 2021). The BFLPE likely occurs because students base self-evaluations more on local comparisons indicating their rank relative to other students in their school than general comparisons indicating rank of their school relative to other schools (i.e., local dominance; Zell & Alicke, 2009, 2010).

As of present, most studies on the BFLPE have focused on educational contexts, where the reference group is one's school or classroom (e.g., Pekrun et al., 2019; Zell & Lesick, 2021). Surprisingly few studies have examined whether the BFLPE occurs in other contexts. In the present research, we explore whether the BFLPE occurs in regional contexts, where the reference group is one's nation or subnation (i.e., a smaller regional group within one's country). Only one prior study examined whether the BFLPE occurs in a regional context (Marsh et al., 2019). Further, because this prior study was non-experimental, it remains unclear whether the regional BFLPE has a causal impact on self-evaluations. People often consider their standing in regional groups, such as their neighborhood, county, or country (Kraus et al., 2013; Zell et al., 2018). Nonetheless, it remains unclear whether a BFLPE occurs in these settings.

Educational Context

Nearly 40 years of research suggests that the BFLPE is a robust and generalizable phenomenon when studied in educational contexts. For example, an early study in the United States and Australia found that equally able students reported more positive academic self-concepts in low achieving schools where they had high rank than in high achieving schools where they had low rank (Marsh, 1987). Thus, students who ranked above average in a below average school evaluated their academic ability and intelligence more favorably than students

who ranked below average in an above average school. Subsequent research replicated this BFLPE across a variety of countries and cultures, including European-American, Asian, and Middle Eastern societies (Marsh et al., 2015). The effect also replicates across a variety of age groups, including elementary, middle, and high school students (Fang et al., 2018). Beyond self-evaluations, the BFLPE extends to other outcomes such as achievement emotions (Pekrun et al., 2019), academic motivation (Trautwein et al., 2006), and career aspirations (Nagengast & Marsh, 2012).

Research decomposing the BFLPE suggests that is arises from the differential impact of two types of social comparison (Marsh et al., 2014; Zell & Alicke, 2010, 2020). First, students compare their academic performance to others in their school (i.e., intragroup or local comparison). Students with high rank in their school are considered "big fish" and students with low rank in their school are considered "little fish". Second, students compare the overall academic performance of their school relative to other schools (i.e., intergroup or general comparison). Schools with high rank are considered "big ponds" and schools with low rank are considered "little ponds". Whereas local comparisons have a strong impact on student self-evaluations and affective reactions, general comparisons often have little influence on self-evaluations (Huguet et al., 2009; Zell & Alicke, 2009). Thus, students focus on their rank relative to others in their school but neglect the rank of their school relative to other schools during self-evaluation (Zell & Lesick, 2021).

Most educational studies on the BFLPE are cross-sectional, which prohibits causal conclusions (Marsh & Seaton, 2015). Nonetheless, prospective studies suggest that the BFLPE predicts self-evaluations 2-4 years after graduation (Marsh et al., 2007) and longitudinal studies suggest that changes in student's relative position over time predict changes in self-evaluations

(Koivuhovi et al., 2022; von Keyserlingk et al., 2019). Further, a few experiments provide causal evidence for the BFLPE (e.g., Alicke et al., 2010; Zell & Alicke, 2009). That is, students who are told that they have high rank at a low rank school on a given test evaluate themselves more favorably than students told that they have low rank at a high rank school. Despite robust support, research on the BFLPE has primarily been conducted in educational settings, where the school or class was the reference group (Fang et al., 2018; Marsh et al., 2015). Social comparisons occur in a variety of intergroup contexts, especially regional contexts where people compare themselves with others in their neighborhood, county, or country (Kraus et al., 2013; Zell et al., 2018). As such, it is important to examine whether and to what extent the BFLPE occurs in regional contexts.

Regional Context

Only one prior study examined whether the BFLPE extends from an educational context to a regional context (Marsh et al., 2019). This study utilized cross-sectional data from adolescents in 68 nations to examine whether the BFLPE occurs in math self-concepts both when the reference group is one's school and one's nation. Consistent with the BFLPE in educational settings, students with high rank in a low rank school evaluated themselves more favorably than students with low rank in a high rank school. Moreover, extending this effect to a regional context, students with high rank in low rank nation evaluated themselves more favorably than students with low rank in a high rank nation. Although the two effects were not directly compared, standardized path coefficients were similar when examining the BFLPE in school (.352) and regional contexts (.328), which suggests that the two effects are comparable. Altogether, these data suggest that the BFLPE occurs when the reference group is one's nation, but the use of a non-experimental design prevents causal conclusions about this effect.

To our knowledge, no additional research has examined the BFLPE in a regional context. However, three experiments are consistent with the argument that the BFLPE may occur in non-educational contexts. One experiment gave participants personalized feedback about a group that they rated in an earlier portion of the study (e.g., their gender, age, race, or religion) and found that participants who had high rank in a low rank group assessed themselves more favorably than participants who had low rank in a high rank group (McFarland & Buehler, 1995, Study 4). Another experiment focused on racial groups in the United States and found that participants who had high rank in a racial group that performed poorly assessed themselves more favorably than participants who had low rank in a racial group that performed well (Lesick & Zell, 2023). Finally, the BFLPE occurs even when participants are randomly assigned to minimal groups (i.e., by drawing letters out of a hat; Alicke et al., 2010), which suggests that mere identification with a group is sufficient to produce the effect.

More broadly, research on social status (i.e., income, education, and occupational prestige in comparison to others) suggests that people make regional social comparisons and that these comparisons have implications for well-being. Specifically, people who have higher status relative to others in their community (Hagerty, 2000; Luttmer, 2005) and country (Oishi et al., 2011; Tay et al., 2014) are happier than people with lower status. These relative status effects remain after adjustment for objective status (e.g., absolute income or education; Cundiff & Matthews, 2017; Senik, 2009), and in some cases, have a stronger association with happiness than objective status (Anderson et al., 2012). Lastly, data suggests that people sacrifice objective status (e.g., higher income or achievement) to occupy groups in which they will have higher status (Frank, 1985; Gladwell, 2013). In sum, although prior studies on social status are non-experimental and do not directly test the BFLPE, they are consistent with the possibility that

people prefer arrangements where they have high rank in a low rank regional group than low rank in a high rank regional group.

The Current Research

Extensive research supports the BFLPE in educational contexts (Marsh et al., 2021; Pekrun et al., 2019), however, it remains unclear whether this effect occurs in other contexts. The present set of experiments test for the first time whether a BFLPE occurs in regional contexts and whether this regional BFLPE differs from the traditional BFLPE studied in educational settings. As a proof of concept, Study A1 tested whether the BFLPE occurs when the reference group is one's nation, that is, whether people evaluate themselves more favorably when they have high rank in a low rank nation than low rank in a high rank nation. Study A2 went further by testing whether a regional BFLPE occurs when the reference group is either one's nation or subnation, and whether the size of the effect differs across these regional groups. Finally, Study A3 once again explored whether a regional BFLPE occurs when the reference group is one's nation and whether this effect differs from the traditional BFLPE that occurs when the reference group is one's school.

One (non-experimental) study suggests that the BFLPE occurs in a national context and is similar in size to the BFLPE in an educational context (Marsh et al., 2019). Additionally, other (non-experimental) research indicates that social status in comparison to regional group members is associated with health and well-being (Cundiff & Matthews, 2017; Zell et al., 2018). Informed by prior work, we expected that a robust BFLPE would occur both when the reference group was one's nation and one's subnation. Further, we thought that the BFLPE might be either (1) larger in more local groups such as one's subnation and school than one's nation, since these groups may be more salient, or (2) comparable across different reference groups, given the general

tendency for people to focus on their rank within groups during self-evaluation rather than group rank (Alicke et al., 2010; Zell & Lesick, 2021).

For exploratory purposes, we also tested whether identification with one's subnation or nation moderates the regional BFLPE. Prior work suggests that the BFLPE is weaker among students who strongly identify with their school because they are more attentive to school rank information during self-evaluation (McFarland & Buehler, 1995). Nonetheless, this prior work was limited by low statistical power and subsequent research failed to find moderation of the BFLPE as a function of group identity (Lesick & Zell, 2023). Given the salience of regional identities and the tendency for some people to strongly identify with regional groups such as their nation or subnation (David & Bar-Tal, 2009; Simon et al., 1995), we explored whether identification with one's regional group moderates the BFLPE.

We report all manipulations, measures, and exclusions in these studies. Samples sizes for each study were determined prior to data collection. Studies A2-A3 were pre-registered prior to data collection. All materials, data, analysis code, pre-registrations, and supplemental analyses are publicly available on OSF. In all three of our studies, results were very similar when excluded participants were retained (see the supplemental analyses).

Study A1

Our first experiment tested whether the BFLPE occurs when the reference group is one's nation. Specifically, American adults recruited on Prolific completed a verbal test and were told that they had high rank in a low rank nation or low rank in a high rank nation. Consistent with the BFLPE, we anticipated that participants in the former condition would report more favorable self-evaluations and affect than participants in the latter condition.

Method

Participants and design. We recruited American adults on Prolific who met the following criteria: 19 to 36 years old, had a Prolific study approval rating of 80% or higher (indicating that at least 80% of their prior Prolific studies were completed satisfactorily), were born in the United States, identified English as their first language, identified as either male or female, and indicated that they were willing to participate in deception studies. We restricted the study to young adults since this age group has been the focus of experimental research on the BFLPE (Alicke et al., 2010; Zell & Lesick, 2021). Further, we restricted the study to people who identified as male or female, since this is required to enable a feature in Prolific that promotes gender "balance" (i.e., automatically obtains a similar number of males and females).

We collected data until about 50 participants were obtained in both experimental conditions, after exclusions. Participants were paid $0.60 in exchange for their participation in the five-minute study ($7.20/hour), which was similar to the federal minimum wage in the US. In total, we collected data from 144 adults. Based on exclusion criteria, we excluded 41 participants for failing a feedback manipulation check or expressing suspicion about the provided feedback. These exclusions resulted in a final sample of 103 participants (53 women, 75 White, M_{age} = 28.1).

A sensitivity power analysis indicated that our sample of 103 participants provided 80% power to detect a medium effect (d = 0.56), which was anticipated for the regional BFLPE in Study A1. Prior experiments found large BFLPEs when tested in educational contexts (Zell & Alicke, 2009; Zell & Lesick, 2021). Participants were randomly assigned to one of two feedback conditions. Two additional conditions were collected for another project examining how people

think others would respond to social comparison feedback. These conditions are irrelevant to the research questions tested here and are not discussed further.

Procedures. Participants completed study measures online and were told that the purpose of the study was to measure verbal reasoning ability. Participants first completed a verbal reasoning test consisting of 40 sentence-completion items (Zell & Lesick, 2021; Zell & Strickhouser, 2020). For each item, participants were presented with a sentence that was missing a word (e.g., *In parts of the Arctic, the land grades into the landfast ice so _____ that you can walk off the coast*…) and were asked to select which of five words best completed the sentence (*permanently*, *imperceptibly*, *irregularly*, *precariously*, *slightly*; the correct answer is imperceptibly). For exploratory purposes, we scored participant's actual performance on the test as the total number of items they answered correctly out of 40.

Next, participants received manipulated feedback about their own and their nation's performance (see Table A1 for sample feedback). In a *big-fish-little-pond condition* (BFLP, n = 52), participants were told that they performed better than 68% of participants from USA (206 so far), and that as a group, USA had performed better than 32% of countries taking part in the study (39 so far). Summary information then described that based on this feedback, they ranked above average relative to others in USA and that USA ranked below average relative to other countries. In a *little-fish-big-pond condition* (LFBP, n = 51), participants were told that they performed better than 32% of participants from USA and that as a group, USA had performed better than 68% of countries. Summary information in this condition noted that participants ranked below average relative to others in USA and that USA ranked above average relative to other countries. In both conditions, the order of individual and country feedback was counterbalanced.

After reviewing the feedback, participants completed two dependent measures (Zell & Lesick, 2021). First, participants completed a 3-item measure of self-evaluations, where they evaluated their performance, ability, and task-specific ability (e.g., *How well do you think you performed on the Verbal Reasoning Test?*) using a nine-point scale (1 = *very poorly/bad*, 9 = *very well/good*; α = .83). Second, participants completed a 6-item measure of performance-related affect, indicating how *satisfied*, *proud*, *happy*, *distressed*, *sad*, and *discouraged* they were with their performance (e.g., *How satisfied do you feel about your Verbal performance?*) using a nine-point scale (1 = *not at all*, 9 = *extremely*; α = .89). Negative affect items were reverse scored.

Next, participants completed two manipulation checks asking them to recall their own performance relative to others from USA and the performance of USA relative to other nations (*above average*, *below average*). Participants were also asked if anything seemed suspicious during the study, and if yes, to describe their suspicion in a text box. Lastly, participants completed demographic questions (e.g., age, gender, race-ethnicity), were debriefed, and dismissed.

Participant exclusions. We excluded data from 7 participants for failing the individual manipulation check, 19 participants for failing the group manipulation check, and 15 participants for expressing suspicion about the provided feedback (e.g., "I assume that the feedback given to me had nothing to do with my performance on the test").

Results and Discussion

Consistent with the BFLPE, self-evaluations were much more favorable in the BFLP condition (M = 5.93, SD = 1.24) than the LFBP condition (M = 4.22, SD = 1.46), $t(101) = 6.42$, $p < .001$, d = 1.27, 95% CI [0.84, 1.69]. Similarly, performance-related affect was significantly

more positive in the BFLP condition (M = 5.78, SD = 1.65) than the LFBP condition (M = 4.50, SD = 1.66), $t(101) = 3.92$, $p < .001$, $d = 0.77$, 95% CI [0.37, 1.17]. Suggesting that random assignment was effective, actual performance on the verbal reasoning test did not significantly differ across the BFLP condition (M = 19.92, SD = 9.06) versus the LFBP condition (M = 17.49, SD = 8.18), $t(101) = 1.43$, $p = .156$, $d = 0.28$, 95% CI [–0.11, 0.67]. Further, the effect of feedback condition on self-evaluations and performance-related affect remained significant after adjusting for participant's actual performance on the verbal reasoning test, $F(1, 100) = 38.16$, $p < .001$, $\eta_p^2 = .28$, and $F(1, 100) = 14.61$, $p < .001$, $\eta_p^2 = .13$. Altogether, data from our first experiment demonstrate that a large BFLPE occurs when examined in a regional context involving one's nation.

Study A2

Study A1 provided the first experimental evidence that the BFLPE has a causal impact on self-evaluations and affective reactions when the reference group is one's country, rather than one's school or classroom. Going further, our second experiment tested whether the BFLPE occurs when the reference group is one's nation (USA) as well as one's subnation (e.g., New York, California). We anticipated that if the BFLPE occurs when the reference group is one's nation, it likely also occurs when the reference group is one's subnation. Furthermore, we examined whether the size of the BFLPE varies across national and subnational levels, and for exploratory purposes, whether the degree to which people identify with their nation or subnation moderates the BFLPE.

Further, a limitation of Study A1 was that a relatively large percentage of participants were excluded for manipulation check failures (18%) or for reporting suspicion about the feedback (10%). To reduce manipulation check failures in Study A2, we screened out

participants who were at risk of providing low quality data (i.e., had successfully completed fewer than 20 prior studies on Prolific) and we required participants to correctly enter the feedback they received before proceeding. To reduce suspicion, we used a motion graphic and 7 second delay after the test to indicate that their performance was being "graded".

Method

Participants and design. We recruited American adults on Prolific who met the following criteria: resided in California, Texas, Florida, or New York, had a Prolific study approval rating of 80% or higher, had 20 or more approved studies, were born in the United States, identified English as their first language, and indicated that they were willing to participate in deception studies. Because the feedback in this study had to be tailored to the specific subnation of the participant, we selected four subnations for data collection rather than all 50 subnations or "states" within the United States (which would have created logistical challenges). The four subnations that we selected are the most populous ones in the US and represent four different geographic locations (i.e., Pacific, South-Central, Southeast, and Northeast).

Further, because this study was restricted to four subnations in the US, we chose not to further restrict the sample by age or gender (as we did in Study A1), out of concern that this might limit the size of the participant pool and slow data collection. Participants were paid $1.00 in exchange for their participation in the five-minute study ($12/hour). The payment rate was slightly increased in Study A2 (versus Study A1) to account for a higher minimum wage in some of the subnations examined.

As in Study A1, we collected data until about 50 participants were obtained in each experimental condition, after exclusions. An a priori power analysis indicated that a sample of

210 participants would provide 95% power to detect a medium effect (f = .25), which was the smallest effect we were interested in detecting. In total, we collected data from 217 adults. We excluded data from 5 participants for failing a manipulation check, attention check, or suspicion probe, resulting in a final sample of 212 participants (170 women, 160 White, M_{age} = 27.4). We obtained a similar number of participants from each of the four subnations (California = 51, Texas = 54, Florida = 55, New York = 52). Participants were randomly assigned to one condition in a 2 (feedback condition: BFLP, LFBP) by 2 (reference group: nation, subnation) between-subjects design.

Procedures. Unless noted otherwise, Study A2 procedures were identical to Study A1. That is, participants completed a verbal reasoning test and received manipulated feedback about their performance. To promote believability, a motion graphic was displayed after participants completed the test for 7 seconds to indicate that their test was being "graded" (i.e., a spinning and flashing wheel accompanied by the text "Please Wait").

Participants in *BFLP conditions* were told that they ranked better than 68% of 206 participants from their group and that their group performed better than 32% of 39 groups. Further, participants in *LFBP conditions* were told that they ranked better than 32% of 206 participants from their group and that that their group performed better than 68% of 39 groups. As in Study A1, some participants received feedback indicating how well they performed relative to others in USA and how well USA performed relative to other nations (*nation conditions*). Unique to Study A2, other participants received feedback indicating how well they performed relative to others in their subnation (e.g., California) and how well their subnation performed relative to other subnations (*subnation conditions*). In all conditions, summary information noted that participants ranked above average in a below average nation/subnation or

below average in an above average nation/subnation. To promote recall, participants were required to correctly enter the feedback and summary information they received before proceeding.

Participants next completed the same measures of self-evaluations (α = .86) and performance related-affect as Study A1 (α = .92). Participants also completed a 4-item measure of the degree to which they identified with the reference group they received feedback about (e.g., *I often regret that I'm a resident of USA [Florida]*) using a seven-point scale (1 = *strongly disagree*, 7 = *strongly agree*). Negatively worded items were reverse scored (α = .90).

Participants then completed a feedback manipulation check, which asked them to recall the scores they received (*above average in a below average group*, *below average in an above average group*, or *average in an average group*) and an attention check, which asked them to select the 4th option in a seven-point scale. We used a single manipulation check that combined individual and group feedback in Study A2 since participants were required to manually enter both sets of feedback they received and thus responding to two separate items (as in Study A1) would have been redundant with this process. Finally, participants completed the same suspicion probe and demographic items as Study A1.

Participant exclusions. We excluded data from 3 participants for failing the feedback manipulation check, 0 participants for failing the attention check, and 2 participants for expressing suspicion about the provided feedback.

Results

Self-evaluations. We conducted a 2 (feedback condition) by 2 (reference group) ANOVA on self-evaluations. Consistent with the BFLPE, there was a significant main effect of feedback condition, $F(1, 208) = 127.82$, $p < .001$, $\eta_p^2 = .38$. Specifically, participants in the

BFLP conditions (M = 5.00, SD = 1.05) evaluated themselves much more favorably than participants in the LFBP conditions (M = 3.31, SD = 1.13). However, the main effect of reference group and the reference group by feedback condition interaction were nonsignificant, $F(1, 208) = 2.52$, $p = .114$, $\eta_p^2 = .01$ and $F(1, 208) = 0.23$, $p = .634$, $\eta_p^2 < .01$ (see Figure A1).

Despite the nonsignificant interaction, we used planned comparisons to examine the effect of feedback condition in each reference group. As in Study A1, there was a very large BFLPE when examining the nation conditions, $t(208) = 8.25$, $p < .001$, $d = 1.60$, 95% CI [1.34, 2.19]. Unique to Study A2, there was also a very large BFLPE when examining the subnation conditions, $t(208) = 7.73$, $p < .001$, $d = 1.51$, 95% CI [1.21, 2.04]. Further, whereas participants in the BFLP conditions evaluated themselves favorably regardless of whether the reference group was their nation or subnation, $t(208) = 0.79$, $p = .428$, $d = .16$, 95% CI [–0.25, 0.58], participants in the LFBP conditions evaluated themselves unfavorably regardless of whether the reference group was their nation or subnation, $t(208) = 1.45$, $p = .150$, $d = .27$, 95% CI [–0.11, 0.73]. Thus, Study A2 yielded a strong regional BFLPE on self-evaluations, and this effect was similar in nation and subnation conditions.

Affect. Parallel results were obtained for performance-related affect. That is, a 2 by 2 ANOVA yielded a significant main effect of feedback condition, $F(1, 208) = 121.62$, $p < .001$, $\eta_p^2 = .37$, a non-significant main effect of reference group, $F(1, 208) = 0.93$, $p = .336$, $\eta_p^2 < .01$, and a non-significant interaction, $F(1, 208) = 0.17$, $p = .678$, $\eta_p^2 < .01$. As expected, performance-related affect was significantly more positive in the BFLP condition than the LFBP condition, both when the reference group was one's nation, $t(208) = 7.431$, $p < .001$, $d = 1.42$, 95% CI [1.27, 2.19], and one's subnation, $t(208) = 8.17$, $p < .001$, $d = 1.61$, 95% CI [1.42, 2.32]. Further, whereas positive affect was relatively high in the BFLP conditions regardless of the

reference group, $t(208) = 0.99$, $p = .325$, $d = .19$, 95% CI [–0.23, 0.68], it was relatively low in the LFBP conditions regardless of the reference group, $t(208) = 0.38$, $p = .701$, $d = .07$, 95% CI [–0.37, 0.55].

Exploratory analyses. Indicative of successful random assignment, a one-way ANOVA found that there were no significant differences across the four feedback conditions in actual performance on the verbal reasoning test ($p = .858$). Further, when verbal performance was entered as a covariate in a 2 (feedback condition) X 2 (reference group) ANOVA, the main effect of feedback condition remained significant for self-evaluations ($p < .001$), as well as performance-related affect ($p < .001$). Thus, the effect of feedback condition cannot be accounted for by actual differences in verbal reasoning performance. Further, exploratory 2 (feedback condition) by 4 (participant subnation: CA, TX, FL, NY) ANOVAs on self-evaluations and performance-related affect found that the feedback condition by subnation interaction was non-significant (ps > .440; see Table A2). Thus, a regional BFLPE emerged and was similar in size across samples from four US subnations.

To explore potential effects of group identity, we conducted a 2 (feedback condition) by 2 (reference group) by group identity (continuous) mixed model on self-evaluations and performance-related affect. For self-evaluations, both two-way interactions involving group identity, as well as the three-way interaction, were non-significant (ps > .065; see Table A3). However, for performance-related affect, the three-way interaction was significant, $F(1, 204) = 7.56$, $p = .007$, $\eta_p^2 = .04$. Analyses decomposing the three-way interaction found that whereas the feedback condition by group identity interaction was not significant when the reference group was one's subnation, $F(1, 105) = 1.10$, $p = .297$, it was significant when the reference group was one's nation, $F(1, 104) = 6.30$, $p = .014$ (see Figure A2). More specifically, the impact of

feedback condition on performance related-affect was greater among participants who strongly identified with their nation (+1 *SD*; $b = 2.24$, $SE = 0.33$, $p < .001$) than participants who weakly identified with their nation (–1 *SD*; $b = 1.06$, $SE = 0.33$, $p = .002$). Thus, identification with one's nation exacerbated the BFLPE.

Discussion

Results from Study A2 provide additional support for the regional BFLPE. That is, participants who had high rank in a low rank region evaluated themselves significantly more favorably, and reported more positive performance-related affect, than participants who had low rank in a high rank region. These effects occurred and were highly robust both the when the reference group was one's nation and one's subnation. Finally, we found some evidence that group identification moderates the BFLPE, such that the effect was larger on performance-related affect when people strongly identified with their nation.

Study A3

Studies A1-A2 provided support for the regional BFLPE when the reference group was one's nation and Study A2 suggests that this effect is comparable to the regional BFLPE when the reference group is one's subnation. In Study A3, we once again examined whether a regional BFLPE emerges at the nation level, but this time also examined how it might differ from the traditional BFLPE examined at the school-level. Based on a (non-experimental) study which suggests that these effects both occur and are comparable in size (Marsh et al., 2019), we anticipated that the regional BFLPE would occur and would be either similar or somewhat smaller in magnitude than the traditional BFLPE at the school level.

Method

Participants and design. We recruited students from a racially diverse public university in the Southeastern United States to participate in exchange for credit in introductory psychology courses. An a priori power analysis indicated that a sample of 210 participants would provide 95% power to detect a medium effect ($f = .25$), which was the smallest effect we were interested in detecting. In total, we collected data from 236 students. We excluded data from 13 students for failing a manipulation check, failing an attention check, or for reporting suspicion about the provided feedback, resulting in final sample of 223 students (169 women, 70 White, 83 African American, $M_{age} = 19.2$). Participants were randomly assigned to one condition in a 2 (feedback condition: BFLP, LFBP) by 2 (reference group: nation, school) between-subjects design.

Procedures. Participants completed study measures online. Participants first completed a verbal reasoning test consisting of 40 sentence-completion items, saw a motion graphic indicating that their test was being "graded" for 7 seconds, and then received manipulated feedback about their performance. Participants in *BFLP* (*LFBP*) *conditions* were told that they ranked better than 68% (32%) of 206 participants from their group and that their group performed better than 32% (68%) of 39 groups. As in the *nation conditions* described previously, some participants received feedback indicating how well they performed relative to others in USA and how well USA performed relative to other nations. Unique to Study A3, other participants in *university conditions* received feedback indicating how well they performed relative to other students at their university and how well students at their university performed relative to other universities. All participants received summary information noting that they ranked above average in a below average nation/university or below average in an above average

nation/university. As in Study A2, participants were required to correctly enter their feedback and summary information before proceeding.

Following the feedback, participants completed the same measures of self-evaluations (α = 0.81) and performance related affect as Studies A1-A2 (α = 0.85). Participants also completed a parallel measure of group identification, to indicate the degree to which they identified with their nation or university (α = 0.82). Participants next completed the same manipulation check as Study A2. They also completed three exploratory measures, namely, a 3-item measure of group-evaluations, where they evaluated their group's performance, ability, and task-specific ability (α = 0.92), an 11-item measure of social comparison orientation (i.e., the tendency to make and be impacted by social comparisons; Gibbons & Buunk, 1999), and a second 10-item verbal reasoning test (post-feedback). The second verbal test was included to assess whether feedback condition significantly impacted subsequent performance. Lastly, participants completed the same suspicion probe and similar demographic items to Studies A1-A2.

Participant exclusions. We excluded data from 9 participants for failing the feedback manipulation check, 3 participants for failing the attention check, and 1 participant for expression about the provided feedback.

Results

Self-evaluations. We conducted a 2 (feedback condition) by 2 (reference group) ANOVA on self-evaluations. Consistent with the BFLPE, there was a significant main effect of feedback condition, $F(1, 219) = 105.19$, $p < .001$, $\eta_p^2 = .32$, indicating that participants in the BFLP condition ($M = 4.29$, $SD = 1.16$) evaluated themselves more favorably than participants in the LFBP condition ($M = 2.77$, $SD = 1.02$). However, the main effect of reference group was nonsignificant, $F(1, 219) = 0.04$, $p = .845$, $\eta_p^2 < .01$, as was the feedback condition by reference

group interaction, $F(1, 219) = 0.49$, $p = .486$, $\eta_p^2 < .01$ (see Figure A3). Thus, the BFLPE on self-evaluations did not vary as a function of the reference group examined.

As in Studies A1-A2, planned comparisons found a robust BFLPE when the reference group was one's nation, $t(219) = 7.98$, $p < .001$, $d = 1.53$, 95% CI [1.22, 2.01]. Further, the BFLPE was also significant and similar in size when the reference group was one's university, $t(219) = 6.57$, $p < .001$, $d = 1.23$, 95% CI [0.99, 1.83]. Whereas participants in BFLP conditions evaluated themselves favorably regardless of whether the reference group was one's nation or university, $t(219) = –0.36$, $p = .723$, $d = 0.06$, 95% CI [–0.48, 0.34], participants in the LFBP conditions evaluated themselves unfavorably regardless of whether the reference group was one's nation or university, $t(219) = 0.63$, $p = .528$, $d = 0.13$, 95% CI [–0.28, 0.54].

Affect. Similar results were obtained for performance-related affect. That is, a 2 by 2 ANOVA yielded a very large main effect of feedback condition, $F(1, 219) = 93.46$, $p < .001$, $\eta_p^2 = .30$. There was also a small but significant main effect main effect of reference group, $F(1, 219) = 5.98$, $p = .015$, $\eta_p^2 = .03$, such that participants reported slightly more positive affect when the reference group was their university ($M = 4.33$, $SD = 1.40$) than their nation ($M = 4.00$, $SD = 1.34$). Finally, the feedback condition by reference group interaction was not significant, $F(1, 219) = 0.21$, $p = .651$, $\eta_p^2 < .01$, which suggests that the BFLPE was constant across reference groups.

Indeed, a substantial BFLPE on performance-related affect occurred in both the nation conditions, $t(219) = 6.71$, $p < .001$, $d = 1.26$, 95% CI [1.01, 1.84], and the university conditions, $t(219) = 6.96$, $p < .001$, $d = 1.33$, 95% CI [1.12, 2.01]. Once again, participants reported relatively low levels of positive affect in the LFBP conditions regardless of the reference group, $t(219) = 1.41$, $p = .160$, $d = 0.27$, 95% CI [–0.12, 0.74]. Further, participants reported relatively

high levels of positive affect in the BFLP conditions, but this tendency was somewhat more pronounced when the reference group was one's university than one's nation, $t(219) = 2.05$, $p = .041$, $d = 0.39$, 95% CI [0.02, 0.88].

Exploratory analyses. One-way ANOVAs found no significant differences in pre-feedback performance ($p = .152$), post-feedback performance ($p = .189$), or social comparison orientation ($p = .054$) across the four conditions. Further when (pre-feedback) verbal performance was entered as a covariate in a 2 by 2 ANOVA, the main effect of feedback condition on self-evaluations ($p < .001$) and performance-related affect ($p < .001$) remained significant. Thus, the effect of feedback condition remained after adjusting for any initial group differences in performance. Further, a 2 by 2 ANOVA on group-evaluations yielded only a significant main effect of feedback condition, $F(1, 219) = 136.75$, $p < .001$, $\eta_p^2 = .38$, such that participants in the BFLP condition ($M = 3.40$, $SD = 1.24$) rated their group much less favorably than participants in the LFBP condition ($M = 5.17$, $SD = 0.99$). Therefore, although participants appeared to neglect group-rank information during self-evaluations, they did utilize this information during group-evaluations.

As in Study A2, we conducted a 2 (feedback condition) by 2 (reference group) by group identity (continuous) mixed model on both outcomes to explore potential moderating effects of group identity. For self-evaluations, both two-way interactions and the three-way interaction involving group identity were non-significant (ps > .150; see Table A4). Similarly, there were no interactions involving group identity for performance-related affect (ps > .340). Thus, the degree to which participants identified with their country or school did not substantially moderate the BFLPE.

Discussion

Study A3 replicated the regional BFLPE, demonstrated that it occurs when the reference group is both one's nation and one's school, and suggests that it is similar in size across these reference groups. Further, although Study A2 yielded some evidence that the BFLPE is moderated by group identity, Study A3 did not replicate this pattern. That is, the BFLPE was robust regardless of group identity. Finally, exploratory analyses in Study A3 demonstrate that participants use group rank information to evaluate their group, but not themselves. This finding suggests that participants adequately processed and understood group rank information, but still neglected it during self-evaluation.

General Discussion

The big-fish-little-pond effect (BFLPE) is one of the most robust and oft-cited phenomena in the social comparison literature. Dozens of studies indicate that students evaluate themselves more favorably when they have high rank in a low rank school than low rank in a high rank school (e.g., Marsh et al., 2021; Pekrun et al., 2019). Additionally, one non-experimental study documented a BFLPE when the reference group was one's country (Marsh et al., 2019). Nonetheless, most studies on the BFLPE have been conducted in educational settings, where the reference group was one's class or school. People often make social comparisons in regional groups, such as their neighborhood, county, or country (Zell et al., 2018). Therefore, in the present research, we conducted three experiments testing whether the BFLPE occurs in a regional context and whether the size of the effect varies across different contexts.

As a proof of concept, Study A1 explored whether a BFLPE occurs when the reference group is one's nation and found that participants with high rank in a low rank nation evaluated themselves more favorably than participants with low rank in a high rank country. Study A2

replicated this regional BFLPE and found that it is comparable to the BFLPE that emerges when the reference group is one's subnation. Finally, Study A3 found highly robust and similar BFLPEs when the reference group was one's nation versus one's school. Exploratory analyses found that this regional BFLPE was robust after adjusting for differences in actual verbal performance across conditions and occurred both for participants who strongly and weakly identified with their group.

Implications

This research makes several contributions to theoretical perspectives on social comparison, especially the BFLPE. First, we uniquely tested whether the BFLPE occurs when the reference group is one's region instead of one's class or school. Our studies provide highly robust evidence for the effect both when the reference group is one's nation (Studies A1-A3) and one's subnation (Study A2), which strongly supports the existence of a regional BFLPE. These data bolster a prior study, which provided non-experimental evidence for a BFLPE at the nation level (Marsh et al., 2019) as well as prior experiments, which suggest that the BFLPE may extend to minimal groups (Alicke et al., 2010) and demographic groups such as one's racial-ethnic or gender group (Lesick & Zell, 2023; McFarland & Buehler, 1995). Advancing this literature, the present work provides the first causal evidence for a regional BFLPE at both the nation and subnation level, and therefore suggests that social comparisons with regional group members have a substantial impact on self-evaluations.

Relatedly, our data also suggest that regional groups may be an important frame of reference for self-evaluations. That is, in each study, participants evaluated themselves much more favorably, and reported much more favorable affect, when they had high rank rather than low rank in their regional group. Therefore, just as schools and classrooms may be an important

frame of reference for self-evaluations (Marsh et al., 2014; Pekrun et al., 2019), nations and subnations may be salient frames of reference. As such, the present research complements prior non-experimental studies indicating that perceived status in regional groups, such as one's community or country, is positively associated with mental and physical health (Cundiff & Matthews, 2017; Zell et al., 2018).

Second, the present research suggests that the magnitude of the BFLPE is similar across different regional and educational contexts. Specifically, a large and comparable BFLPE was obtained when the reference group was one's nation (Studies A1-A3), subnation (Study A2), and school (Study A3). These findings are consistent with a non-experimental study which found comparable BFLPEs when the reference group was one's school and one's nation (Marsh et al., 2019). More broadly, the present findings are consistent with the argument that intragroup or local comparisons have a stronger influence on self-evaluations than intergroup comparisons (i.e., local dominance; Marsh et al., 2014; Zell & Alicke, 2010). Prior work indicated that students neglect the rank of their school relative to other schools during self-evaluation and instead focus on their within-school rank (Zell & Alicke, 2009; Zell & Lesick, 2021). The present data suggest further that people neglect the rank of their nation (subnation) relative to other nations (subnations) during self-evaluation and instead focus on their within nation (subnation) rank.

Third, exploratory analyses tested whether identification with one's group reduces the magnitude of the BFLPE. Although limited by low statistical power (*n*s < 15 per group), prior work suggests that the BFLPE is significantly *smaller* among people who strongly identify with their group, likely because they are more attentive to group rank information (McFarland & Buehler, 1995). However, subsequent research in a racial-ethnic context found that the BFLPE

was not significantly influenced by group identification (Lesick & Zell, 2023). Further, in the present research, we found that group identification generally did not influence the BFLPE, except in one case where participants who strongly identified with their group evidenced a *larger* BFLPE (Study A2, affective reactions). Thus, the present research argues against the assertion that group identification reduces the BFLPE.

Fourth and finally, the BFLPE has mostly been examined non-experimentally and early experimental studies were limited by small sample sizes (Alicke et al., 2010; McFarland & Buehler, 1995). The present work provides robust support for the BFLPE in relatively high-powered experimental designs with sample sizes of over 50 participants per condition. Such experiments provide important causal support for the BFLPE (Zell & Lesick, 2021) and help address alternative explanations for non-experimental studies, such as reverse causation and third-variable explanations. More generally, our research is unique in its experimental approach to investigating the effects of social comparison processes on self-evaluations, which is especially rare in the field of educational psychology. All materials and measures for the present work are publicly available, which may facilitate additional experiments.

Limitations

A limitation of the present research was that there were a few methodological differences between Studies A1 and A2, despite both being conducted on the same platform (Prolific). Specifically, Study A1 was limited to young adults, was gender balanced, utilized two rather than one manipulation check, had a slightly lower payment rate, and was missing some methodological refinements used in Study A2 that likely lowered manipulation check failures and suspicion. However, the regional BFLPE was substantial in both studies, which suggests that the effect is robust to these differences across studies. Further, given that Study A2 (as well as

Study A3) primarily included women, a unique strength of Study A1 was its demonstration of the regional BFLPE in a gender balanced design. Another limitation was that our responses to item 3 (dependent measures) and item 6 (data exclusions) in our pre-registrations were somewhat vague (see OSF for pre-registrations). The dependent measures and criteria for data exclusions were derived in advance from previous research (Zell & Lesick, 2021; Zell & Strickhouser, 2020) and thus we did think it was necessary to elaborate upon them. Nonetheless, we acknowledge that our pre-registrations may lack the level of precision required for full pre-registration.

Although a key strength of the present research was its use of an experimental approach, this approach necessitated a somewhat artificial situation where participants received precise feedback about their own and their regional group's rank on a verbal reasoning test. In everyday life, people may not have clear knowledge about their own or their regional group's rank in the verbal domain. However, non-experimental studies conducted in real-world settings indicate that people who perceive themselves as having high social status in regional groups (i.e., high levels of education, income, occupational prestige, respect, and admiration) evidence better mental and physical health than those who perceive themselves as having low status (Anderson et al., 2012; Kraus et al., 2013). Our experimental studies uniquely complement this work by demonstrating that social comparisons in regional contexts have a causal influence on self-evaluations and emotions.

Lastly, it is important to recognize that most research on the BFLPE is done by scholars in educational psychology, who use multi-level modelling to examine the association of school-average achievement with academic self-concepts after adjusting for individual student achievement, in large datasets consisting of thousands of students across many schools (Marsh et

al., 2008; Marsh & Seaton, 2015). Conversely, the present research used an experimental approach to investigate the social comparison processes underlying the BFLPE in smaller samples of students and adults, typically in a single region or school. Therefore, educational studies on the BFLPE have much greater external validity than the present work, given that they are conducted in real-world settings and in larger samples. However, by randomly assigning people to social comparison conditions, the present work enables causal conclusions and is therefore higher than past work in internal validity.

Future Directions

This research may inspire additional non-experimental and experimental studies exploring whether the BFLPE obtains when the reference group is one's region instead of one's school. Moreover, although the present work examined large regional groups such as one's nation or subnation, future research should explore whether the BFLPE occurs in more local groups, such as one's neighborhood, county, or community. Finally, given that the BFLPE has received strong support in educational settings (Marsh et al., 2015; Pekrun et al., 2019) and initial support in regional settings (Marsh et al., 2019), future research should examine whether the BFLPE occurs in other group contexts. For example, research could examine whether people who have high rank in low rank companies, sports teams, and religious communities evaluate themselves more favorably than people low rank in high rank groups. Further investigations of the BFLPE would help assess whether the effect universally occurs across a variety of intergroup contexts or whether it is restricted to a narrow set of group contexts, such as educational and regional groups.

Future research is also needed to explore whether and to what extent the BFLPE varies across different reference groups. Prior non-experimental research found that the BFLPE was

comparable when the reference group was one's school and one's nation (Marsh et al., 2019). Additionally, the present work found that the BFLPE was comparable when the reference group was one's nation, subnation, and school. Altogether, the available evidence suggests that the BFLPE is highly robust regardless of the reference group examined. Nonetheless, given the narrow set of reference groups examined as of present, it remains possible that the effect varies in contexts that have yet to be explored. For example, the BFLPE may be larger when the reference group is highly cohesive, salient, or relevant to the performance domain (Zell & Alicke, 2010).

Finally, research is needed to further evaluate group identity as a moderator of the BFLPE. In the present research, group identification did not consistently moderate the regional BFLPE after a verbal test. However, it remains possible that group identification moderates the regional BFLPE when examining other performance domains that relate more closely to people's regional identities (e.g., intelligence, social skills, and athleticism).

Conclusions

Many studies have examined the BFLPE in educational settings, documenting its effects on outcomes such as self-evaluations and achievement emotions (e.g., Pekrun et al., 2019; Zell & Lesick, 2021). The present research is among the first to examine whether a regional BFLPE occurs when the reference group is one's nation or subnation. We also tested whether the BFLPE fluctuates across regional and educational settings. Overall, we found strong support for the claim that the BFLPE occurs when the reference group is one's nation, subnation, and school and that the effect is comparable across these different group contexts. Future research is needed to explore the universality of the BFLPE in other regional and non-regional settings. Being a big

fish in a little pond may boost self-evaluations and affect not just in educational settings, but in a variety of settings in which people make social comparisons.

CHAPTER III: DO GROWTH MINDSETS REDUCE THE BIG-FISH-LITTLE-POND EFFECT?

Research on the big-fish-little-pond effect (BFLPE) indicates that students evaluate themselves more favorably when they have high rank in low rank schools than low rank in high rank schools. Nonetheless, it remains unclear how the detrimental effects of being a little fish in a big pond might be reduced. To address this gap, we conducted two experiments testing whether growth mindsets reduce the BFLPE. Participants in two studies read an article suggesting that intellectual abilities are malleable or fixed prior to a standard manipulation of the BFLPE (total $N = 539$). Results yielded a highly robust BFLPE on self-evaluations and affect (ds > 1.02). However, the BFLPE was not substantially altered by growth mindsets in either study. Further, despite using parallel materials and measures, the growth mindset manipulation was somewhat less impactful in the present study than in prior studies. These data provide the first experimental evidence suggesting that the BFLPE is robust to growth mindset interventions. Implications for research on the BFLPE and growth mindset interventions more generally are discussed.

Introduction

Social comparison theories have long stipulated that people frequently compare their own attributes and abilities to those of relevant peers, and that these comparisons have a pronounced impact on affect and cognition (Festinger, 1954; Wheeler & Suls, 2020). Consistent with this notion, considerable research indicates that people report more positive self-evaluations and mood following downward comparisons with inferior others than upward comparisons with superior others (Alicke et al., 2013; Gerber et al., 2018). The big-fish-little-pond effect (BFLPE) is one of the most studied manifestations of social comparison. According to dozens of educational studies (e.g., Marsh, 1987; Marsh et al., 2015) and several experiments on this

phenomenon (McFarland & Buehler, 1995; Zell & Lesick, 2021), students evaluate themselves more favorably when they have high rank in low rank schools than low rank in high rank schools. The BFLPE occurs because students focus on their rank in local groups during self-evaluation, but neglect more general information indicating the rank of their group relative to other groups (i.e., local dominance; Zell & Alicke, 2010, 2020).

Although the BFLPE is beneficial for students with high rank in low rank schools, it is detrimental for students with low rank in high rank schools. That is, students who are little fish in big ponds experience deflated self-evaluations, mood, and aspirations (Nagengast & Marsh, 2012; Pekrun et al., 2019), which may contribute to decrements in performance and productivity (Marsh & Martin, 2011; Valentine et al., 2004). Thus, it is important to identify factors that reduce the harmful effects of being a little fish in a big pond. Along these lines, scholars have recently noted that there is a need to "more fully recognize the negative implications of the BFLPE and to pursue research evaluating strategies designed to counter the negative effects" (Marsh et al., 2019, p. 237).

Here we test whether adopting a growth mindset—that is, viewing one's abilities as malleable and capable of change (Burnette et al., 2023a; Dweck, 2008)—reduces the BFLPE. On the one hand, prior experiments found that activating a growth mindset reduces the adverse impact of failure experiences on affect and motivation (Park & Kim, 2015; Plaks & Stecher, 2007; Snyder et al., 2014). On the other hand, one study in an educational context found no moderation of the BFLPE by individual differences in growth mindsets (Marsh et al., 2021). Thus, it remains unclear whether the harmful effects of being a little fish in a big pond are ameliorated by growth mindsets. We address this gap by reporting the first set of experiments testing whether activation of a growth mindset reduces the BFLPE.

Big-Fish-Little-Pond Effect

The BFLPE is the tendency for students to perceive their academic abilities more favorably when they have high rank in low rank schools or classes than low rank in high rank schools or classes (Marsh & Parker, 1984; Marsh & Seaton, 2015). Extensive research in educational settings suggests that the BFLPE emerges when examining academic self-concepts (Koivuhovi et al., 2022; Marsh et al., 2019, 2021). In a representative study of 5th and 6th grade students in Germany (Pekrun et al., 2019, Study 2), there was a negative association between class-average math performance and math self-concepts, after adjusting for individual student performance. In other words, average students who were surrounded by strong performers reported less positive math self-concepts, whereas average students surrounded by poorer performers reported more positive math self-concepts. The BFLPE is often interpreted to suggest that all else being equal, students with high rank in low rank groups evaluate their performance and ability more favorably than students with low rank in high rank groups (Zell & Alicke, 2010, 2020).

Beyond self-evaluations, research suggests that the BFLPE extends to achievement emotions (Basarkod et al., 2023; Pekrun et al., 2019) and aspirations (Nagengast & Marsh, 2012). Moreover, educational studies suggest that the BFLPE is robust across gender, age, and cultural groups (Fang et al., 2018; Marsh et al., 2015). Most studies on the BFLPE are cross-sectional, but prospective studies suggest that the effect remains two to four years after graduation (Marsh et al., 2007) and longitudinal studies suggest that the BFLPE is associated with changes in self-evaluations and achievement emotions across two years (Koivuhovi et al., 2022; von Keyserlingk et al., 2019). Although less numerous, experimental studies provide causal evidence for the BFLPE (Alicke et al., 2010; Zell & Alicke, 2009), and both experimental

and non-experimental studies suggest that the effect occurs because people neglect group rank information during self-evaluation (Huguet et al., 2009; Marsh et al., 2014; Zell & Lesick, 2021).

In short, whereas being a big fish in a little pond (BFLP) inflates self-evaluations and mood, being a little fish in a big pond (LFBP) deflates them. Given these adverse impacts, scholars have called for research on strategies to reduce the BFLPE in high performing groups where students are regularly demoralized by comparisons with "superstar" peers (Marsh et al., 2021; Marsh & Seaton, 2015). If identified, successful interventions could be used to boost the motivation and performance of students grappling with the challenges of being an LFBP. Surprisingly, however, few experiments have tested strategies to the reduce the BFLPE. As one of the first attempts to fill this gap, we examine whether adopting a growth mindset attenuates detrimental impacts of the BLFPE on self-evaluations and mood.

Growth Mindsets

Whereas people with a fixed mindset view their personality and abilities as set in stone, people with a growth mindset view these same dimensions as malleable and capable of improvement (Burnette et al., 2023a; Dweck, 2008). For example, a student with a fixed mindset likely perceives her intelligence as unchangeable and perhaps determined by genetic factors. Conversely, a student with a growth mindset likely perceives her intelligence as improvable with sustained effort and training. Research in educational settings demonstrates that individual differences in growth mindsets are associated with greater resilience and achievement (Burnette et al., 2023b; Yeager & Dweck, 2020; Zhao et al., 2021). For example, early research found that students with a growth mindset regarding intelligence were less likely to attribute a poor intellectual performance to a lack of ability—and were more likely to seek methods to improve their future performance—than students with a fixed mindset (Hong et al., 1999). Further, more

recent work in a representative sample of 10th graders in Chile found that growth mindsets regarding intelligence were associated with standardized test performance in math and language, and that this effect was comparable in size to the effects of socioeconomic status (Claro et al., 2016).

In addition to individual differences in growth mindsets, research also suggests that activating growth mindsets has beneficial effects. For example, one study examined the impact of a growth mindset intervention, specifying that intellectual abilities can be developed with effort, in a nationally representative sample of 9th graders in the US (Yeager et al., 2019). Lower-achieving students who received the intervention earned better grades in core classes such as math, science, and English than lower-achieving students in a control condition. More generally, a recent meta-analysis found statistically significant impacts of growth mindset interventions on academic performance, mental health, and social functioning, with effect sizes (*d*s) ranging from .09 to .36 (Burnette et al., 2023a). There remains debate, however, regarding the size and replicability of growth mindset intervention effects (Brez et al., 2020; Macnamara & Burgoyne, 2023), with more work needed to understand the conditions under which such effects are most pronounced (Miller, 2019; Yeager et al., 2019).

Can Growth Mindsets Reduce the BFLPE?

Although little research has directly tested whether growth mindsets offset the harmful effects of being an LFBP, relevant research suggests that growth mindset interventions may promote adaptive responses to poor performance. In one project, for example, participants read a brief article suggesting that intelligence is fixed or malleable and then received feedback indicating that their performance on an intelligence test was stable, improved, or declined across two tests (Plaks & Stecher, 2007, Study 2). The harmful impact of declining feedback on anxiety

and effort on a third test was less pronounced among students in the growth mindset condition. In a related project, participants read a brief article suggesting that intelligence is fixed or malleable, received feedback indicating that their performance on an intelligence test was below average, and completed another intelligence test (Park & Kim, 2015, Study 4). Participants in the growth mindset condition reported fewer self-critical thoughts and performed better on the second intelligence test than participants in the fixed mindset condition.

Data from two additional studies suggest that growth mindset interventions may soften adverse reactions to poor performance. Specifically, in one study, participants were told that problem solving ability is fixed or malleable and that they correctly answered a relatively high or low number of items on a problem-solving test (Snyder et al., 2014). The harmful effect of failure feedback on self-handicapping (i.e., intentionally sabotaging one's performance to provide a convenient excuse for failure) was less pronounced among participants in the growth mindset condition. The other study had participants read a brief article suggesting that math ability is fixed or malleable and then read about a 7th grade student who performed poorly on a math test (Rattan et al., 2012, Study 2). Participants in the growth mindset condition were less likely to perceive the student as having low math ability and were less likely to endorse comfort-oriented strategies that might reduce the student's achievement in math (e.g., assigning the student less math homework).

In sum, although not examining the BFLPE specifically, research suggests that people may react more adaptively to failure following activation of a growth mindset. To the extent that being an LFBP is experienced as a failure (Zell & Alicke, 2010, 2020), one might expect growth mindsets to also reduce negative self-evaluations and affect in these conditions. As of present, only one study has directly tested whether growth mindsets reduce the BFLPE (Marsh et al.,

2021). Specifically, a sample of East Asian secondary school students indicated the degree to which they thought their math ability was fixed versus malleable. In contrast with prior studies on the benefits of growth mindsets, results found a robust BFLPE on math self-concepts, and this effect was not significantly moderated by mindsets regarding math ability. Although this study was limited in that did not manipulate growth mindsets or the BFLPE, it suggests that individual differences in growth mindsets may not reduce the BFLPE.

The Current Research

There is extensive support for the robustness and generalizability of the BFLPE, including its tendency to deflate self-evaluations and emotions of students who are an LFBP (Marsh et al., 2015; Pekrun et al., 2019). However, interventions that might reduce these adverse impacts have yet to be explored despite calls for research on this topic (Marsh et al., 2019). Activation of a growth mindset has been found to yield a host of beneficial effects in educational settings (Burnette et al., 2023a; Yeager et al., 2019), including the tendency to soften adverse affective and motivational reactions to failure (Park & Kim, 2015; Plaks & Stecher, 2007; Snyder et al., 2014). In two studies, we examined whether activation of a growth mindset softens adverse impacts of the BFLPE on self-evaluations and emotions.

As noted above, one prior study found that individual differences growth mindsets failed to reduce the BFLPE when studied in an educational setting (Marsh et al., 2021). However, this study did not manipulate growth mindsets or the BFLPE, which prevents any causal conclusions about the impact (or lack thereof) of growth mindsets on the BFLPE. In two experiments, we provide the first direct test of whether activation of a growth versus a fixed mindset regarding intellectual abilities reduces adverse effects of the BFLPE. Given the conflicting findings and novelty of our research question, we remained open to two possibilities: (1) growth mindsets

moderate the BFLPE, by reducing the deflating effect of being an LFBP on self-evaluations and emotions (i.e., reducing the impact of failure feedback; Park & Kim, 2015; Plaks & Stecher, 2007; Snyder et al., 2014) or (2) that growth mindsets would not substantially moderate the BFLPE and thus the deflating effect of being an LFBP emerges regardless of whether people adopt a growth versus fixed mindset (Marsh et al., 2021).

We report how we determined our sample sizes and all data exclusions, manipulations, and measures in these studies. Sample sizes for each study were determined prior to data collection. Each study was pre-registered. All materials, data, analysis code, pre-registrations, and supplemental analyses for each study are publicly available on OSF. This research was approved by the Institutional Review Board (IRB) at the authors' university. All participants provided informed consent prior to their participation.

Study B1

In our first study, we tested whether the BFLPE was reduced following activation of a growth versus fixed mindset. Specifically, participants read an article indicating that intelligence is either malleable or fixed (Plaks & Stecher, 2007, Study 2). Then, participants completed a verbal reasoning test and received feedback indicating that they were either a BFLP or an LFBP. This design allowed us to examine whether the impact of performance feedback on self-evaluations and affect was altered by mindsets.

Method

Participants and design. Participants were 295 undergraduates at a university in the Southeastern United States, who completed the 30-min study for credit in introductory psychology courses during the Fall 2022 semester. According to our pre-registered plan, we collected data until we obtained at least 259 participants who passed attention, feedback, and

suspicion checks. This allowed us to achieve at least 80% power to detect a small-to-medium effect (f = .175), which was the smallest effect we were interested in detecting. After excluding 29 participants who failed one or more of the checks (exclusion details below), our final sample contained 266 participants (195 women, 91 White, 93 Black, M_{age} = 18.8). Participants were randomly assigned to one condition in a 2 (mindset: growth, fixed) by 2 (feedback: BFLP, LFBP) between-subjects design.

Procedures. Participants completed all study measures online. To begin, participants were told that the purpose of the study was to measure the academic abilities of students at their university, and that they would complete a reading comprehension test and verbal reasoning test as part of this effort. Both tests were described as valid measures of academic abilities that were predictive of success in a variety of occupations.

Next, participants completed what was ostensibly a reading comprehension test, which provided the mindset manipulation. Specifically, participants were presented with a brief article that was supposedly published in "Psychology Now" and that was adapted from previous research (Park & Kim, 2015; Plaks & Stecher, 2007; Rattan et al., 2012). In the *growth mindset conditions*, participants read an article that described intelligence as malleable and presented research indicating that intelligence is altered by environmental factors and training. In the *fixed mindset conditions*, participants read an article that described intelligence as highly stable across time and presented research indicating that intelligence is not altered by educational interventions or training. Similar types of evidence were presented in both articles (e.g., longitudinal studies, case studies, intervention programs) but findings were modified to endorse either a growth or fixed mindset.

After reading the article, participants completed a 10-item reading comprehension test, indicating whether a series of statement were true or false according to the article they just read (e.g., *Intellectual abilities are malleable and develop significantly over time*). As a manipulation check, participants provided their personal belief about stability of intelligence (*To what extent do you believe that a person's intelligence is stable?*), using a 1 (*not at all stable*) to 7 (*extremely stable*) scale (Park & Kim, 2015; Plaks & Stecher, 2007; Rattan et al., 2012). For exploratory purposes, participants completed article credibility (*The content of this article is credible*), comprehension difficulty (*The content is easy to understand*), and attention items (*I paid great attention while reading this article*), using 1 (*strongly disagree*) to 7 (*strongly agree*) scales.

Participants next completed a verbal reasoning test consisting of 40 sentence-completion items. For each item, participants were presented with a sentence that was missing a word (e.g., *The modern age is a permissive one in which things can be said explicitly, but the old tradition of _____ dies hard*) and were asked to select which of five words best completed the sentence (*garrulousness*, *exaggeration*, *excoriation*, *bombast*, *euphemism*; the correct answer is *euphemism*). Prior research found this test to be sufficiently ambiguous to promote the believability of manipulated performance feedback (Zell & Lesick, 2021; Zell & Strickhouser, 2020). Prior research also found that participants on average completed the test in 16.0 minutes (*SD* = 4.6 minutes).

After the test, participants received manipulated feedback about their performance (Stockus & Zell, 2023b; Zell & Lesick, 2021). In the *BFLP conditions*, participants were told that they ranked better than 68% of 206 students at their university and that their university ranked better than 32% of 39 universities. In the *LFBP conditions*, participants were told that they ranked better than 32% of 206 students at their university and that their university ranked

better than 68% of 39 universities. Summary information indicated that the participant ranked above (below) average in a below (above) average school. To promote comprehension, participants recorded the feedback and summary information they received. The order of student and school feedback was counterbalanced.

Participants then completed the primary outcome measures of verbal self-evaluations and performance-related affect (Stockus & Zell, 2023b; Zell & Strickhouser, 2020). Specifically, participants completed a 3-item measure of self-evaluations (e.g., *How well do you think you performed on the Verbal Reasoning Test?*; $\alpha = .88$), using 1 (*very poorly [bad]*) to 7 (*very well [good]*) scales. Participants also completed a 6-item measure of performance-related affect ($\alpha = .90$), indicating how *satisfied*, *proud*, *sad*, *discouraged*, *happy*, and *distressed* they felt about their test performance (e.g., *How satisfied do you feel about your Verbal performance?*), using 1 (*not at all*) to 7 (*extremely*) scales. Negative affect items were reverse scored, so that higher values indicate more favorable affect.

Lastly, participants completed demographic items (e.g., age, race, and gender), a feedback manipulation check, which asked them to recall the feedback they received about their performance (*I ranked above [below] average* and UNCG *ranked below [above] average*), an attention check, which asked them to select the fifth circle in a 7-point scale, and a suspicion probe. Participants who expressed suspicion about the provided feedback ("My ranking seemed a bit too good to be true") were excluded from data analyses. Finally, participants were debriefed and awarded credit.

Participant exclusions. We excluded data from 18 participants for failing the attention check, 6 participants for failing the manipulation check, and 5 participants for expressing

suspicion about the feedback. Results were very similar when excluded participants were retained (see the Supplemental Analyses).

Results

Mindset manipulation check. A 2 (mindset: growth, fixed) by 2 (feedback: BFLP, LFBP) ANOVA yielded a significant main effect of mindset, $F(1, 262) = 30.75$, $p < .001$, $\eta_p^2 = .11$ (see Figure B1). As expected, participants believed that intelligence was more stable in the fixed mindset condition ($M = 4.79$, $SD = 1.03$) than the growth mindset condition ($M = 4.03$, $SD = 1.25$). This indicates that the mindset manipulation was successful, and was medium-to-large in size, $d = 0.67$. Further, the main effect of feedback and the two-way interaction were both non-significant, $F(1, 262) = 0.06$, $p = .807$, $\eta_p^2 < .01$ and $F(1, 262) = 3.47$, $p = .064$, $\eta_p^2 = .01$.

Self-evaluations. A 2 by 2 ANOVA yielded a very large main effect of feedback, $F(1, 262) = 130.73$, $p < .001$, $\eta_p^2 = .33$ (see Figure B2). Consistent with the BFLPE, self-evaluations were much more pronounced in the BFLP conditions ($M = 4.55$, $SD = 1.17$) than the LFBP conditions ($M = 2.88$, $SD = 1.20$). The main effect of mindset and the two-way interaction were both non-significant, $F(1, 262) = 0.34$, $p = .559$, $\eta_p^2 < .01$ and $F(1, 262) = 2.12$, $p = .146$, $\eta_p^2 < .01$. Planned comparisons found that the BFLPE was very large in both the fixed mindset condition, $t(262) = 6.99$, $p < .001$, $d = 1.23$ and the growth mindset condition, $t(262) = 9.20$, $p < .001$, $d = 1.58$. Moreover, the mindset manipulation did not have a significant impact on self-evaluations in either the BFLP condition, $t(262) = 1.48$, $p = .139$, $d = 0.25$, or the LFBP condition, $t(262) = 0.60$, $p = .548$, $d = 0.11$.

Affect. Parallel results were obtained for affect (see Figure B3). Specifically, a 2 by 2 ANOVA yielded a very large main effect of feedback, $F(1, 262) = 132.81$, $p < .001$, $\eta_p^2 = .34$, but no significant main effect of mindset or two-way interaction, $F(1, 262) = 3.01$, $p = .084$, η_p^2

= .01 and $F(1, 262) = 0.19$, $p = .663$, $\eta_p^2 < .01$. The BFLPE was highly robust across both fixed and growth mindset conditions, $t(262) = 7.77$, $p < .001$, $d = 1.36$ and $t(262) = 8.54$, $p < .001$, $d = 1.47$. Further, the mindset manipulation did not have a significant influence on affect either in the BFLP condition or the LFBP condition, $t(262) = 1.58$, $p = .116$, $d = 0.27$ and $t(262) = 0.90$, $p = .372$, $d = 0.16$.

Exploratory analyses. Reading comprehension was significantly better in the growth mindset condition ($M = 7.92$, $SD = 1.41$) than the fixed mindset condition ($M = 5.95$, $SD = 2.41$), $t(264) = 8.13$, $p < .001$, $d = 1.00$, which suggests that the growth mindset article may have been more intuitive. However, there were no significant differences between the growth and fixed mindset conditions with respect to participant ratings of article credibility, $t(264) = 1.25$, $p = .212$, $d = 0.15$, comprehension difficulty, $t(264) = 1.54$, $p = .125$, $d = 0.19$, or amount of attention paid to the article, $t(264) = 1.45$, $p = .148$, $d = 0.18$ (see the Supplemental Analyses). In addition to the categorical analyses of growth mindsets reported above, we also tested whether individual differences in growth mindsets via responses to the manipulation check moderated the BFLPE. Results for these analyses are very similar to the primary analyses we report above (see the Supplemental Analyses).

Discussion

Study B1 provided the first experimental test of whether growth mindsets reduce the BFLPE. Consistent with previous research (Stockus & Zell, 2023b; Zell & Lesick, 2021), a very large BFLPE obtained on both self-evaluations and affect. Furthermore, these effects were generally similar across growth and fixed mindset conditions. Thus, our initial study suggests that the BFLPE is robust to growth mindsets. Nonetheless, despite using a similar manipulation of growth mindsets and manipulation check, the effect of the growth mindset manipulation on

the perceived stability of intelligence was somewhat smaller in Study B1 (d = 0.67) than in related studies by Park and Kim (2015; d = 1.26), Plaks and Stecher (2007; d = 2.34), and Rattan and colleagues (2012; d = 1.41). To address this issue, we strengthened the manipulation of growth mindsets in our next study.

Study B2

Study B2 once again tested whether the magnitude of the BFLPE was reduced following activation of a growth versus fixed mindset. As in Study B1, participants read an article indicating that intelligence is either malleable or fixed as part of a reading comprehension test, completed a verbal reasoning test, and received manipulated feedback indicating that they were either a BFLP or an LFBP on the verbal test. Further, to strengthen the mindset manipulation, participants were given corrective feedback on article comprehension items that were dispersed throughout the article.

Method

Participants and design. Participants were 299 undergraduates who completed the study for course credit during the Spring 2023 and Fall 2023 semesters. After excluding 26 participants who failed attention, feedback manipulation, or suspicion checks, our final sample contained 273 participants (193 women, 95 White, 90 Black, M_{age} = 19.3). As in Study B1, participants were randomly assigned to one condition in a 2 (mindset: growth, fixed) by 2 (feedback: BFLP, LFBP) between-subjects design.

Procedures. Unless noted otherwise, Study B2 procedures were identical to Study B1. Specifically, participants read one of two articles describing intelligence as either malleable or fixed. To bolster the impact of the article, participants responded to 10 true/false statements about its content that were dispersed evenly across various sections of the article. Incorrect

responses were flagged, and participants had to correct them before continuing. These statements were identical to the comprehension test items used in Study B1, but were administered during rather than after the article, and used to provide corrective feedback rather than measure participants' actual knowledge. All remaining portions of the study were identical to Study B1. That is, participants completed a mindset manipulation check, exploratory items in response to the article, and a verbal reasoning test. After the test, participants received feedback indicating that they were either a BFLP or an LFBP and completed measures of verbal self-evaluations (α = .82) and performance-related affect (α = .85). Lastly, participants completed demographic items and a feedback manipulation check, attention check, and suspicion probe.

Participant exclusions. We excluded data from 8 participants for failing the attention check, 11 participants for failing the manipulation check, and 7 participants for expressing suspicion about the feedback. As in Study B1, results were very similar when excluded participants were retained (see the Supplemental Analyses).

Results

Mindset manipulation check. A 2 (mindset: growth, fixed) by 2 (feedback: BFLP, LFBP) ANOVA yielded a significant main effect of mindset, $F(1, 269) = 100.05$, $p < .001$, $\eta_p^2 = .27$. Participants believed that intelligence was more stable in the fixed mindset condition ($M = 4.82$, $SD = 1.18$) than the growth mindset condition ($M = 3.35$, $SD = 1.25$), with an effect size that was 82% larger than our first study, $d = 1.22$. The main effect of feedback and the two-way interaction were both non-significant, $F(1, 269) = 3.73$, $p = .055$, $\eta_p^2 = .01$ and $F(1, 269) = 0.45$, $p = .503$, $\eta_p^2 < .01$.

Self-evaluations. A 2 (mindset) by 2 (feedback) ANOVA yielded a very large main effect of feedback, $F(1, 269) = 120.27$, $p < .001$, $\eta_p^2 = .31$. Consistent with the BFLPE, self-

evaluations were much more pronounced in the BFLP conditions (M = 4.23, SD = 1.25) than the LFBP conditions (M = 2.75, SD = 0.96). The main effect of mindset and the two-way interaction were both non-significant, $F(1, 269) = 0.07$, $p = .794$, $\eta_p^2 < .01$ and $F(1, 269) = 0.04$, $p = .833$, $\eta_p^2 < .01$. Planned comparisons found that the BFLPE was very large in both the fixed mindset condition, $t(269) = 7.84$, $p < .001$, $d = 1.35$ and the growth mindset condition, $t(269) = 7.67$, $p < .001$, $d = 1.30$. Moreover, the mindset manipulation did not have a significant impact on self-evaluations in either the BFLP condition, $t(269) = 0.04$, $p = .972$, $d = 0.01$, or the LFBP condition, $t(269) = 0.34$, $p = .737$, $d = 0.06$.

Affect. Parallel results were obtained for affect. Specifically, a 2 (mindset) by 2 (feedback) ANOVA yielded a very large main effect of feedback, $F(1, 269) = 91.29$, $p < .001$, $\eta_p^2 = .25$, but no significant main effect of mindset or two-way interaction, $F(1, 269) = 0.00$, $p = .955$, $\eta_p^2 < .01$ and $F(1, 269) = 1.16$, $p = .282$, $\eta_p^2 < .01$. The BFLPE was highly robust across both fixed and growth mindset conditions, $t(269) = 7.45$, $p < .001$, $d = 1.29$ and $t(269) = 6.05$, $p < .001$, $d = 1.03$. Further, the mindset manipulation did not have a significant influence on affect either in the BFLP or LFBP conditions, $t(269) = –0.80$, $p = .428$, $d = 0.14$ and $t(269) = 0.73$, $p = .467$, $d = 0.12$.

Exploratory analyses. Unlike Study B1, significant differences emerged between the growth and fixed mindset conditions with respect to participant ratings of article credibility, $t(271) = 3.88$, $p < .001$, $d = 0.47$, comprehension difficulty, $t(271) = 4.24$, $p < .001$, $d = 0.51$, and amount of attention paid to the article, $t(271) = 3.52$, $p < .001$, $d = 0.43$ (see the Supplemental Analyses). Participants in the growth mindset conditions reported that the article was more credible, easier to understand, and that they paid more attention to it than participants in the fixed mindset conditions. Further, as in Study B1, individual differences in growth mindsets beliefs

assessed via the manipulation check did not significantly moderate the BFLPE (see the Supplemental Analyses).

Discussion

Study B2 replicated the results of Study B1. That is, although we effectively strengthened the mindset manipulation in Study B2, the BFLPE was once again robust to activation of growth versus fixed mindsets.

Combined Analyses

To maximize statistical power, we conducted combined analyses across Studies B1-B2 for the primary outcome measures (total $N = 539$, after exclusions). As above, a 2 (mindset) by 2 (feedback) ANOVA on self-evaluations yielded a very large main effect of feedback, $F(1, 535) = 252.19$, $p < .001$, $\eta_p^2 = .32$, but no significant main effect of mindset condition or interaction, $F(1, 535) = 0.42$, $p = .520$, $\eta_p^2 < .01$ and $F(1, 535) = 0.84$, $p = .360$, $\eta_p^2 < .01$. Similarly, for performance-related affect, we found a very large main effect of feedback, $F(1, 535) = 222.29$, $p < .001$, $\eta_p^2 = .29$, but no significant main effect of mindset condition or interaction, $F(1, 535) = 1.55$, $p = .214$, $\eta_p^2 < .01$ and $F(1, 535) = 0.15$, $p = .696$, $\eta_p^2 < .01$. Thus, when looking across Studies B1 and B2, activation of a growth mindset did not substantially alter the BFLPE. Moreover, when examined across studies, individual differences in growth mindset beliefs did not significantly moderate the BFLPE (see the Supplemental Analyses).

General Discussion

The BFLPE is the tendency for students to evaluate themselves more favorably and report more favorable emotions when they have high rank in low rank schools than low rank in high rank schools (Koivuhovi et al., 2022; Pekrun et al., 2019). Substantial evidence supports the generalizability and robustness of the BFLPE (Fang et al., 2018; Marsh et al., 2021). However,

little is known about factors that might reduce harmful effects of being an LFBP, despite repeated calls for research on this topic (e.g., Marsh et al., 2019). In two studies, we tested whether the harmful effects of being an LFBP are reduced following activation of a growth versus fixed mindset. Specifically, we tested whether inducing the belief that intelligence is malleable buffers the negative impact of being an LFBP on self-evaluations and affect.

In both studies, participants completed a standard manipulation of growth versus fixed mindsets (Park & Kim, 2015; Plaks & Stecher, 2007; Rattan et al., 2012), where they read a research article ostensibly indicating that intelligence is fixed or malleable. Participants then received feedback indicating that they had high rank in a low rank school (BFLP) or low rank in a high rank school (LFBP) on a verbal reasoning test. A highly robust BFLPE obtained on self-evaluations and performance-related affect in both studies. However, analyses within and across the two studies indicated that activation of a growth versus fixed mindset did not substantially alter the BFLPE. In sum, our data suggest that growth mindset interventions may not be sufficient to dislodge harmful effects of being an LFBP on self-evaluations and mood.

Implications

The above studies contribute to research on the BFLPE and work on growth mindsets. First, the present research joins only one prior study to examine whether growth mindsets reduce the BFLPE. This prior study found that individual differences in growth mindsets do not substantially moderate the BFLPE when tested in an educational context in East Asia (Marsh et al., 2021). That is, a robust BFLPE obtained on math self-concepts regardless of whether students perceived their math ability as fixed or malleable. Nonetheless, because this prior work used non-experimental methods, no conclusions could be drawn regarding the potential impact of growth mindset interventions on the BFLPE.

Addressing this concern, the present experiments provide the first tests of whether activation of a growth versus fixed mindset moderates the BFLPE. As in prior experiments (Stockus & Zell, 2023b; Zell & Lesick, 2021), we found robust causal evidence for the BFLPE on self-evaluations and mood. We also found, consistent with prior work, that a growth mindset reading exercise successfully increases endorsement of a growth versus fixed mindset regarding intelligence (e.g., Plaks & Stecher, 2007). However, the BFLPE was highly robust across growth versus fixed mindset conditions and did not substantially vary across these conditions. Thus, our studies provide the first evidence suggesting that growth mindset interventions may not be a viable strategy to reduce the BFLPE.

Second, our research contributes to an emerging literature on whether growth mindset interventions soften adverse reactions to failure. Along these lines, prior research found that activation of a growth mindset reduces adverse effects of negative performance feedback on anxiety (Plaks & Stecher, 2007), self-critical thoughts (Park & Kim, 2015), and self-handicapping (Snyder et al., 2014) or other-handicapping strategies (Rattan et al., 2012). Prior research indicates that being an LFBP sharply drops self-evaluations and mood, and should therefore be conceptualized as a failure experience (Zell & Alicke, 2010, 2020). Nonetheless, the present research found that activation of a growth mindset did not soften adverse reactions to the BFLPE. Thus, the present findings suggest that some failure experiences may be sufficiently robust to resist potential effects of growth mindset interventions. As such, the present findings provide a counterpoint to emerging findings in the literature demonstrating beneficial effects of growth mindsets in the context of failure. Further, they contribute to ongoing debates about the size and replicability of growth mindset effects (Brez et al., 2020; Macnamara & Burgoyne, 2023).

Third, this research contributes to emerging experiments on the activation of growth versus fixed mindsets. Although the present research used a similar manipulation of growth versus fixed mindsets and manipulation check as past work (Park & Kim, 2015; Plaks & Stecher, 2007; Rattan et al., 2012), the impact of the manipulation in the present work was somewhat smaller than in some prior studies. When viewed in combination with past work, this finding suggests that the effect of standard growth mindset interventions may vary substantially across samples and contexts. Additionally, we found some evidence that an article on growth mindsets elicited better recall (Study B1) and was perceived as more credible, attention grabbing, and easier to understand (Study B2) than an article on fixed mindsets. Although more work is needed in other samples and contexts, these data suggest that growth mindset manipulations may be more persuasive than fixed mindset manipulations. Thus, more refinement may be needed in the development of growth versus fixed mindset manipulations, and perhaps other comparison conditions rather than fixed mindsets may be useful to explore in future work (e.g., Yeager et al., 2019).

Fourth, and more generally, the present research contributes to emerging work on potential moderators of the BFLPE (Marsh & Seaton, 2015; Marsh et al., 2021). Prior study identified a few student factors (i.e., age, country, gender) and teacher factors (i.e., student-teacher relationships, differentiated instruction) that consistently moderate the BFLPE (Stockus & Zell, 2024b). Little work, however, has examined contextual factors or interventions that alter the BFLPE. In testing the impact of growth versus fixed mindsets on the BFLPE, the present work developed a rigorous experimental protocol that could be used to test other contextual moderators. Further, by publicly sharing the materials used in this protocol, the present work should stimulate additional study on moderators.

Limitations and Future Directions

A few limitations of the present work could be profitably explored in future study. The present research was conducted in a university setting and in a single region in the United States. Future research in other academic settings (e.g., K-12 schools) and in other regions, countries, and cultures is needed to further explore whether growth mindsets moderate the BFLPE. Although growth mindsets may not generally reduce the BFLPE, it remains possible that growth mindsets reduce the BFLPE under specific conditions or contexts, for example, when examining lower-achieving students or when school norms advocate persistence in the face of academic difficulties (Yeager et al., 2019). Thus, future research should seek to identify optimal conditions in which growth mindsets reduce the BFLPE.

The present research was also conducted online. Although the present work found that growth versus fixed mindsets can be successfully activated using Internet methods, as have prior studies (e.g., Park & Kim, 2015), research is needed to examine whether such manipulations are more (or less) effective when delivered online. Relatedly, we found that the manipulation of growth versus fixed mindsets was somewhat stronger when participants were required to complete comprehension questions that were dispersed throughout the reading activity (Study B2). This approach could be used in future study to increase the effectiveness of the growth mindset manipulations.

Another potential limitation was that the growth mindset manipulation was administered before the verbal reasoning task. Although the verbal task was relatively brief (i.e., about 16 minutes), and prior work suggests that growth mindset effects are durable (Park & Kim, 2015; Plaks & Stecher, 2007), it remains possible that the effect of the growth mindset manipulation faded over time. Thus, future work is needed to explore how long growth mindset manipulations

remain effective and how to re-activate such manipulations, if necessary. Moreover, the present work utilized a standard manipulation of growth mindsets, where participants read and engaged with a short research article. Future study is needed to examine whether other manipulations, for example, those that span a longer length of time or that have greater interactivity (e.g., Yeager et al., 2019), produce stronger effects.

Finally, future work is needed to examine whether alignment between the growth mindset manipulation and performance task helps promote its impact. As in prior studies, the present research utilized a general manipulation of growth mindsets in the broad domain of intellectual abilities, and then had participants complete a specific intellectual ability test (i.e., verbal reasoning; Plaks & Stecher, 2007; Rattan et al., 2012). Nonetheless, it remains possible that growth mindset manipulations have a greater impact when they are precisely linked to the performance task (Snyder et al., 2014). For example, a growth mindset manipulation regarding verbal reasoning could promote adaptive responses to failure on a verbal reasoning test. Nonetheless, a significant downside to this approach is that the two tasks are more obviously linked to participants, thus increasing the odds of suspicion and socially desirable responding (i.e., participants in the growth mindset condition may realize that they should respond more optimistically to failure).

Conclusions

A large research literature on the BFLPE suggests that being an LFBP has adverse impacts on self-evaluations and mood (Pekrun et al., 2019; Zell & Lesick, 2021). The present research tested whether activation of a growth versus fixed mindset softens these adverse impacts. Consistent with a recent non-experimental study conducted in an educational context (Marsh et al., 2021), the present experiments found that the BFLPE is largely resistant to growth

mindsets. Research in other samples and contexts is needed to further evaluate whether the BFLPE is robust to growth mindsets, and to identify other interventions that alter its impacts on self-evaluations, mood, and motivation.

CHAPTER IV: EMPATHIC FORECASTING OF THE BIG-FISH-LITTLE-POND EFFECT

Stockus, C. A., & Zell, E. (2024a). Empathic forecasting of the big-fish-little-pond effect. *Cognition & Emotion*, *38*(1), 44–58. https://doi.org/10.1080/02699931.2023.2258581

The big-fish-little-pond effect (BFLPE) is the tendency for students to evaluate themselves more favorably when they have high rank in a low rank school than low rank in a high rank school. Extensive research has documented the BFLPE on experienced emotions. We conducted three studies that examined forecasts of how the BFLPE influences other people's emotions (i.e., empathic forecasts). In Study C1, participants received performance feedback either about themselves or another person and reported either their own affect or anticipated the other person's affect. Results extended the BFLPE to empathic forecasting. Moreover, forecasters anticipated that the BFLPE had a stronger influence on negative emotion than it actually did. Study C2 tested whether neglect of group rank information underlies the BFLPE in empathic forecasting. As predicted, empathic forecasts were strongly influenced by another person's rank in their group, but only weakly influenced by group rank, resulting in a large BFLPE. Finally, Study C3 tested whether extremity of group ranks exacerbates the BFLPE in empathic forecasting. Consistent with predictions, empathic forecasts were especially favorable (unfavorable) when a target had very high (low) rank in a very low (high) rank group. Taken together, these data strongly support the BFLPE in empathic forecasting, but also illustrate ways in which it is both similar to and different from actual experience of the effect.

Introduction

Nearly 70 years ago, Festinger's (1954) social comparison theory proposed that people compare their own attributes and abilities to those of relevant peers and that these comparisons powerfully shape thoughts and feelings about the self. A large empirical literature now supports

these assertions, demonstrating that people compare themselves to others in numerous domains (Križan & Gibbons, 2014; Suls et al., 2020), especially those involving academic and vocational performance, and that these comparisons strongly impact self-evaluations and emotions (Gerber et al., 2018; Zell & Strickhouser, 2020). Perhaps the most frequently studied manifestation of social comparison is the big-fish-little-pond effect (BFLPE), namely, the tendency for students with high rank in a low rank school to evaluate themselves more favorably, as well as report greater positive emotions and fewer negative emotions, than students with low rank in a high rank school (Marsh et al., 2021; Pekrun et al., 2019). The BFLPE likely occurs because social comparisons with local group members, such as others in one's class or school, overwhelm and negate the affective impact of other standards (i.e., local dominance; Zell & Alicke, 2009, 2010).

Research on the BFLPE has focused on its impact on the self—that is, *how people feel* when they are a big fish in a little pond versus a little fish in a big pond (Pekrun et al., 2019; Zell & Lesick, 2021). Here we test for the first time *how people think others will feel* when they are a big fish in a little pond versus a little fish in a big pond. Prior work on empathic forecasting has examined how people think others will feel in emotionally charged situations (e.g., after failing an exam or contemplating negative stereotypes about one's group) and whether such forecasts match how people actually feel in these situations (Boucher et al., 2015; Pollmann & Finkenauer, 2009). Building upon this framework, the present research explores whether the BFLPE occurs in empathic forecasting and whether empathic forecasts of the BFLPE match actual emotional experiences of people experiencing it. Further, we test why the BFLPE may occur in empathic forecasting and whether extremity of group ranks exacerbates it.

Big-Fish-Little-Pond Effect

A large research literature demonstrates that the BFLPE has a robust impact on how students think and feel about themselves. In an early study, 10th grade males in USA and Australia evaluated their academic abilities more favorably when they had high rank in a low rank school than low rank in a high rank school (Marsh, 1987). Research has since replicated this effect across a variety of cultural groups (e.g., European-American, Asian, and Middle Eastern societies) and age groups (e.g., elementary, middle, high-school, and college students; Fang et al., 2018; Marsh et al., 2019). In addition to its contribution to self-evaluations, research indicates that the BFLPE extends to other outcomes, such as motivation and aspirations (Nagengast & Marsh, 2012; Trautwein et al., 2006). Further, the BFLPE extends to achievement emotions, leading students with high rank in a low rank school to experience higher levels of positive achievement emotions (e.g., enjoyment, pride) and lower levels of negative achievement emotions (e.g., anger, anxiety) than students with low rank in a high rank school (Holm et al., 2020; Pekrun et al., 2019). Although most BFLPE studies are non-experimental, experimental studies provide causal support for its impact on self-evaluations and achievement emotions (Zell & Alicke, 2009).

Research examining underlying mechanisms of the BFLPE suggests that it can be decomposed into two types of social comparison: (1) intragroup comparisons indicating the student's rank relative to others in their class or school, and (2) intergroup comparisons indicating the rank of one's class or school (Huguet et al., 2009; Marsh et al., 2014). Whereas intragroup comparisons with local peers have a substantial impact on self-evaluations and emotions, the impact of intergroup comparisons is much smaller. Thus, students appear to focus on their local standing but neglect the standing of their group (i.e., local dominance; Zell &

Alicke, 2010, 2020). This tendency ironically leads to situations where, despite lower objective performance levels, students with high rank in a low rank group feel better about their performance than students with low rank in a high rank group (Alicke et al., 2010). Moreover, students with very high rank in a very low rank group feel especially good about their performance, likely because they focus on within-group standing (Zell & Lesick, 2021).

So far, research on the BFLPE has examined its consequences for self-evaluations and emotions among students directly experiencing it. Surprisingly, research has not explored forecasts of other people's emotions in contexts where they are a big fish in a little pond versus a little fish in a big pond. Given that people make important decisions based on empathic forecasts (e.g., whether to approach or avoid another person; Green et al., 2013; Moons et al., 2017), it is necessary explore empathic forecasts of the BFLPE. Further, intuitions about the hedonic impact of the BFLPE later influence people's entry decisions (Wu et al., 2018) and may alter the advice they give to others making academic or vocational choices. Along these lines, parents, educators, academic advisors, and career counselors may base advice regarding which school or job to select on forecasts about the BFLPE. Nonetheless, it remains unclear whether people generally appreciate the degree to which being a big fish in a little pond uplifts other people's emotions or the degree to which being a little fish in a big pond depresses them.

More generally, accurately assessing other people's emotions can promote the development and maintenance of close relationships (Le et al., 2020; Sened et al., 2017). People who accurately infer another person's emotions may engage in more effective interactions with that person, resulting in a closer connection. The BFLPE is a pervasive phenomenon, as illustrated by its frequent coverage in the popular press and extensive research attention (Frank, 1985; Gladwell, 2013). Moreover, as noted above, evidence suggests that the BFLPE occurs

across a variety of age groups and regions (Marsh et al., 2019). Thus, it is important to assess whether people generally have accurate forecasts of how this fundamental social comparison experience influences other people's emotions.

Empathic Forecasting

Empathic forecasting is the prediction of other people's affective state in response to a specific emotionally charged event (Boucher et al., 2015; Pollmann & Finkenauer, 2009). For example, people may predict another person's affective state after receiving a promotion, forming a new romantic relationship, failing a test, or experiencing the death of a loved one. Empathic forecasting can be distinguished from related work on empathic accuracy (e.g., Atzil-Slonim et al., 2019; Hodges & Kezer, 2021). Whereas research on empathic accuracy generally explores the accuracy of people's inferences regarding other's thoughts and feelings as they unfold over time, empathic forecasting research specifically examines whether people accurately estimate, overestimate, or underestimate other people's emotions in response to a concrete stimulus or event. Nonetheless, the two literatures are related in their interest in how well people assess other's emotional states.

Similar to research suggesting that people overestimate the impact of emotionally charged events on *their own affect* (i.e., affective forecasting; Gilbert et al., 1998; Lench et al., 2019), research on empathic forecasting suggests that people overestimate the impact of emotionally charged events on *other people's affect*. In one set of studies, for example, participants forecasted the affect of another person in response to positive or negative feedback on a cognitive ability test, which was later compared with that person's actual affect (Pollmann & Finkenauer, 2009). Participants overestimated the affective impact of the feedback, estimating that positive feedback had a stronger beneficial effect, and that negative feedback had a stronger

detrimental effect, than it actually did. Further, this bias occurred both in empathic forecasts for a friend and a stranger, suggesting that it extends to a variety of targets.

Two studies focusing on negative events further support the argument that people overestimate the impact of emotionally charged events on other people's emotions. In one of these studies, participants forecasted the affective experience of women who took a difficult math test after being told that women are generally worse than men at math (i.e., stereotype threat; Boucher et al., 2015). These forecasts were later compared with the actual experience of women in the same situation. As expected, forecasters significantly overestimated the intensity of negative affect experienced by women, assuming that women would be more anxious in this situation than they actually were. In the other study, participants forecasted the degree to which their romantic partner would experience sadness following several relationship transgressions (e.g., physical aggression, failing to provide support when needed; Green et al., 2013). These forecasts were later compared with affective responses of partners who had actually experienced each transgression. As expected, forecasters significantly overestimated the amount of sadness that was reported by romantic partners.

More generally, research indicates that empathic forecasts can be biased by aspects of both the perceiver and target. For example, research on target effects had participants forecast the amount of anger experienced by members of different gender and racial groups in several scenarios (e.g., being insulted by a stranger; Moons et al., 2017). Participants forecasted that men would be angrier than women and that Black men would be angrier than White men. These findings are buttressed by other work indicating that stereotypes about racial and political groups bias empathic forecasts in a stereotype consistent direction (Deska et al., 2020; Lau & Morewedge, 2016). Research examining perceiver effects suggests that people who are high

versus low in social anxiety forecast that other people will have stronger negative affective reactions to their social blunders (Hall et al., 2018, 2020). Further, people who react calmly to distressing events assume that other people will also react calmly (Qian et al., 2020). Thus, empathic forecasts may in part reflect projections of one's own emotions onto others (Atzil-Slonim et al., 2019; Gonzalez et al., 2021; Igou, 2008).

Altogether, research suggests that empathic forecasts are susceptible to several biasing influences. People often overestimate the intensity of other people's affect, especially negative affect, and aspects of both the perceiver and target can further bias empathic forecasts. Nonetheless, although research has explored empathic forecasts in several contexts, it has not examined empathic forecasts of the BFLPE. That is, research has yet to explore whether people recognize the hedonic benefits of being a big fish in a little pond or hedonic costs of being a little fish in a big pond. Moreover, no work has tested whether estimates of other people's emotion in these circumstances match the emotions of people who actually experience them.

The Current Research

The present research examines forecasts of other people's emotions when they have high rank in a low rank group versus low rank in a high rank group, and in doing so, tests 3 interrelated questions. First, we tested whether forecasts of other people's emotions resulting from the BFLPE match actual emotions produced by it (Study C1). Given the robustness of the BFLPE on one's own emotions (Pekrun et al., 2019; Zell & Lesick, 2021), and its frequent discussion in the popular press (Frank, 1985; Gladwell, 2013), we anticipated that forecasts of other people's emotions would be more positive when others had high rank in a low rank group than low rank in a high rank group. Additionally, informed by prior work suggesting that people overestimate the impact of emotionally charged events on other people's emotions (Pollmann &

Finkenauer, 2009), we anticipated that people would overestimate the affective impact of the BFLPE. More specifically, past work suggests that negative events exert less influence than forecasters anticipate (Boucher et al., 2015; Green et al., 2013). As such, forecasters might overestimate the adverse effect of having low rank in a high rank group, but not necessarily the uplifting effect of having high rank in a low rank group.

Second, focusing on an underlying mechanism, our next study explored why the BFLPE may occur in empathic forecasting (Study C2). Prior work on experienced emotions suggests that affective reactions are strongly influenced by people's rank within groups, but only weakly influenced by the rank of their group in comparison to other groups, resulting in a large BFLPE (Marsh et al., 2014; Zell & Alicke, 2009). In the present work, we examined whether forecasters also neglect group rank information, resulting in a large BFLPE. Third, to examine a potential moderating factor, our final study tested whether extremity of group ranks alters the BFLPE in empathic forecasting (Study C3). Prior research found that the BFLPE has a stronger effect on experienced emotions when group ranks are extreme rather than moderate, that is, when people have very high rank in a very low rank group than moderately high rank in a moderately low rank group (Zell & Lesick, 2021). Here we tested whether empathic forecasts of the BFLPE are also larger when group ranks are extreme.

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the studies. Sample sizes for each study were determined prior to data collection. Each study was pre-registered. All materials, data, analysis code, pre-registrations, and supplemental analyses for each study are publicly available on OSF. Prior research suggests that empathic forecasting biases are pronounced when examining negative emotions (Boucher et al., 2015; Green et al., 2013). Therefore, we separately examined forecasts

of positive and negative emotions for targets that were big fish in little ponds and little fish in big ponds. We did not pre-register separate analyses of positive and negative affect but came to consider this strategy after the data were collected. This research was approved by the Institutional Review Board (IRB) at the authors' university. All participants provided informed consent prior to their participation.

Study C1

Our first study tested whether the BFLPE occurs in empathic forecasting and whether forecasts of the BFLPE differ from actual experiences of it. Participants were randomly assigned to either a forecaster or experiencer condition. Participants also received feedback indicating that the person they were evaluating had high rank in a low rank group or low rank in a high rank group. We anticipated that the BFLPE would occur in empathic forecasting and that this effect would be significantly larger than the BFLPE reported by people who actually experienced it.

Method

Participants and design. Participants were 259 undergraduates at a university in the Southeastern United States, who completed the study for credit in introductory psychology courses. According to our pre-registered plan, we collected data until we obtained at least 210 participants who passed attention and manipulation checks. This allowed us to achieve at least 95% power to detect a medium effect ($f = .25$), which was the smallest effect we were interested in detecting. After excluding 21 participants who failed attention or manipulation checks, our final sample contained 238 participants (181 women, 89 White, 79 Black, $M_{age} = 19.4$).

Participants were randomly assigned to one condition in a 2 (feedback: BFLP, LFBP) by 2 (perspective: experiencer, forecaster) between-subjects design. Whereas *experiencers* completed a verbal test, received feedback about their own performance, and provided affective

reactions to the test, forecasters reviewed the verbal test of a supposed previous participant at their school, received feedback about that student's performance, and forecasted that student's affective reactions. A visual depiction of the study procedures across experiencer and forecaster conditions is provided in Figure C1.

Procedures. Participants completed study measures online using Qualtrics. Participants were told that the purpose of the study was to measure the verbal reasoning skills of students at their university and that verbal reasoning ability is predictive of success in a variety of domains. In the *experiencer conditions*, participants next completed a 40-item verbal reasoning test. Prior research found this test to be sufficiently ambiguous to promote the believability of performance feedback (e.g., Zell & Lesick, 2021). In the *forecaster conditions*, participants reviewed a verbal reasoning test that was supposedly completed by a previous participant at their university named "Alex", whose response to each item was based on the modal response from a pilot study in the same population ($N = 128$, $M_{correct} = 14$).

After the test, participants received written feedback either about their own performance (experiencers) or Alex's performance (forecasters; see Table C1 for sample feedback). In *big-fish-little-pond (BFLP) conditions*, participants were told that the person they were evaluating ranked better than 68% of 206 students at their university and that their university ranked better than 32% of 39 universities. In *little-fish-big-pond (LFBP) conditions*, participants were told that the person they were evaluating ranked better than 32% of 206 students at their university and that their university ranked better than 68% of 39 universities. Summary information indicated that the person they were evaluating ranked above (below) average in a below (above) average school. To promote comprehension, participants recorded the feedback and summary information they received. The order of student and school feedback was counterbalanced. The

feedback levels (32%, 68%) and group sizes (206 students, 39 universities) were derived from previous research and were selected to promote the believability of manipulated performance feedback (Zell & Lesick, 2021).

Participants then completed the primary outcome measures of positive and negative affect. For equivalence, experiencers completed affect items immediately after reviewing their feedback and forecasters anticipated how Alex would complete affect items immediately after Alex reviewed their feedback. Specifically, participants completed a 3-item measure of positive affect (α = .92), indicating how *satisfied*, *proud*, and *happy* either themselves or Alex felt about their test performance (e.g., *How satisfied do you [did Alex] feel about your [their] Verbal performance?*), using 1 (*not at all*) to 7 (*extremely*) scales. Participants also completed a 3-item measure of negative affect (α = .88), indicating how *sad*, *discouraged*, and *distressed* either themselves or Alex felt about their performance using the same scales. To allow direct comparisons of positive and negative affect, negative affect items were reverse scored (i.e., higher values reflect more favorable affect on both indices). The measures of positive and negative affect were derived from previous research on reactions to social comparison information (Zell & Strickhouser, 2020).

Participants then completed a manipulation check, which asked them to recall the feedback they received about either their own or Alex's performance (*above average in a below average school*, *below average in an above average school*, *average in an average school*) and an attention check, which asked them to select the sixth circle in a 7-point scale. For exploratory purposes, participants completed a measure of target-evaluations, anticipated how well they [Alex] would perform on a second verbal test, completed a second verbal test, and completed a

social anxiety scale. Finally, participants completed demographic items (e.g., age, race, and gender) and an open-ended question asking if anything seemed suspicious about the study.

Participant exclusions. We excluded data from 19 participants for failing the manipulation check, 1 participant for failing the attention check, and 1 participant for expressing suspicion about the provided feedback ("I wasn't sure about the accuracy of the score.").

Results and Discussion

None of the hypothesized effects in this research were significantly moderated by age, race-ethnicity, or gender. Thus, we do not provide additional commentary on these factors.

Preliminary analyses. Indicative of successful random assignment, the number of items answered correctly on the verbal reasoning test was similar when comparing experiencers in the BFLP condition (M = 13.77, SD = 6.92) to experiencers in the LFBP condition (M = 14.79, SD = 7.27), $t(112) = 0.77$, $p = .446$, $d = 0.14$, 95% CI [–0.23, 0.51]. Further, the test viewed by forecasters evidenced a similar overall performance level (14 correct items) than the test completed by experiencers (M = 14.28, SD = 7.08), $t(113) = 0.42$, $p = .673$, $d = 0.04$, 95% CI [–0.14, 0.22]. Thus, actual verbal performance was largely equivalent across experiencer and forecaster conditions.

Primary analysis. Our primary analysis examined whether forecasted affect differed from experienced affect across feedback conditions. Along these lines, we conducted a 2 (affect type: positive, negative) by 2 (feedback: BFLP, LFBP) by 2 (perspective: experiencer, forecaster) ANOVA with repeated measures on the first factor (i.e., affect type). This analysis yielded a significant three-way interaction, $F(1, 234) = 14.39$, $p < .001$, $\eta_p^2 = .06$ (see Table C2 and Figure C2). We decomposed the three-way interaction with separate analyses on positive and negative affect.

For positive affect, a 2 (feedback) by 2 (perspective) ANOVA yielded a very large main effect of feedback, $F(1, 234) = 114.84$, $p < .001$, $\eta_p^2 = .33$. Consistent with the BFLPE, positive affect was much more pronounced in the BLFP conditions ($M = 4.90$, $SD = 1.23$) than the LFBP conditions ($M = 3.10$, $SD = 1.38$). There was also a very small but significant main effect of perspective, $F(1, 234) = 4.25$, $p = .040$, $\eta_p^2 = .02$, such that positive affect was slightly more pronounced among forecasters ($M = 4.20$, $SD = 1.42$) than experiencers ($M = 3.82$, $SD = 1.73$). Finally, the two-way interaction was non-significant, $F(1, 234) = 2.73$, $p = .100$, $\eta_p^2 = .01$.

For negative affect, a 2 (feedback) by 2 (perspective) ANOVA yielded a significant main effect of feedback, $F(1, 234) = 47.27$, $p < .001$, $\eta_p^2 = .17$. Supporting the BFLPE, negative affect was less pronounced among participants in the BFLP conditions ($M = 5.46$, $SD = 1.33$, reverse scored) than participants in the LFBP conditions ($M = 4.23$, $SD = 1.53$, reverse scored). Further, there was a significant main of perspective, $F(1, 234) = 13.79$, $p < .001$, $\eta_p^2 = .06$, whereby negative affect was somewhat less pronounced among experiencers ($M = 5.19$, $SD = 1.59$, reverse scored) than forecasters ($M = 4.56$, $SD = 1.47$, reverse scored). Lastly, these main effects were qualified by a significant two-way interaction, $F(1, 234) = 4.37$, $p = .038$, $\eta_p^2 = .02$. Planned comparisons on negative affect found that the BFLPE was much larger in the forecaster conditions, $t(234) = 6.47$, $p < .001$, $d = 1.16$, 95% CI [0.80, 1.53] than the experiencer conditions, $t(234) = 3.32$, $p = .001$, $d = 0.62$, 95% CI [0.25, 0.99]. Moreover, whereas forecasted negative affect was similar to experienced negative affect in the BFLP condition, $t(234) = 1.16$, $p = .247$, $d = 0.21$, 95% CI [–0.15, 0.57], forecasted negative affect was much more pronounced than experienced negative affect in the LFBP condition, $t(234) = 4.06$, $p < .001$, $d = 0.75$, 95% CI [0.38, 1.12].

In sum, Study C1 found a significant and large BFLPE both among experiencers and forecasters. That is, participants either felt better after having high rank in a low rank group than low rank in a high rank group (experiencers) or anticipated that another person would feel better after having high rank in a low rank group than low rank in a high rank group (forecasters). Nonetheless, forecasters significantly overestimated the size of the BFLPE on negative emotions. Specifically, the harmful impact of being a little fish in a big pond was less dramatic than forecasters predicted it would be.

Study C2

Having demonstrated that a robust BFLPE occurs in empathic forecasting, Study C2 next tested an underlying mechanism for this effect. Prior work on the experience of the BFLPE indicates that a neglect of group rank information leads people report more favorable emotions when they have high rank in a low rank group than low rank in high rank group (Zell & Alicke, 2009, 2010). Study C2 examined whether neglect of group rank also explains the BFLPE in empathic forecasting. Participants were told that another student had high or low rank at their school and that the student's school had high or low rank. We anticipated that student rank would have a stronger influence on empathic forecasting than school rank, resulting in a BFLPE.

Method

We recruited American adults on Prolific who met the following criteria: age 18–25, approval rate ≥ 80%, approved HITs ≥ 20, and identified as either male or female. Participants were paid $0.60 in exchange for their participation in the three-minute study ($12/hour). An a priori power analysis indicated that a sample of 210 participants would provide 95% power to detect a medium effect (f = .25), which was the smallest effect we were interested in detecting. In total, we collected data from 238 adults. We excluded data from 15 participants for failing a

manipulation check, resulting in a final sample of 223 participants (131 women, 137 White, M_{age} = 21.43). Participants were randomly assigned to one condition in a 2 (student rank: above average, below average) by 2 (school rank: above average, below average) between-subjects design. We restricted the study to participants aged 18-25, since this group is more likely to relate to the educational context studied in the present work.

Participants were told that they would evaluate a college student who completed a verbal reasoning test and that, although confidential, the student would be referred to as "Alex" and the student's school would be referred to as "Pembroke University". Next, participants received scores indicating that Alex ranked better than 68% or 32% of 206 students at their university and that their university ranked better than 72% or 28% of 39 schools. As in Study C1, participants received summary information, indicating that Alex ranked above or below average at an above or below average university, and the order of the scores was counterbalanced. Next, participants completed the same 3-item measures of forecasted positive affect (α = .96) and forecasted negative affect (α = .91) as Study C1; negative affect was reverse scored in preparation for analysis. Finally, participants completed exploratory measures (target-evaluations, forecasted self-evaluations), similar manipulation and attention checks as Study C1, as well as demographic items. We excluded data from 15 participants for failing the manipulation check; no participant failed the attention check.

Results and Discussion

We conducted a 2 (affect type) by 2 (school rank) by 2 (student rank) ANOVA with repeated measures on the first factor (i.e., affect type). Consistent with predictions, this analysis yielded a very large main effect of student rank, $F(1, 219) = 494.70$, $p < .001$, $\eta_p^2 = .69$, and a relatively small effect of school rank, $F(1, 219) = 23.47$, $p < .001$, $\eta_p^2 = .10$ (see Table C3; see

Figure C3). Participants forecasted much more favorable affect when Alex ranked above average (M = 5.69, SD = 0.79) than below average (M = 3.14, SD = 0.97). Further, participants forecasted somewhat more favorable affect when Alex's school ranked above average (M = 4.76, SD = 1.47) than below average (M = 4.08, SD = 1.57). Thus, empathic forecasts were based much more on student rank than school rank.

Additionally, although unexpected, there was a significant main effect of affect type, $F(1, 219) = 80.41$, $p < .001$, $\eta_p^2 = .27$, that was qualified by an affect type by student rank interaction, $F(1, 219) = 8.43$, $p = .004$, $\eta_p^2 = .04$. We decomposed the interaction by examining the effect of student rank on positive and negative affect. Student rank had a very strong influence on both affect types, but its influence was somewhat larger on positive affect than negative affect, $F(1, 219) = 495.58$, $p < .001$, $\eta_p^2 = .69$ and $F(1, 219) = 282.61$, $p < .001$, $\eta_p^2 = .56$, respectively. Thus, forecasters anticipated that student rank information would have a stronger influence on positive affect than negative affect. All other two-way interactions as well as the three-way interaction were non-significant, ps > .390.

Altogether, Study C2 supported our argument that group-rank neglect underlies the BFLPE in empathic forecasting. Specifically, empathic forecasts were influenced far more by another person's rank in their group than their group's rank. Thus, people anticipate that another person will have more favorable achievement emotions when they have high rank in a low rank group than low rank in a high rank group because they focus on that person's within-group rank but neglect group information. We also found that student rank had a slightly stronger impact on forecasted positive affect than forecasted negative affect, but the impact of student rank was highly pronounced in both instances.

Study C3

Prior work on experienced emotions suggests that the BFLPE is larger when group ranks are extreme rather than moderate (Zell & Lesick, 2021). To further explore potential parallels between the BFLPE in experienced and forecasted emotions, Study C3 tested whether this critical moderator also affects the magnitude of the BFLPE in empathic forecasting. Given that the BFLPE in empathic forecasting is likely driven by group rank neglect, we anticipated that forecasted emotions would be more favorable for big fish in little ponds, and less favorable for little fish in big ponds, when group ranks are extreme rather than moderate.

Method

Unless noted otherwise, methods in Study C3 were identical to Study C2. We recruited American adults on Prolific who met the following criteria: age 18–25, approval rate ≥ 95%, approved HITs ≥ 20, and identified as either male or female. In total, we collected data from 249 adults. After excluding 32 participants who failed a manipulation check, our final sample contained 217 participants (113 women, 134 White, M_{age} = 22.25). Participants were randomly assigned to one condition in a 2 (feedback type: BFLP, LFBP) by 2 (feedback extremity: moderate, extreme) between-subjects design. As in Study C2, participants evaluated a college student named "Alex" who attended "Pembroke University" and completed a verbal reasoning test.

In *moderate feedback conditions*, participants received scores indicating that Alex ranked better than 65% of 206 students at their school and that their school ranked better than 35% of 39 schools (BFLP) or that Alex ranked better than 35% of students at their school and that their school ranked better than 65% of schools (LFBP). Alternatively, in *extreme feedback conditions*, participants were told that Alex ranked better than 85% of students at their school and that their

school ranked better than 15% of schools (BFLP) or that Alex ranked better than 15% of students at their school and that their school ranked better than 85% of other schools (LFBP). Participants then completed the same measures of forecasted positive affect (α = .96) and forecasted negative affect (α = .91, reverse scored) as Studies C1-C2, as well as a similar manipulation check, attention check, exploratory measures, and demographic items as Study C2. We excluded data from 32 participants for failing the manipulation check; no participant failed the attention check.

Results and Discussion

We conducted a 2 (affect type) by 2 (feedback extremity) by 2 (feedback type) ANOVA with repeated measures on the first factor (i.e., affect type). This analysis yielded a non-significant main effect of feedback extremity, $F(1, 213) = 0.00$, $p = .976$, $\eta_p^2 < .01$, a very large main effect of feedback type, $F(1, 213) = 257.06$, $p < .001$, $\eta_p^2 = .55$, and as predicted, a significant feedback extremity by feedback type interaction, $F(1, 213) = 19.04$, $p < .001$, $\eta_p^2 = .08$ (see Table C4 and Figure C4). In the moderate feedback conditions, forecasted affect was much more favorable when Alex was a BFLP than an LFBP, $t(213) = 8.68$, $p < .001$, $d = 1.59$, 95% CI [1.20, 1.98]. However, this effect was even more pronounced in the extreme feedback conditions, $t(213) = 13.77$, $p < .001$, $d = 2.78$, 95% CI [2.30, 3.26]. Additionally, forecasted affect was *more favorable* in the extreme feedback condition than the moderate feedback condition when Alex was a BFLP, $t(213) = 3.13$, $p = .002$, $d = 0.60$, 95% CI [0.22, 0.98], but was *less favorable* in the extreme feedback condition than the moderate feedback condition when Alex was an LFBP, $t(213) = 3.04$, $p = .003$, $d = 0.59$, 95% CI [0.21, 0.98]. Thus, as anticipated, forecasted affect was especially favorable when Alex had very high rank in a very low rank group and especially unfavorable when Alex had very low rank in a very high rank group.

Additionally, there was a significant main effect of affect type, $F(1, 213) = 50.75$, $p < .001$, $\eta_p^2 = .19$, that was qualified by an affect type by feedback type interaction, $F(1, 213) = 5.75$, $p = .017$, $\eta_p^2 = .03$. Although feedback type had a very strong influence on both positive and negative affect, its influence was slightly larger on the former than the latter, $F(1, 213) = 255.14$, $p < .001$, $\eta_p^2 = .55$ and $F(1, 213) = 146.45$, $p < .001$, $\eta_p^2 = .41$, respectively. Therefore, as in Study C2, forecasters anticipated that rank information would have a slightly stronger influence on positive affect than negative affect. The remaining two-way interaction as well as the three-way interaction were non-significant, ps > .350.

Overall, Study C3 provided strong support for our hypothesis that the BFLPE in empathic forecasting is larger when group ranks are extreme rather than moderate. This finding suggests that extremity of group ranks moderates the BFLPE in forecasts of other people's emotions much as it does reports of one's own emotions (Zell & Lesick, 2021) and thus provides additional evidence suggesting that the BFLPE operates via a similar process in empathic forecasting as it does in actual experience (i.e., group rank neglect).

General Discussion

The BFLPE, namely the tendency for students to evaluate themselves more favorably and report more favorable emotions when they have high rank in a low rank school than low rank in a high rank school (Marsh et al., 2021; Pekrun et al., 2019), is one of the most robust and oft-studied phenomena in the social comparison literature. Research on the BFLPE has documented its emotional impact for people actually experiencing it. Here we present the first set of studies examining forecasts of how the BFLPE impacts other people's emotions. Along these lines, participants in Study C1 received feedback indicating that they or another person had high rank in a low rank school or low rank in a high rank school, and then either reported their own affect

or anticipated the other person's affect. Results extended the BFLPE to empathic forecasting and suggest that people overestimate the influence of the BFLPE on negative emotions experienced by others.

We next tested an underlying mechanism and critical moderator of the BFLPE in empathic forecasting. Specifically, in Study C2 we examined whether neglect of group rank underlies the BFLPE in empathic forecasting, as it does in actual emotional experience (Marsh et al., 2014; Zell & Alicke, 2009). Consistent with predictions, results suggested the forecasts of other people's emotions are influenced to a much greater extent by that person's rank in their group than group rank information, resulting in a BFLPE. Additionally, in Study C3 we examined whether extremity of group ranks exacerbates the BFLPE in empathic forecasting, as it does in actual emotional experience (Zell & Lesick, 2021). As expected, results demonstrated that empathic forecasts were especially favorable when a target had very high rank in a very low rank group, and especially unfavorable when a target had very low rank in a very high rank group. Taken together, these data strongly support the BFLPE in empathic forecasting, but also illustrate ways in which it is both similar to and different from actual experience of the effect.

Implications

The present research advances the BFLPE literature as well as work on empathic forecasting more generally. First, our studies charted new territory by examining whether the BFLPE occurs in empathic forecasting. All three studies found that the BFLPE occurs in forecasts of other people's emotions, and the size of this effect was consistently very large. Therefore, people not only feel more positive achievement emotions and less negative achievement emotions when they are a big fish in a little pond versus a little fish in a big pond (Holm et al., 2020; Pekrun et al., 2019), but they also assume that other people will feel more

positive emotions and fewer negative emotions in the case of the former versus the later. Given that research on empathic forecasting is relatively new and so far limited to forecasts of other people's emotions in only a few contexts, such as reactions to social blunders (Hall et al., 2020) or insults (Moons et al., 2017), the present work extends this literature by examining empathic forecasts in the context of a well-established social comparison effect.

Second, our studies contribute to a growing literature suggesting that empathic forecasts are often more extreme than experienced emotions. Along these lines, prior work suggests that people overestimate other people's emotional reactions to succeeding or failing a test (Pollmann & Finkenauer, 2009). Moreover, research focusing on negative affect found that people overestimate the degree to which women feel anxiety following stereotype threat (Boucher et al., 2015) and the degree to which relationship partners feel sadness following relationship transgressions (Green et al., 2013). Study C1 builds upon these findings by demonstrating that people overestimate the degree to which others feel negative affect when they have low rank in a high rank group. To be fair, forecasters correctly predicted that having low rank in high rank group would be an unpleasant experience. Nonetheless, the amount of negative affect reported by experiencers in these conditions was substantially less than the amount predicted by forecasters.

Third, despite the tendency to overestimate the impact of the BFLPE identified in Study C1, our subsequent studies found that the BFLPE in empathic forecasting likely operates via a similar underlying mechanism, and is influenced in a similar way by a critical moderating factor, than the BFLPE in actual emotional experience. Along these lines, Study C2 found that forecasts of another person's affect were very strongly influenced by that person's rank in their group, but only relatively weakly influenced by the rank of that person's group in relation to other groups.

These results mirror findings on emotional reactions to the BFLPE (Marsh et al., 2014; Zell & Alicke, 2009) and thus suggest that group rank neglect is a mechanism underlying both the experience of the BFLPE and forecasts of it. Further, Study C3 found that the forecasted impact of being a big fish in a little pond was *more favorable*, and the forecasted impact of being a little fish in a big pond was *less favorable*, when group ranks were extreme. Once again, these results are parallel to findings on the actual experience of the BFLPE (Zell & Lesick, 2021) and thus suggest that extremity of ranks moderates both the experience of the BFLPE and forecasts of it. In sum, the present findings buttress prior work on mechanisms and moderators of the BFLPE, but also extend these effects to a novel context.

Fourth, although unexpected, we found a relatively small but consistent difference in the impact of social comparison feedback on forecasts of positive versus negative affect. More specifically, Studies C2 and C3 found that the forecasted impact of being a big fish in a little pond versus a little fish in a big pond was somewhat larger for positive affect than negative affect. Thus, forecasters appear to assume that positive achievement emotions are slightly more responsive to social comparison than negative achievement emotions. However, this difference by affect type was relatively small and was dwarfed by the effect of social comparison. In each study, forecasters predicted that *both* positive and negative achievement emotions would be profoundly influenced by social comparison. Research should continue to explore the consequences of social comparison for both positive and negative achievement emotions.

Beyond these theoretical contributions, the present findings may have practical implications for empathic forecasting in daily life. Research indicates that empathic accuracy facilitates close relationships (Le et al., 2020; Sened et al., 2017). Further, the BFLPE is a prominent manifestation of social comparison and may occur regularly in daily life (Frank, 1985;

Gladwell, 2013). Our data suggests that people may overestimate the impact of the BFLPE on negative emotions. Thus, when interacting with someone who has low rank in a high rank group, people may anticipate that this person is feeling worse than they really are, which could lead to maladaptive communications and behaviors. Additionally, biased empathic forecasts may lead people to give poor advice to others who are considering entry into environments where they would have low rank in a high rank group (e.g., recommending that they decline the opportunity because of anticipated negative affect).

Limitations and Future Directions

Despite these important contributions, the present research has limitations that could be addressed in future inquiry. Specifically, although a major strength of Study C1 was its inclusion of both experiencers and forecasters, Studies C2 and C3 were restricted to forecasters. This allowed us to examine a critical mechanism and moderator of the BFLPE in empathic forecasting in efficient experimental designs. Further, when comparing our work to prior studies on experienced affect (Zell & Alicke, 2009; Zell & Lesick, 2021), we found that group rank neglect and extremity of ranks had a very similar influence on affective forecasts. Nonetheless, although beyond the scope of the present work, future studies that include both experiencers and forecasters could be conducted to examine whether the impact of group rank neglect and extremity of ranks is parallel across experiencers and forecasters when tested simultaneously in the same experimental conditions.

Moreover, research suggests that the BFLPE is potentially moderated by other factors besides the extremity of ranks, such as student motivation level and culture (see Fang et al., 2018; Marsh et al., 2021). Thus, future work should examine whether forecasters appreciate that the BFLPE has a greater affective impact for some students and under certain conditions than

others. Additionally, the present work did not explore whether there are individual differences in empathic forecasts of the BFLPE or whether some forecasters evidence greater accuracy than others. Prior work suggests that aspects of the perceiver, such as their dispositions, influence empathic forecasting (Hall et al., 2018, 2020). Thus, an important question for future research is the degree to which perceiver effects emerge in empathic forecasting of the BFLPE. Finally, future work could identify downstream implications of the BFLPE in empathic forecasts, for example, whether it influences decisions to approach or avoid another person (Moons et al., 2017) as well as the advice people give to others making academic or vocational choices (Wu et al., 2018).

In terms of methods, a limitation of Study C3 was that the extremity manipulation involved differences in both intragroup (student rank) and intergroup (school rank) comparison information. Future study is needed to clarify whether the stronger BFLPE obtained in extreme rather than moderate conditions is attributable to extreme intragroup comparisons, intergroup comparisons, or both. We also acknowledge that the three studies described in this paper involve samples of students or young adults. Future research should assess empathic forecasts and actual experiences of the BFLPE in older samples. Relatedly, although most work on the BFLPE involves ranks in school settings, research is needed to explore whether a BFLPE occurs in other settings, such as when people receive feedback indicating their rank in their country (Stockus & Zell, 2023b).

A potential limitation of the performance feedback used in this work was that it exclusively used a "better than" framing to reflect relative performance levels (e.g., participants were told that they ranked *better* than 68% of students at their school). Future work could explore whether a "worse than" framing alters the BFLPE (e.g., participants could be told that

they ranked *worse* than 32% of students at their school). Moreover, a limitation of this research was that participants forecasted the affect of another person named "Alex" whom they never met and who was in fact not a real person. Although we think participants believed that Alex was a real person, and the feedback ostensibly given to Alex mirrored the feedback given to actual experiencers in Study C1, future research is needed to explore whether empathic forecasts of the BFLPE differ for real targets versus artificial targets.

Conclusions

The BFLPE has a powerful impact on achievement emotions, increasing positive emotions such as enjoyment and pride, and decreasing negative emotions such as anger and anxiety (Holm et al., 2020; Pekrun et al., 2019). In the present research, we examined empathic forecasts of the BFLPE, that is, whether people appreciate how being a big fish in a little pond versus a little fish in a big pond impacts other people's affective state. Our results indicated that a robust BFLPE occurs in empathic forecasting, that forecasters overestimate this effect on negative emotions, that group rank neglect underlies the effect, and that the effect is magnified when group ranks are extreme. More broadly, future research is needed to explore the degree to which the BFLPE in empathic forecasting varies across different perceivers and targets. The BFLPE has long fascinated scientists and laypersons (Frank, 1985; Gladwell, 2013). Thus, future work should continue to examine not just the experience of the effect, but also forecasts of how it impacts others and the implications of these forecasts for subsequent behavior.

CHAPTER V: INTEGRATED DISCUSSION

In this integrated dissertation, I provide three papers examining the BFLPE, which refers to the tendency for students to evaluate themselves more favorably when they have high rank in low rank schools than low rank in high rank schools (Marsh et al., 2015; Marsh & Parker, 1984; Zell & Lesick, 2021). Specifically, I describe work examining experienced BFLPEs in different regional contexts, whether the deflating experience of being an LFBP is reduced by growth mindsets, and whether the BFLPE occurs in empathic forecasts of other people's self-evaluations and mood. Whereas papers one and two examine experiences of the BFLPE (Stockus & Zell, 2023a, 2023b), paper three examines forecasts of the effect (Stockus & Zell, 2024a). Papers one and two provide evidence for the occurrence of experienced BFLPEs across novel reference groups (Stockus & Zell, 2023b) and following activation of a growth mindset (Stockus & Zell, 2023a). Paper three suggests that the BFLPE obtains in empathic forecasting, and that the mechanism under which forecasted BFLPEs occur is similar to that which occurs in actual experience (Stockus & Zell, 2024a).

Three key findings emerge from this program of research. First, in paper one, we found that a strong experienced BFLPE emerged at the school, state, and country level and that the size of the effect did not significantly vary according to the locality of the reference group (Stockus & Zell, 2023b). That is, the BFLPE was not larger at the state level than at the country level (Study 2) or at the school level than at the country level (Study 3). Second, in paper two, we found that the BFLPE is robust to moderation by growth mindsets (Stockus & Zell, 2023a). A robust BFLPE emerged across studies, and we found little support for the possibility that the detrimental effects of being an LFBP are reduced following activation of a growth mindset. Third, in paper three, we observed that the BFLPE occurs in empathic forecasting (Stockus &

Zell, 2024a). Along these lines, we found that forecasters anticipate that the BFLPE has a larger impact on people's negative emotions than experiencers actually report (Study 1), that group rank neglect underlies forecasted BFLPEs (Study 2), and that forecasted BFLPEs are more pronounced under extreme social comparison conditions (Study 3).

Implications

When viewed collectively, this research has several implications for work on the BFLPE. First, this set of papers examining experiences and forecasts of the BFLPE contributes to our understanding of the conditions and contexts in which the BFLPE is likely to emerge. For example, we found that the BFLPE obtains in different regional contexts in a similar way as it does in educational contexts (Stockus & Zell, 2023b), suggesting that the effect is not confined to school settings. Furthermore, we provide evidence suggesting that the BFLPE occurs in empathic forecasting as well as actual experience (Stockus & Zell, 2024a). This suggests that the BFLPE is not confined to one's own experience. These papers bolster the generalizability of the BFLPE by demonstrating that the effect extends to novel regional contexts (Stockus & Zell, 2023b) and extends from one's own experience to empathic forecasting of other people's experiences (Stockus & Zell, 2024a).

Along these lines, empirical paper one provides among the first evidence suggesting that the BFLPE occurs in novel reference groups (Stockus & Zell, 2023b). Most papers on the BFLPE are conducted in academic contexts, where the reference group is one's school (Marsh, 1987; Marsh et al., 2015). Only one prior paper provides non-experimental evidence indicating that the BFLPE occurs at both the school and country level (Marsh et al., 2019). In paper one, we provide experimental evidence indicating that the BFLPE occurs in the context of one's school, state, and country, which suggests that the BFLPE is not limited to academic reference groups.

Future investigations of the BFLPE should determine reference groups in which the effect is most likely to emerge and explore other potentially viable reference groups, such as one's neighborhood, workplace, or friend group.

Relatedly, in empirical paper three, we found that the BFLPE is an effect that people experience themselves but also one that people anticipate for others (Stockus & Zell, 2024a). By demonstrating that the BFLPE occurs in novel (forecasting) contexts, our work suggests that the BFLPE is not confined only to actual experiences of the effect. Thus, future tests of the BFLPE should not limit examination of the effect to only experiences but should also consider whether the BFLPE emerges from the perspective of empathic forecasters. More work is needed to identify the consequences of forecasted BFLPEs for person perception and behavior (Moons et al., 2017). For example, forecasted BFLPEs may influence whether one chooses to approach or avoid another person.

Second, this collection of papers provides evidence in support of the robustness of the BFLPE. Empirical paper one demonstrated that the BFLPE is robust across regional and academic contexts (Stockus & Zell, 2023b). That is, the BFLPE occurred at the school, state, and country level and was comparable in size across these reference groups. Empirical paper two showed that the BFLPE is not substantially reduced by growth mindsets (Stockus & Zell, 2023a). That is, the negative effects of being an LFBP emerged regardless of whether participants adopted a growth or fixed mindset. This suggests that the BFLPE has highly robust effects on self-evaluations and mood that are not easily dislodged by growth mindset interventions. Altogether, this program of research supports prior work suggesting that the BFLPE is robust across regional and academic contexts (Marsh et al., 2019) and that it is robust to growth mindset interventions (Marsh et al., 2021). Further, one novel addition to the BFLPE literature is our

demonstration of the robustness of the BFLPE across both experiences and forecasts of the effect (Stockus & Zell, 2024a).

Third, the experimental nature of this collection of work has important implications for establishing the causal effect of the BFLPE on key outcomes, namely self-evaluations and performance-related affect (e.g., Zell & Lesick, 2021). Most BFLPE studies are non-experimental in that they only measure associations of student and school ranks with academic self-concepts, preventing causal conclusions (e.g., Marsh et al., 2015). In the papers described above, we manipulated student and group ranks, allowing causal conclusions about the impact of having either high rank in a low rank group or low rank in a high rank group on key outcomes (Stockus & Zell, 2023b). Therefore, whereas most previous work is unable to draw causal conclusions about impacts of the BFLPE on outcomes such as academic self-concepts and emotion (Marsh et al., 2015, 2019; Pekrun et al., 2019), our experiments allow for such conclusions.

A fourth and final implication of this work stems from our finding that the effect of student rank information displaced that of school rank information in empathic forecasts of the BFLPE (intergroup comparison neglect; Stockus & Zell, 2024a). Prior work on experiences of the BFLPE found that people tend to focus on their rank within a group and neglect the rank of their group relative to other groups during self-evaluation (Zell & Alicke, 2009, 2010, 2020). That is, the within-group comparison, where the student compares their performance to another student in their class or school is more salient than the between-group comparison, where the student compares the rank of their school relative to other schools. The present work suggests that intergroup comparison neglect also underlies forecasts of the BFLPE (Stockus & Zell,

2024a). Additional study is needed to further establish group rank neglect as a mechanism of the BFLPE in empathic forecasting.

Limitations

There are a few limitations associated with this collection of empirical papers. First, each of the studies in the three papers were experiments. This is a notable strength in that the studies afforded us greater control over the precise social comparison conditions experienced by participants and the ability to identify causal effects. That is, being a BFLP had positive effects on self-evaluations and mood, whereas being an LFBP had negative effects. The control afforded by these experiments does however present a compromise in terms of external validity. Our experimental designs, although adequately powered, used relatively small convenience samples of mostly young, undergraduate college students or Internet respondents in the United States. Applied BFLPE studies have more external validity than the present work given that they are conducted in real-world settings with larger samples (Marsh, 1987; Marsh et al., 2015, 2019). This speaks to the need for experimental investigations of the BFLPE using varied samples of participants and not exclusively young, educated, online samples of participants from Western populations.

A second limitation of this work is that each of our papers used verbal reasoning performance feedback to manipulate the BFLPE. Although people likely have a general idea of how their verbal reasoning ability compares to other people in their group, and how their group ranks relative to other groups, they may not have precise knowledge about these ranks. Thus, the present studies may involve somewhat artificial situations that are uncommon in everyday life. The creation of these artificial situations was necessary to precisely manipulate social comparison conditions in each of our experiments (BFLP, LFBP). In applied BFLPE studies, the

association between student and school rank with academic self-concepts is measured without having to create such artificial situations. In our studies, these situations were an essential component of our manipulation that subsequently allowed us to make causal inferences about experienced and forecasted BFLPEs.

A third potential limitation of this set of papers is that each used an exclusively "better than" framing to reflect relative performance levels. Specifically, participants were told that they or the person they were evaluating ranked "better than" 68% or 32% of 206 previous participants. The feedback percentages and group sizes used in our studies were derived from previous research and were selected to promote the believability of manipulated performance feedback (Zell & Lesick, 2021). It is unclear whether manipulating the phrasing of the verbal reasoning performance feedback might alter the BFLPE. It seems possible that a "worse than" framing may elicit a stronger forecasted BFLPE for an LFBP target (e.g., "Alex ranked *worse* than 68% of 206 participants at their school") whereas a "better than" framing may elicit a stronger BFLPE for a BFLP target (e.g., "Alex ranked *better* than 68% of 206 participants at their school"). Future research is needed to explore this possibility in experiencing and forecasting contexts.

Future Directions

Several viable avenues for future study emerge from this program of research. First, there is a need for experimental tests of the BFLPE in non-academic settings and from the perspective of empathic forecasters. Empirical paper one tested the BFLPE in different regional contexts, specifically at the school, state, and country level (Stockus & Zell, 2023b). This is the first paper to our knowledge testing the occurrence of a regional BFLPE using experimental designs. Only one prior paper tested the occurrence of the BFLPE at the school and country level non-

experimentally (Marsh et al., 2019). Such a lack of scholarship on the regional BFLPE points to the need for additional work. Further, in our study of the regional BFLPE, we did not find any evidence in support of the notion that the BFLPE varies according to the locality of the reference group. Future investigations of the regional BFLPE should test the effect in more local reference groups (e.g., one's neighborhood, workplace, or peer group). Relatedly, empirical paper three tested whether the BFLPE emerges in both experiencing and forecasting perspectives (Stockus & Zell, 2024a). To our knowledge, this is the only paper examining how the BFLPE varies according to the perspective one adopts. Additional experimental work examining experienced and forecasted BFLPEs may be helpful in elucidating any differences observed between these two perspectives (e.g., negative emotions).

Second, additional work is needed to test the robustness of the BFLPE. More specifically, whether the BFLPE is robust to moderation by growth mindsets (Stockus & Zell, 2023a) and whether people consistently underestimate the impact of the BFLPE on other people's emotions (Stockus & Zell, 2024a). In empirical paper two, we tested whether negative effects resulting from the experience of being an LFBP are reduced following activation of a growth mindset (Stockus & Zell, 2023a). Brief growth mindset interventions have been shown to be effective at reducing harmful impacts of failure feedback on emotions and motivation (Plaks & Stecher, 2007), however, our paper is the first to test whether growth mindsets reduce the BFLPE specifically. Additional experiments with different manipulations of growth mindsets are needed to determine whether the BFLPE is moderated by growth mindsets or if the effect is robust to growth mindsets as has been argued in previous research (Marsh et al., 2021). Empirical paper three found that perceivers were generally aware of the robust impact of the BFLPE on other people's positive emotions (Stockus & Zell, 2024a). However, forecasters overestimated the

impact of the BFLPE on others' negative emotions. Future research should examine the degree to which empathic forecasters underestimate the impact of the BFLPE on other people's positive versus negative emotions.

Third, and finally, future research should explore untested moderators of experienced and forecasted BFLPEs. Self-enhancement, where people inflate their own positive characteristics, may act as a moderator reducing the BFLPE. People high in narcissism have been found to experience weaker BFLPEs than those low in narcissism (Jonkmann et al., 2012), so the possibility exists that other forms of self-enhancement might moderate experienced BFLPEs. Social comparison orientation may also moderate experienced BFLPEs, such that people who score high on this individual difference dimension engage in comparisons with others more often and thus may be more affected by the BFLPE. There is no direct support for social comparison orientation as a moderator of experienced BFLPEs, but it does seem plausible given its moderation of other social comparison effects (Gibbons & Buunk, 1999). In empirical paper three, we tested whether the BFLPE occurs in empathic forecasting, however, we did not test specifically whether there are individual differences in empathic forecasts of the BFLPE (Stockus & Zell, 2024a). Past research suggests that certain characteristics of the perceiver (e.g., social anxiety) may influence empathic forecasting in general (Hall et al., 2018, 2020). More work is needed to examine whether there are individual differences that may either enhance or attenuate empathic forecasts of the BFLPE specifically.

Conclusions

Across three papers in this integrated dissertation, I provide evidence in support of the emergence of experienced and forecasted BFLPEs. Empirical papers one and two examine experiences of the effect, that is, consequences of the BFLPE on people's own self-evaluations

and affect (Stockus & Zell, 2023a, 2023b), whereas paper three examines forecasts of the effect, or perceptions of how the BFLPE impacts other people's self-evaluations and affect (Stockus & Zell, 2024a). We provide evidence suggesting that experienced BFLPEs occur in different regional contexts (Stockus & Zell, 2023b), that these BFLPEs are robust to moderation by growth mindsets (Stockus & Zell, 2023a), and that the BFLPE occurs in empathic forecasting (Stockus & Zell, 2024a). These papers contribute to the BFLPE literature by demonstrating that the BFLPE extends beyond academic contexts, occurs in experiences as well as forecasts of other people's experiences, and is likely not altered substantially by growth mindset interventions. Altogether, the findings that emerge from this program of research expand our understanding of how the BFLPE operates in different contexts and from different perspectives.

REFERENCES

Alicke, M. D., Zell, E., & Bloom, D. L. (2010). Mere categorization and the frog-pond effect. *Psychological Science*, *21*(2), 174–177. https://doi.org/10.1177/0956797609357718

Alicke, M. D., Zell, E., & Guenther, C. L. (2013). Social self-analysis: Constructing, protecting, and enhancing the self. In J. M. Olson & M. P. Zanna (Eds.), *Advances in experimental social psychology*. (Vol. 48, pp. 173–234). Elsevier Academic Press.

Anderson, C., Kraus, M. W., Galinsky, A. D., & Keltner, D. (2012). The local-ladder effect: Social status and subjective well-being. *Psychological Science*, *23*(7), 764–771. https://doi.org/10.1177/0956797611434537

Atzil-Slonim, D., Bar-Kalifa, E., Fisher, H., Lazarus, G., Hasson-Ohayon, I., Lutz, W., Rubel, J., & Rafaeli, E. (2019). Therapists' empathic accuracy toward their clients' emotions. *Journal of Consulting and Clinical Psychology*, *87*(1), 33–45. https://doi.org/10.1037/ccp0000354

Basarkod, G., Marsh, H. W., Guo, J., Parker, P. D., Dicke, T., & Pekrun, R. (2023). The happy-fish-little-pond effect on enjoyment: Generalizability across multiple domains and countries. *Learning and Instruction*, *85*. https://doi.org/10.1016/

Bocage-Barthélémy, Y., Selimbegović, L., & Chatard, A. (2018). Evidence that social comparison with the thin ideal affects implicit self-evaluation. *International Review of Social Psychology*, *31*(1), Article 2. https://doi.org/10.5334/irsp.114

Boucher, K. L., Rydell, R. J., & Murphy, M. C. (2015). Forecasting the experience of stereotype threat for others. *Journal of Experimental Social Psychology*, *58*, 56–62. https://doi.org/10.1016/j.jesp.2015.01.002

Breidenthal, A. P., Liu, D., Bai, Y., & Mao, Y. (2020). The dark side of creativity: Coworker

envy and ostracism as a response to employee creativity. *Organizational Behavior and Human Decision Processes*, *161*, 242–254. https://doi.org/10.1016/j.obhdp.2020.08.001

Brez, C., Hampton, E. M., Behrendt, L., Brown, L., & Powers, J. (2020). Failure to replicate: Testing a growth mindset intervention for college student success. *Basic and Applied Social Psychology*, *42*(6), 460–468. https://doi.org/10.1080/01973533.2020.1806845

Burnette, J. L., Billingsley, J., Banks, G. C., Knouse, L. E., Hoyt, C. L., Pollack, J. M., & Simon, S. (2023a). A systematic review and meta-analysis of growth mindset interventions: For whom, how, and why might such interventions work? *Psychological Bulletin*, *149*(3–4), 174–205. https://doi.org/10.1037/bul0000368

Burnette, J. L., Knouse, L. E., Billingsley, J., Earl, S., Pollack, J. M., & Hoyt, C. L. (2023b). A systematic review of growth mindset intervention implementation strategies. *Social and Personality Psychology Compass*, *17*(2), e12723. https://doi.org/10.1111/spc3.12723

Chatard, A., Bocage-Barthélémy, Y., Selimbegović, L., & Guimond, S. (2017). The woman who wasn't there: Converging evidence that subliminal social comparison affects self-evaluation. *Journal of Experimental Social Psychology*, *73*, 1–13. https://doi.org/10.1016/j.jesp.2017.05.005

Claro, S., Paunesku, D., & Dweck, C. S. (2016). Growth mindset tempers the effects of poverty on academic achievement. *PNAS*, *113*(31), 8664–8668. https://doi.org/10.1073/pnas.1608207113

Cundiff, J. M., & Matthews, K. A. (2017). Is subjective social status a unique correlate of physical health? A meta-analysis. *Health Psychology*, *36*(12), 1109–1125. https://doi.org/10.1037/hea0000534

David, O., & Bar-Tal, D. (2009). A sociopsychological conception of collective identity: The

case of national identity as an example. *Personality and Social Psychology Review*, *13*(4), 354–379. https://doi.org/10.1177/1088868309344412

Deska, J. C., Kunstman, J. W., Bernstein, M. J., Ogungbadero, T., & Hugenberg, K. (2020). Black racial phenotypicality shapes social pain and support judgments. *Journal of Experimental Social Psychology*, *90*, 103998. https://doi.org/10.1016/j.jesp.2020.103998

Dweck C. S. (2008). *Mindset: The new psychology of success*. Ballantine Books.

Fang, J., Huang, X., Zhang, M., Huang, F., Li, Z., & Yuan, Q. (2018). The big-fish-little-pond effect on academic self-concept: A meta-analysis. *Frontiers in Psychology*, *9*, 1569–1569. https://doi.org/10.3389/fpsyg.2018.01569

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*(2), 117–140. https://doi.org/10.1177/001872675400700202

Frank, R. H. (1985). *Choosing the right pond: Human behavior and the quest for status*. Oxford University Press.

Garcia, S. M., Tor, A., & Schiff, T. M. (2013). The psychology of competition: A social comparison perspective. *Perspectives on Psychological Science*, *8*(6), 634–650. https://doi.org/10.1177/1745691613504114

Gerber, J. P., Wheeler, L., & Suls, J. (2018). A social comparison theory meta-analysis 60+ years on. *Psychological Bulletin*, *144*(2), 177–197. https://doi.org/10.1037/bul0000127

Gibbons, F. X., & Buunk, B. P. (1999). Individual differences in social comparison: development of a scale of social comparison orientation. *Journal of Personality and Social Psychology*, *76*(1), 129–142. https://doi.org/10.1037/0022-3514.76.1.129

Gilbert, D. T., Giesler, R. B., & Morris, K. A. (1995). When comparisons arise. *Journal of Personality and Social Psychology*, *69*(2), 227–236. https://doi.org/10.1037/0022-

3514.69.2.227

Gilbert, D. T., Pinel, E. C., Wilson, T. D., Blumberg, S. J., & Wheatley, T. P. (1998). Immune neglect: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology*, *75*(3), 617–638. https://doi.org/10.1037/0022-3514.75.3.617

Gladwell, M. (2013). *David and Goliath: Underdogs, misfits, and the art of battling giants*. Hachette Book Group.

Gonzalez, F. J., Jung, M. H., & Critcher, C. R. (2021). Responsibility amplifies empathic forecasts. *Journal of Experimental Psychology: General*, *150*(5), 956–971. https://doi.org/10.1037/xge0000976

Green, J. D., Davis, J. L., Luchies, L. B., Coy, A. E., Van Tongeren, D. R., Reid, C. A., & Finkel, E. J. (2013). Victims versus perpetrators: Affective and empathic forecasting regarding transgressions in romantic relationships. *Journal of Experimental Social Psychology*, *49*(3), 329–333. https://doi.org/10.1016/j.jesp.2012.12.004

Hagerty, M. R. (2000). Social comparisons of income in one's community: Evidence from national surveys of income and happiness. *Journal of Personality and Social Psychology*, *78*(4), 764–771. https://doi.org/10.1037/0022-3514.78.4.764

Hall, K. A. A., Coleman, K., & Timpano, K. R. (2020). Associations between social anxiety and affective and empathic forecasts: A replication and extension in a Mechanical Turk sample. *Behavior Therapy*, *51*(3), 365−374. https://doi.org/10.1016/j.beth.2019.06.004

Hall, K. A. A., Joormann, J., Siemer, M., & Timpano, K. R. (2018). The impact bias in self and others: Affective and empathic forecasting in individuals with social anxiety. *Behaviour Research and Therapy*, *106*, 37−46. https://doi.org/10.1016/j.brat.2018.05.001

Hodges, S. D., & Kezer, M. (2021). It is hard to read minds without words: Cues to use to achieve empathic accuracy. *Journal of Intelligence*, *9*(2), 27. https://doi.org/10.3390/jintelligence9020027

Holm, M. E., Björn, P. M., Laine, A., Korhonen, J., & Hannula, M. S. (2020). Achievement emotions among adolescents receiving special education support in mathematics. *Learning and Individual Differences*, *79*, 101851. https://doi.org/10.1016/j.lindif.2020.101851

Hong, Y.-y., Chiu, C.-y., Dweck, C. S., Lin, D. M.-S., & Wan, W. (1999). Implicit theories, attributions, and coping: A meaning system approach. *Journal of Personality and Social Psychology*, *77*(3), 588–599. https://doi.org/10.1037/0022-3514.77.3.588

Huguet, P., Dumas, F., Marsh, H., Régner, I., Wheeler, L., Suls, J., Seaton, M., & Nezlek, J. (2009). Clarifying the role of social comparison in the big-fish-little-pond effect (BFLPE): An integrative study. *Journal of Personality and Social Psychology*, *97*(1), 156–170. https://doi.org/10.1037/a0015558

Igou, E. R. (2008). "How long will I suffer?" versus "how long will you suffer?" A self-other effect in affective forecasting. *Journal of Personality and Social Psychology*, *95*(4), 899–899. https://doi.org/10.1037/a0011619

Jonkmann, K., Becker, M., Marsh, H. W., Lüdtke, O., & Trautwein, U. (2012). Personality traits moderate the big-fish-little-pond effect of academic self-concept. *Learning and Individual Differences*, *22*(6), 736–746. https://doi.org/10.1016/j.lindif.2012.07.020

Koivuhovi, S., Marsh, H. W., Dicke, T., Sahdra, B., Guo, J., Parker, P. D., & Vainikainen, M.-P. (2022). Academic self-concept formation and peer-group contagion: Development of the big-fish-little-pond effect in primary-school classrooms and peer groups. *Journal of*

Educational Psychology, *114*(1), 198–213. https://doi.org/10.1037/edu0000554

Kraus, M. W., Tan, J. J. X., & Tannenbaum, M. B. (2013). The social ladder: A rank-based perspective on social class. *Psychological Inquiry*, *24*(2), 81–96. https://doi.org/10.1080/1047840X.2013.778803

Krizan, Z., (2018). Social comparison. *Oxford research encyclopedia of psychology: Social processes*. https://doi.org/10.1093/acrefore/9780190236557.013.251

Križan, Z., & Gibbons, F. (Eds.). (2014). *Communal functions of social comparison*. Cambridge: Cambridge University Press.

Lau, T., & Morewedge, C. K. (2016). Overcorrection for social-categorization information moderates impact bias in affective forecasting. *Psychological Science*, *27*(10), 1340–1351. https://doi.org/10.1177/0956797616660292

Le, B. M., Côté, S., Stellar, J., & Impett, E. A. (2020). The distinct effects of empathic accuracy for a romantic partner's appeasement and dominance emotions. *Psychological Science*, *31*(6), 607–622. https://doi.org/10.1177/0956797620904975

Lench, H. C., Levine, L. J., Perez, K., Carpenter, Z. K., Carlson, S. J., Bench, S. W., & Wan, Y. (2019). When and why people misestimate future feelings: Identifying strengths and weaknesses in affective forecasting. *Journal of Personality and Social Psychology*, *116*(5), 724–724. https://doi.org/10.1037/pspa0000143

Lesick, T. L., & Zell, E. (2023). Race-ethnicity and the big-fish-little-pond effect in the United States. *Social Psychological Bulletin*, *18*, 1-15. https://doi.org/10.32872/spb.10261

Luttmer, E. F. P. (2005). Neighbors as negatives: Relative earnings and well-being. *The Quarterly Journal of Economics*, *120*(3), 963–1002. https://doi.org/10.2139/ssrn.571824

Macnamara, B. N., & Burgoyne, A. P. (2023). Do growth mindset interventions impact students' academic achievement? A systematic review and meta-analysis with recommendations for best practices. *Psychological Bulletin*, *149*(3–4), 133–173. https://doi.org/10.1037/bul0000352

Markman, K. D., & McMullen, M. N. (2003). A reflection and evaluation model of comparative thinking. *Personality and Social Psychology Review*, *7*(3), 244–267. https://doi.org/10.1207/S15327957PSPR0703_04

Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of Educational Psychology*, *79*(3), 280–295. https://doi.org/10.1037/0022-0663.79.3.280

Marsh, H. W., Abduljabbar, A. S., Morin, A. J. S., Parker, P., Abdelfattah, F., Nagengast, B., & Abu-Hilal, M. M. (2015). The big-fish-little-pond effect: Generalizability of social comparison processes over two age cohorts from Western, Asian, and Middle Eastern Islamic countries. *Journal of Educational Psychology*, *107*(1), 258–271. https://doi.org/10.1037/a0037485

Marsh, H. W., Kuyper, H., Morin, A. J. S., Parker, P. D., & Seaton, M. (2014). Big-fish-little-pond social comparison and local dominance effects: Integrating new statistical models, methodology, design, theory and substantive implications. *Learning and Instruction*, *33*, 50–66. https://doi.org/10.1016/j.learninstruc.2014.04.002

Marsh, H. W., & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology*, *81*, 59–77. https://doi.org/10.1348/000709910X503501

Marsh, H. W., & Parker, J. W. (1984). Determinants of student self-concept: Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well? *Journal of*

Personality and Social Psychology, *47*(1), 213–231. https://doi.org/10.1037/0022-3514.47.1.213

Marsh, H. W., Parker, P. D., & Pekrun, R. (2019). Three paradoxical effects on academic self-concept across countries, schools, and students: Frame-of-reference as a unifying theoretical explanation. *European Psychologist*, *24*(3), 231–242. https://doi.org/10.1027/1016-9040/a000332

Marsh, H. W., & Seaton, M. (2015). The big-fish-little-pond effect, competence self-perceptions, and relativity: Substantive advances and methodological innovation. In A. Elliot (Ed.), *Advances in motivation science* (Vol. 2, pp. 127−184). Elsevier.

Marsh, H. W., Seaton, M., Trautwein, U., Lüdtke, O., Hau, K. T., O'Mara, A. J., & Craven, R. G. (2008). The big-fish-little-pond-effect stands up to critical scrutiny: Implications for theory, methodology, and future research. *Educational Psychology Review*, *20*(3), 319–350. https://doi.org/10.1007/s10648-008-9075-6

Marsh, H. W., Trautwein, U., Lüdtke, O., Baumert, J., & Köller, O. (2007). The big-fish-little-pond effect: Persistent negative effects of selective high schools on self-concept after graduation. *American Educational Research Journal*, *44*(3), 631–669. https://doi.org/10.3102/0002831207306728

Marsh, H. W., Xu, K. M., Parker, P. D., Hau, K.-T., Pekrun, R., Elliot, A., Guo, J., Dicke, T., & Basarkod, G. (2021). Moderation of the big-fish-little-pond effect: Juxtaposition of evolutionary (Darwinian-economic) and achievement motivation theory predictions based on a Delphi approach. *Educational Psychology Review*, *33*(4), 1353–1378. https://doi.org/10.1007/s10648-020-09583-5

McFarland, C., & Buehler, R. (1995). Collective self-esteem as a moderator of the frog-pond

effect in reactions to performance feedback. *Journal of Personality and Social Psychology*, *68*(6), 1055–1070. https://doi.org/10.1037//0022-3514.68.6.1055

Miller, D. I. (2019). When do growth mindset interventions work? *Trends in Cognitive Sciences*, *23*(11), 910–912. https://doi.org/10.1016/j.tics.2019.08.005

Moons, W. G., Chen, J. M., & Mackie, D. M. (2017). Stereotypes: A source of bias in affective and empathic forecasting. *Group Processes & Intergroup Relations*, *20*(2), 139–152. https://doi.org/10.1177/1368430215603460

Morse, S., & Gergen, K. J. (1970). Social comparison, self-consistency, and the concept of self. *Journal of Personality and Social Psychology*, *16*(1), 148–156. https://doi.org/10.1037/h0029862

Mussweiler, T. (2003). Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review*, *110*(3), 472–489. https://doi.org/10.1037/0033-295X.110.3.472

Mussweiler, T. (2020). How social comparison affects the self: The selective accessibility mechanism. In J. Suls, R. L. Collins, & L. Wheeler (Eds.), *Social comparison, judgment, and behavior*. (pp. 32–51). Oxford University Press. https://doi.org/10.1093/oso/9780190629113.003.0002

Nagengast, B., & Marsh, H. W. (2012). Big fish in little ponds aspire more: Mediation and cross-cultural generalizability of school-average ability effects on self-concept and career aspirations in science. *Journal of Educational Psychology*, *104*, 1033–1053. https://doi.org/10.1037/a0027697

Oishi, S., Kesebir, S., & Diener, E. (2011). Income inequality and happiness. *Psychological Science*, *22*(9), 1095–1100. https://doi.org/10.1177/0956797611417262

Park, D., & Kim, S. (2015). Time to move on? When entity theorists perform better than incremental theorists. *Personality and Social Psychology Bulletin*, *41*(5), 736–748. https://doi.org/10.1177/0146167215578028

Pekrun, R., Murayama, K., Marsh, H. W., Goetz, T., & Frenzel, A. C. (2019). Happy fish in little ponds: Testing a reference group model of achievement and emotion. *Journal of Personality and Social Psychology*, *117*(1), 166–185. https://doi.org/10.1037/pspp0000230

Plaks, J. E., & Stecher, K. (2007). Unexpected improvement, decline, and stasis: A prediction confidence perspective on achievement success and failure. *Journal of Personality and Social Psychology*, *93*(4), 667–684. https://doi.org/10.1037/0022-3514.93.4.667

Pollmann, M. M. H., & Finkenauer, C. (2009). Empathic forecasting: How do we predict other people's feelings? *Cognition and Emotion*, *23*(5), 978–1001. https://doi.org/10.1080/02699930802264895

Qian, G., Zheng, W., Liu, S., Jiang, S., Du, Y., & Yang, S. (2020). Narrowing the empathy gap between doctors and patients using emotion forecasting. *Social Behavior and Personality*, *48*(2), 1–10. https://doi.org/10.2224/sbp.8132

Rattan, A., Good, C., & Dweck, C. S. (2012). "It's ok — Not everyone can be good at math": Instructors with an entity theory comfort (and demotivate) students. *Journal of Experimental Social Psychology*, *48*(3), 731–737. https://doi.org/10.1016/j.jesp.2011.12.012

Sened, H., Lavidor, M., Lazarus, G., Bar-Kalifa, E., Rafaeli, E., & Ickes, W. (2017). Empathic accuracy and relationship satisfaction: A meta-analytic review. *Journal of Family Psychology*, *31*(6), 742–752. https://doi.org/10.1037/fam0000320

Senik, C. (2009). Direct evidence on income comparisons and their welfare effects. *Journal of Economic Behavior and Organization*, *72*(1), 408–408. https://doi.org/10.1016/j.jebo.2009.04.019

Simon, B., Kulla, C., & Zobel, M. (1995). On being more than just a part of the whole: Regional identity and social distinctiveness. *European Journal of Social Psychology*, *25*(3), 325–340. https://doi.org/10.1002/ejsp.2420250306

Snyder, K. E., Malin, J. L., Dent, A. L., & Linnenbrink-Garcia, L. (2014). The message matters: The role of implicit beliefs about giftedness and failure experiences in academic self-handicapping. *Journal of Educational Psychology*, *106*(1), 230–241. https://doi.org/10.1037/a0034553

Stockus, C. A., & Zell, E. (2023a). Do growth mindsets reduce the big-fish-little-pond effect? Manuscript submitted for publication.

Stockus, C. A., & Zell, E. (2023b). The regional big-fish-little-pond effect: Evidence from national and subnational comparisons. *British Journal of Social Psychology*, *62*(3), 1158–1176. https://doi.org/10.1111/bjso.12623

Stockus, C. A., & Zell, E. (2024a). Empathic forecasting of the big-fish-little-pond effect. *Cognition & Emotion*, *38*(1), 44–58. https://doi.org/10.1080/02699931.2023.2258581

Stockus, C. A., & Zell, E. (2024b). Moderators of the big-fish-little-pond effect in educational settings: A scoping review. *Social and Personality Psychology Compass*, *18*(1), e12860. Advance online publication. https://doi.org/10.1111/spc3.12860

Suls, J., Collins, R. L., & Wheeler, L. (Eds.). (2020). *Social comparison, judgment, and behavior*. Oxford: Oxford University Press.

Summerville, A., & Roese, N. J. (2008). Dare to compare: Fact-based versus simulation-based

comparison in daily life. *Journal of Experimental Social Psychology*, *44*(3), 664–671. https://doi.org/10.1016/j.jesp.2007.04.002

Tay, L., Morrison, M., & Diener, E. (2014). Living among the affluent: Boon or bane? *Psychological Science*, *25*(6), 1235–1235. https://doi.org/10.1177/0956797614525786

Televantou, I., Marsh, H. W., Dicke, T., & Nicolaides, C. (2021). Phantom and big-fish-little-pond-effects on academic self-concept and academic achievement: Evidence from English early primary schools. *Learning and Instruction*, *71*. https://doi.org/10.1016/j.learninstruc.2020.101399

Trautwein, U., Lüdtke, O., Marsh, H. W., Köller, O., & Baumert, J. (2006). Tracking, grading, and student motivation: Using group composition and status to predict self-concept and interest in ninth-grade mathematics. *Journal of Educational Psychology*, *98*, 788–806. https://doi.org/10.1037/0022-0663.98.4.788

Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist*, *39*(2), 111–133. https://doi.org/10.1207/s15326985ep3902_3

von Keyserlingk, L., Becker, M., & Jansen, M. (2019). Academic self-concept during the transition to upper secondary school. *Contemporary Educational Psychology*, *56*, 152–160. https://doi.org/10.1016/j.cedpsych.2019.01.001

Wheeler, L., & Miyake, K. (1992). Social comparison in everyday life. *Journal of Personality and Social Psychology*, *62*(5), 760–773. https://doi.org/10.1037/0022-3514.62.5.760

Wheeler, L., & Suls, J. (2020). A history of social comparison theory. In J. Suls, R. L. Collins, & L. Wheeler (Eds.), *Social comparison, judgment, and behavior*. (pp. 5–31). Oxford University Press. https://doi.org/10.1093/oso/9780190629113.003.0001

Wood, J. V. (1989). Theory and research concerning social comparisons of personal attributes. *Psychological Bulletin*, *106*(2), 231–248. https://doi.org/10.1037/0033-2909.106.2.231

Wu, K., Garcia, S. M., & Kopelman, S. (2018). Frogs, ponds, and culture: Variations in entry decisions. *Social Psychological and Personality Science*, *9*(1), 99−106. https://doi.org/10.1177/1948550617706731

Yeager, D. S., & Dweck, C. S. (2020). What can be learned from growth mindset controversies? *American Psychologist*, *75*(9), 1269–1284. https://doi.org/10.1037/amp0000794

Yeager, D. S., Hanselman, P., Walton, G. M., Murray, J. S., Crosnoe, R., Muller, C., Tipton, E., Schneider, B., Hulleman, C. S., Hinojosa, C. P., Paunesku, D., Romero, C., Flint, K., Roberts, A., Trott, J., Iachan, R., Buontempo, J., Yang, S. M., Carvalho, C. M., Hahn, P. R., … Dweck, C. S. (2019). A national experiment reveals where a growth mindset improves achievement. *Nature*, *573*(7774), 364–369. https://doi.org/10.1038/s41586-019-1466-y

Zell, E., & Alicke, M. D. (2009). Contextual neglect, self-evaluation, and the frog-pond effect. *Journal of Personality and Social Psychology*, *97*(3), 467−482. https://doi.org/10.1037/a0015453

Zell, E., & Alicke, M. D. (2010). The local dominance effect in self-evaluation: Evidence and explanations. *Personality and Social Psychology Review*, *14*(4), 368–384. https://doi.org/10.1177/1088868310366144

Zell, E., & Alicke, M. D. (2020). Effects of local and general comparisons on self-assessment. In J. Suls, R. L. Collins, & L. Wheeler (Eds.), *Social comparison, judgment, and behavior* (pp. 141-175). New York: Oxford University Press.

https://doi.org/10.1093/oso/9780190629113.003.0006

Zell, E., & Lesick, T. L. (2021). Taking social comparison to the extremes: The huge-fish-tiny-pond effect in self-evaluations. *Social Psychological and Personality Science*, *12*(6), 1030–1038. https://doi.org/10.1177/1948550620956535

Zell, E., & Strickhouser, J. E. (2020). Comparisons across dimensions, people, and time: On the primacy of social comparison in self-evaluations. *Social Psychological and Personality Science*, *11*(6), 791–800. https://doi.org/10.1177/1948550619884564

Zell, E., Strickhouser, J. E., & Alicke, M. D. (2017). Local dominance effects on self-evaluations and intrinsic motivation. *Self and Identity*, *16*(5), 629–644. https://doi.org/10.1080/15298868.2017.1295100

Zell, E., Strickhouser, J. E., & Krizan, Z. (2018). Subjective social status and health: A meta-analysis of community and society ladders. *Health Psychology*, *37*(10), 979–987. https://doi.org/10.1037/hea0000667

Zhao, Q., Gong, L., & Chen, T. (2021). Impacts of growth mindset and concern with performance on resilience to self-doubt in American and Chinese students. *Social Psychology of Education*, *24*(6), 1585–1606. https://doi.org/10.1007/s11218-021-09671-6

APPENDIX A: TABLES

Table A-A1. Sample Verbal Reasoning Feedback Provided to Participants in Studies A1-A3

Study (Reference Group)	Sample Feedback: BFLP Condition
Study A1 (Nation)	You performed better than 68% of participants from USA. 206 participants from USA have participated so far. As a group, USA performed better than 32% of other countries. Participants from 39 countries have participated so far.
Study A2 (Subnation)	You performed better than 68% of participants from Florida. 206 participants from Florida have participated so far. As a group, Florida performed better than 32% of other states. Participants from 39 states have participated so far.
Study A3 (School)	You performed better than 68% of UNCG students. 206 UNCG students have participated so far. As a group, UNCG performed better than 32% of universities. Students from 39 universities have participated so far.

Note. BFLP = big-fish-little-pond. Subnation = specific "state" of residence within the USA (i.e., California, Texas, Florida, or New York).

Table A-A2. Effects of Feedback and Participant Subnation in Study A2

Outcome	Effect	Result
Self-Evaluations	Feedback	$F(1, 204) = 127.98$, $p < .001$, $\eta_p^2 = .39$
	Subnation	$F(3, 204) = 2.44$, $p = .065$, $\eta_p^2 = .04$
	Feedback*Subnation	$F(3, 204) = 0.54$, $p = .656$, $\eta_p^2 < .01$
Affect	Feedback	$F(1, 204) = 120.19$, $p < .001$, $\eta_p^2 = .37$
	Subnation	$F(3, 204) = 0.50$, $p = .685$, $\eta_p^2 < .01$
	Feedback*Subnation	$F(3, 204) = 0.89$, $p = .448$, $\eta_p^2 = .01$

Note. Feedback = feedback condition (BFLP, LFBP). Subnation = participant subnation (CA, TX, FL, NY).

Table A-A3. Effects of Group Identity, Feedback, and Reference Group in Study A2

Outcome	Effect	Result
Self-Evaluations	Identity	$F(1, 204) = 0.49$, $p = .486$, $\eta_p^2 < .01$
	Feedback	$F(1, 204) = 7.44$, $p = .007$, $\eta_p^2 = .04$
	Reference	$F(1, 204) = 0.09$, $p = .768$, $\eta_p^2 < .01$
	Identity*Feedback	$F(1, 204) = 1.00$, $p = .319$, $\eta_p^2 < .01$
	Identity*Reference	$F(1, 204) = 0.66$, $p = .418$, $\eta_p^2 < .01$
	Feedback*Reference	$F(1, 204) = 2.48$, $p = .117$, $\eta_p^2 = .01$
	Identity*Feedback*Reference	$F(1, 204) = 3.34$, $p = .069$, $\eta_p^2 = .02$
Affect	Identity	$F(1, 204) = 0.07$, $p = .795$, $\eta_p^2 < .01$
	Feedback	$F(1, 204) = 5.65$, $p = .018$, $\eta_p^2 = .03$
	Reference	$F(1, 204) = 0.96$, $p = .330$, $\eta_p^2 < .01$
	Identity*Feedback	$F(1, 204) = 1.80$, $p = .181$, $\eta_p^2 < .01$
	Identity*Reference	$F(1, 204) = 1.75$, $p = .188$, $\eta_p^2 < .01$
	Feedback*Reference	$F(1, 204) = 7.47$, $p = .007$, $\eta_p^2 = .04$
	Identity*Feedback*Reference	$F(1, 204) = 7.56$, $p = .007$, $\eta_p^2 = .04$

Note. Identity = group identity, Feedback = feedback condition (BFLP, LFBP), Reference = reference group (nation, subnation).

Table A-A4. Effects of Group Identity, Feedback, and Reference Group in Study A3

Outcome	Effect	Result
Self-Evaluations	Identity	$F(1, 215) = 0.63$, $p = .427$, $\eta_p^2 < .01$
	Feedback	$F(1, 215) = 4.15$, $p = .043$, $\eta_p^2 = .02$
	Reference	$F(1, 215) = 0.00$, $p = .956$, $\eta_p^2 < .01$
	Identity*Feedback	$F(1, 215) = 0.01$, $p = .906$, $\eta_p^2 < .01$
	Identity*Reference	$F(1, 215) = 0.00$, $p = .960$, $\eta_p^2 < .01$
	Feedback*Reference	$F(1, 215) = 1.34$, $p = .248$, $\eta_p^2 < .01$
	Identity*Feedback*Reference	$F(1, 215) = 1.86$, $p = .174$, $\eta_p^2 < .01$
Affect	Identity	$F(1, 215) = 0.42$, $p = .516$, $\eta_p^2 < .01$
	Feedback	$F(1, 215) = 1.59$, $p = .209$, $\eta_p^2 < .01$
	Reference	$F(1, 215) = 1.67$, $p = .198$, $\eta_p^2 < .01$
	Identity*Feedback	$F(1, 215) = 0.51$, $p = .475$, $\eta_p^2 < .01$
	Identity*Reference	$F(1, 215) = 0.90$, $p = .343$, $\eta_p^2 < .01$
	Feedback*Reference	$F(1, 215) = 0.02$, $p = .891$, $\eta_p^2 < .01$
	Identity*Feedback*Reference	$F(1, 215) = 0.02$, $p = .886$, $\eta_p^2 < .01$

Note. Identity = group identity, Feedback = feedback condition (BFLP, LFBP), Reference = reference group (nation, school).

Table A-C1. Sample Performance Feedback Provided to Participants in Studies C1-C3

Study (Perspective)	Sample Feedback: BFLP Condition
Study C1 (Experiencer)	You performed better than 68% of UNCG students. 206 UNCG students have participated so far. As a group, UNCG performed better than 32% of universities. Students from 39 universities have participated so far.
Study C1 (Forecaster)	Alex performed better than 68% of UNCG students. 206 UNCG students have participated so far. As a group, UNCG performed better than 32% of universities. Students from 39 universities have participated so far.
Study C2	Alex performed better than 68% of Pembroke students. 206 Pembroke students have participated so far. As a group, Pembroke performed better than 32% of universities. Students from 39 universities have participated so far.
Study C3	Alex performed better than 65% of Pembroke students. 206 Pembroke students have participated so far. As a group, Pembroke performed better than 35% of universities. Students from 39 universities have participated so far.

Note. BFLP = big-fish-little-pond. Studies C2 and C3 only had forecaster conditions.

Table A-C2. Effect of Affect Type, Feedback, and Perspective in Study C1

Effect	Result
Affect Type	$F(1, 234) = 100.31$, $p < .001$, $\eta_p^2 = .30$
Feedback	$F(1, 234) = 100.84$, $p < .001$, $\eta_p^2 = .30$
Perspective	$F(1, 234) = 1.13$, $p = .289$, $\eta_p^2 < .01$
Affect Type*Feedback	$F(1, 234) = 10.72$, $p = .001$, $\eta_p^2 = .04$
Affect Type*Perspective	$F(1, 234) = 34.64$, $p < .001$, $\eta_p^2 = .13$
Feedback*Perspective	$F(1, 234) = 0.11$, $p = .745$, $\eta_p^2 < .01$
Affect Type*Feedback*Perspective	$F(1, 234) = 14.39$, $p < .001$, $\eta_p^2 = .06$

Note. Affect Type = positive affect, negative affect. Feedback = BFLP, LFBP. Perspective = experiencer, forecaster. Affect Type is a within-subjects factor. Feedback and Perspective are between-subjects factors.

Table A-C3. Effect of Affect Type, School Rank, and Student Rank in Study C2

Effect	Result
Affect Type	$F(1, 219) = 80.41$, $p < .001$, $\eta_p^2 = .27$
School Rank	$F(1, 219) = 23.47$, $p < .001$, $\eta_p^2 = .10$
Student Rank	$F(1, 219) = 494.70$, $p < .001$, $\eta_p^2 = .69$
Affect Type*School Rank	$F(1, 219) = 0.03$, $p = .860$, $\eta_p^2 < .01$
Affect Type*Student Rank	$F(1, 219) = 8.43$, $p = .004$, $\eta_p^2 = .04$
School Rank*Student Rank	$F(1, 219) = 0.73$, $p = .394$, $\eta_p^2 < .01$
Affect Type*School Rank*Student Rank	$F(1, 219) = 0.25$, $p = .620$, $\eta_p^2 < .01$

Note. Affect Type = positive affect, negative affect. School Rank = above average, below average. Student Rank = above average, below average. Affect Type is a within-subjects factor. School Rank and Student Rank are between-subjects factors.

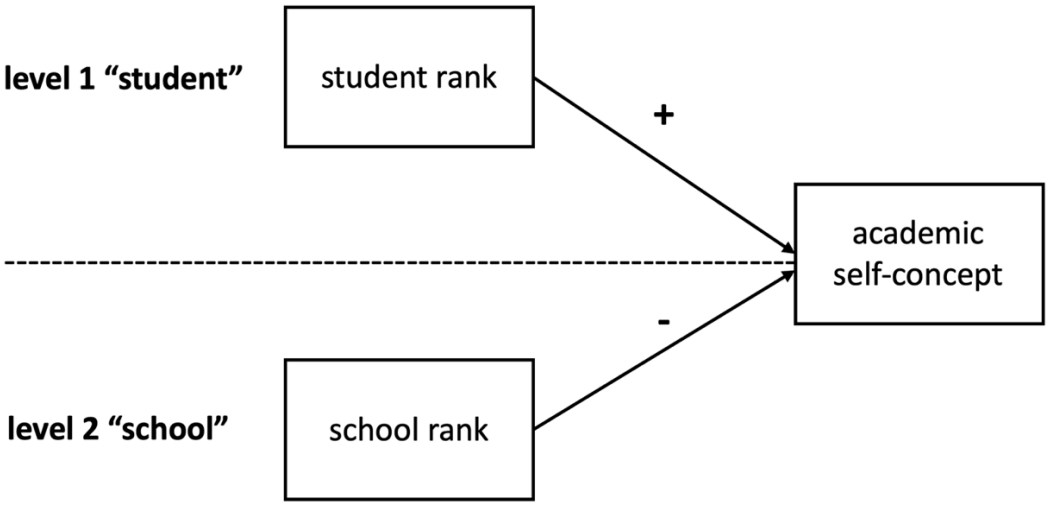
Table A-C4. Effect of Affect Type, Feedback Extremity, and Feedback Type in Study C3

Effect	Result
Affect Type	$F(1, 213) = 50.75$, $p < .001$, $\eta_p^2 = .19$
Feedback Extremity	$F(1, 213) = 0.00$, $p = .976$, $\eta_p^2 < .01$
Feedback Type	$F(1, 213) = 257.06$, $p < .001$, $\eta_p^2 = .55$
Affect Type*Feedback Extremity	$F(1, 213) = 0.16$, $p = .691$, $\eta_p^2 < .01$
Affect Type*Feedback Type	$F(1, 213) = 5.75$, $p = .017$, $\eta_p^2 = .03$
Feedback Extremity*Feedback Type	$F(1, 213) = 19.04$, $p < .001$, $\eta_p^2 = .08$
Affect Type*Feedback Extremity*Feedback Type	$F(1, 213) = 0.85$, $p = .357$, $\eta_p^2 < .01$

Note. Affect Type = positive affect, negative affect. Feedback Extremity = moderate, extreme. Feedback Type = feedback condition (BFLP, LFBP). Affect Type is a within-subjects factor. Feedback Extremity and Feedback Type are between-subjects factors.

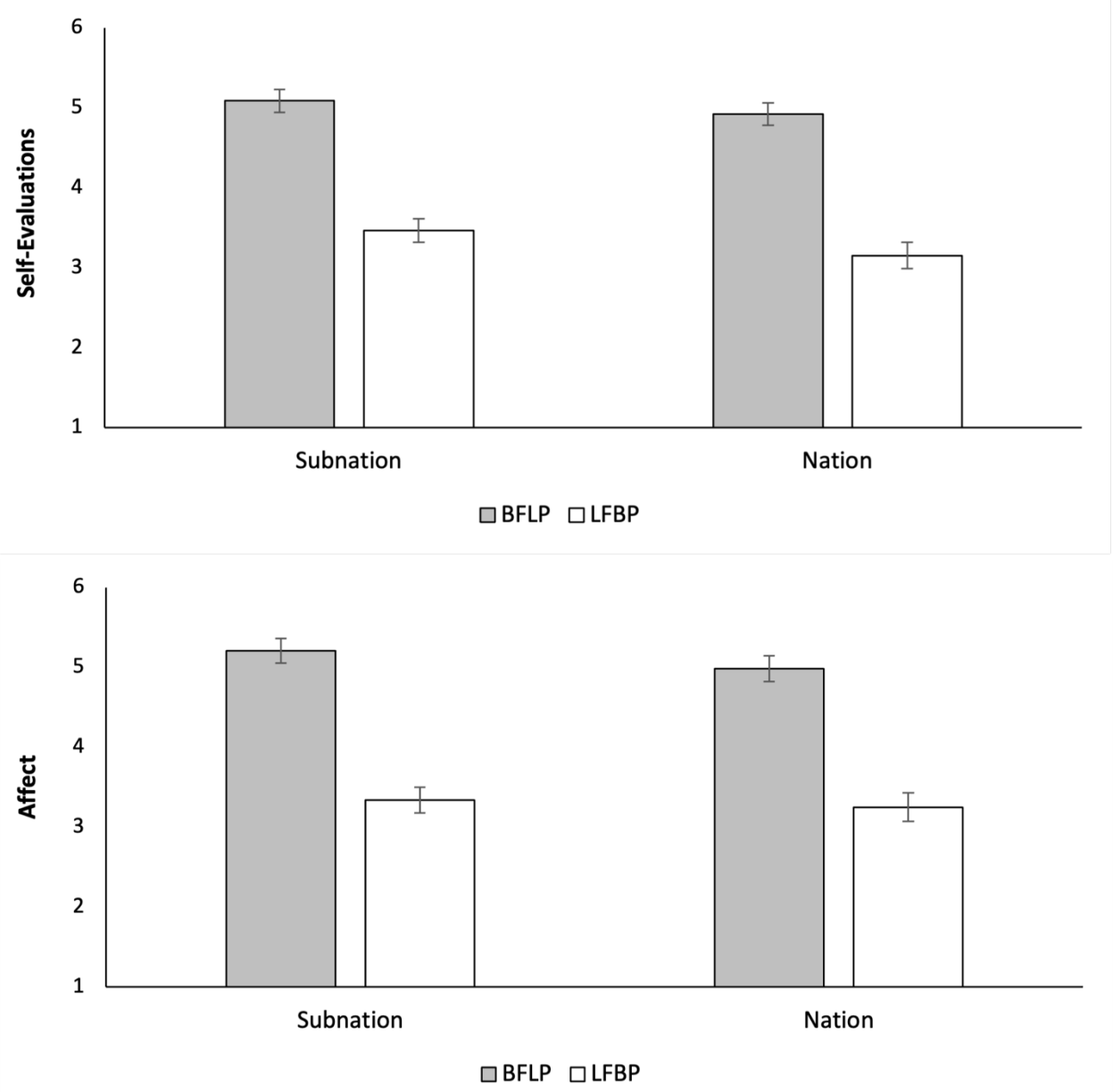
APPENDIX B: FIGURES

Figure B-B1. Visual Depiction of the BFLPE



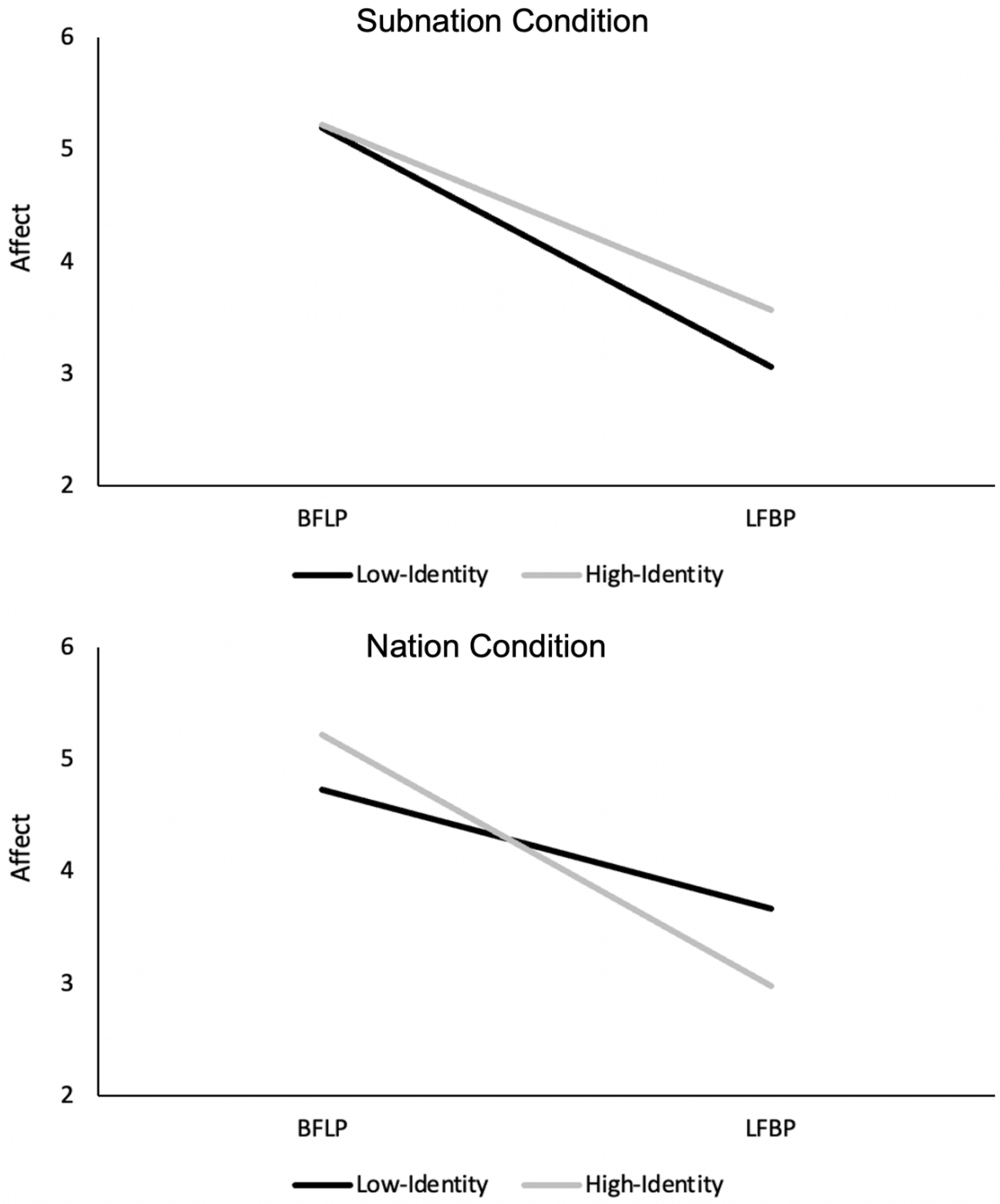
Note. BFLPE = big-fish-little-pond effect. The negative association of school rank with academic self-concepts controlling for student rank. Figure depicts a multi-level model where students (level 1) are nested within schools (level 2).

Figure B-A1. Self-Evaluations and Affect by Reference Group and Feedback Condition in Study A2



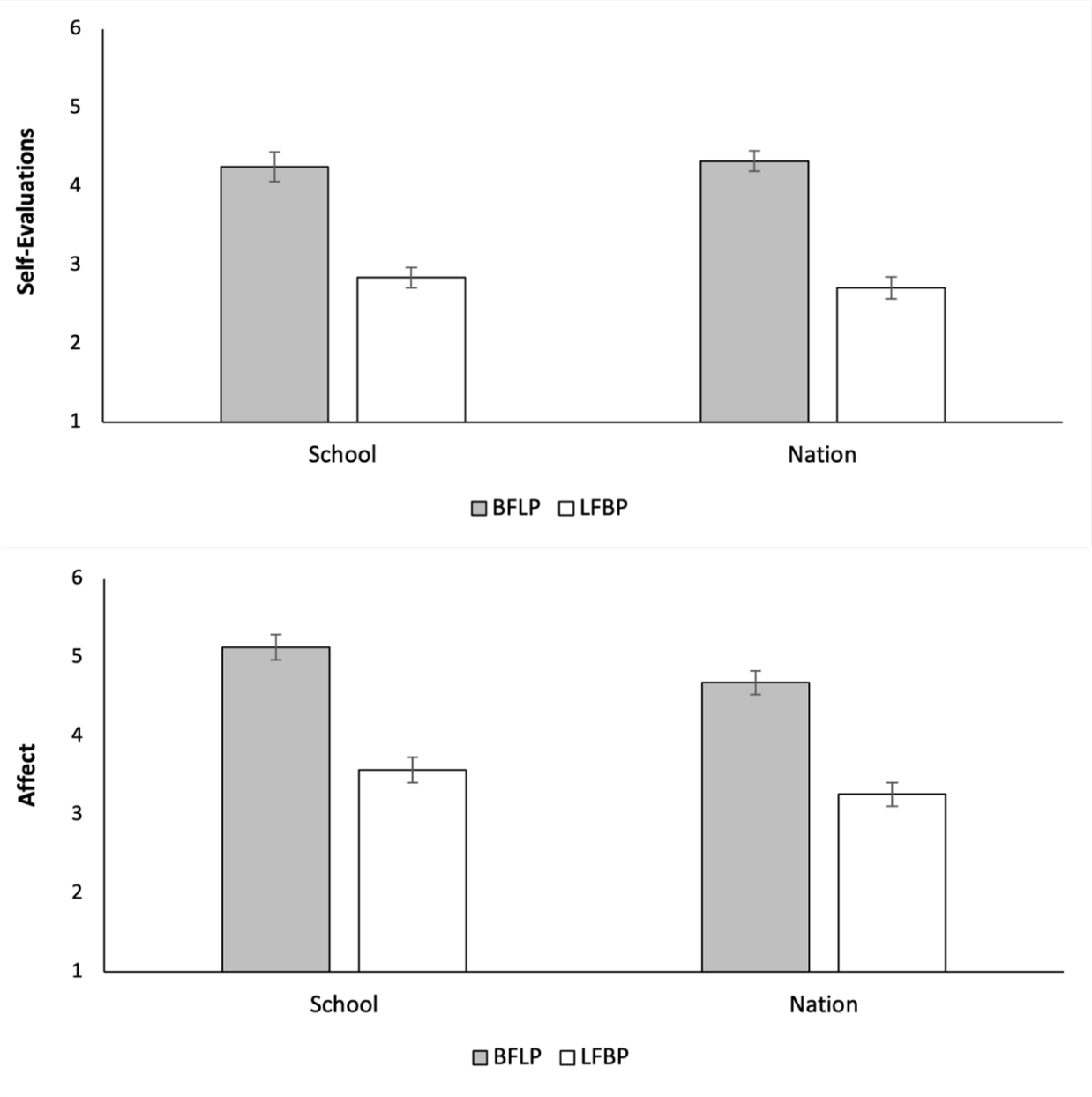
Note. BFLP = big-fish-little-pond, LFBP = little-fish-big-pond. Error bars represent ± 1 SEM.

Figure B-A2. Feedback Condition by Group Identity Plot for Affect in Study A2



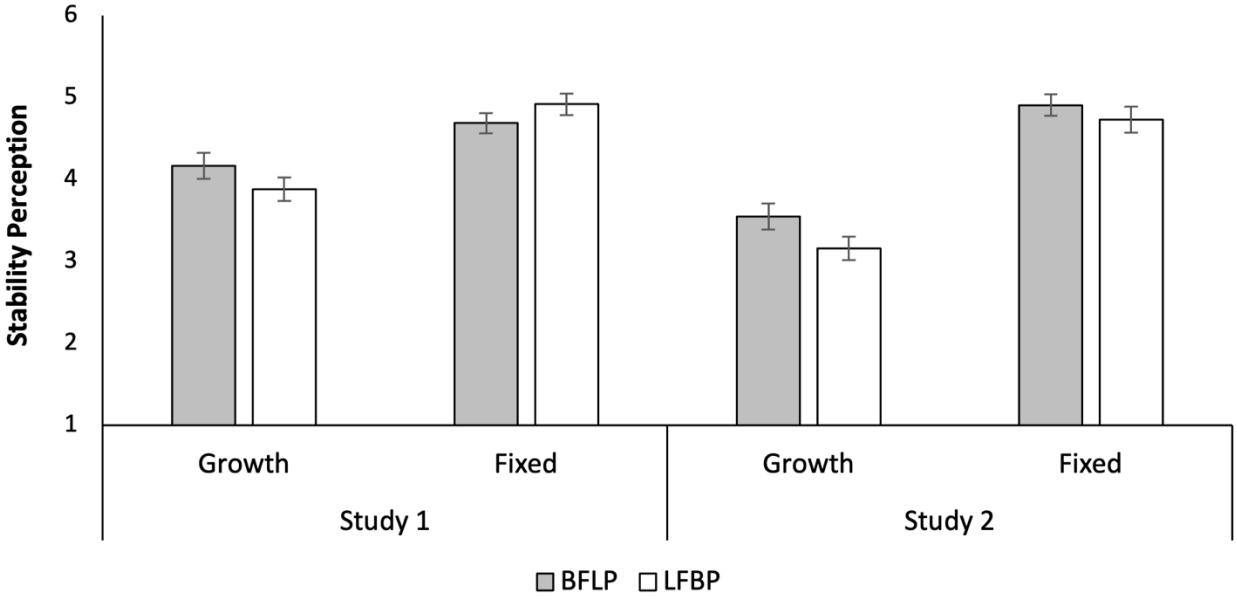
Note. BFLP = big-fish-little-pond, LFBP = little-fish-big-pond.

Figure B-A3. Self-Evaluations and Affect by Reference Group and Feedback Condition in Study A3



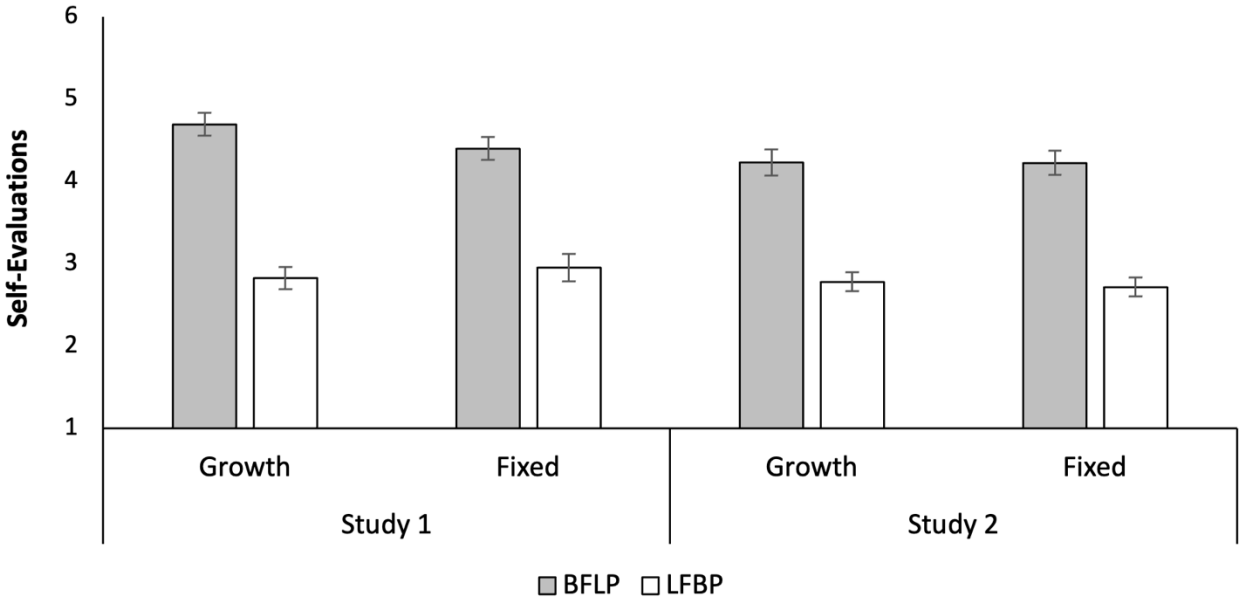
Note. BFLP = big-fish-little-pond, LFBP = little-fish-big-pond. Error bars represent ± 1 SEM.

Figure B-B1. Perceived Stability of Intelligence by Mindset Condition and Feedback Condition in Studies B1-B2



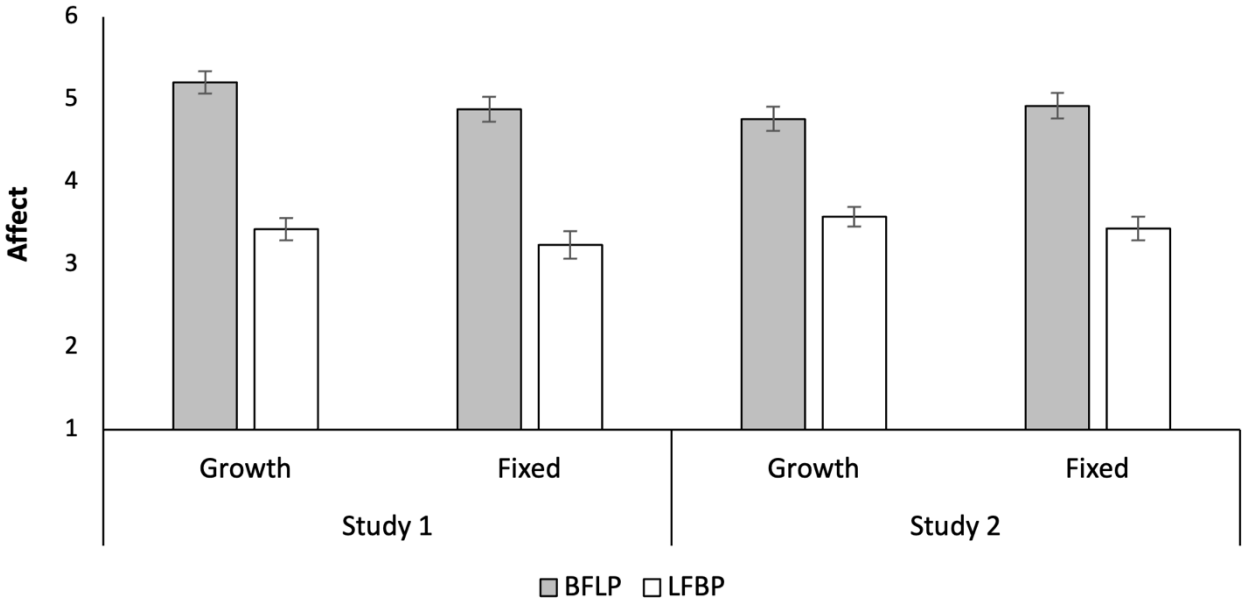
Note. BFLP = big fish in a little pond, LFBP = little fish in a big pond. Error bars represent ± 1 *SEM*. Higher values indicate greater perceived stability of intelligence.

Figure B-B2. Self-Evaluations by Mindset Condition and Feedback Condition in Studies B1-B2



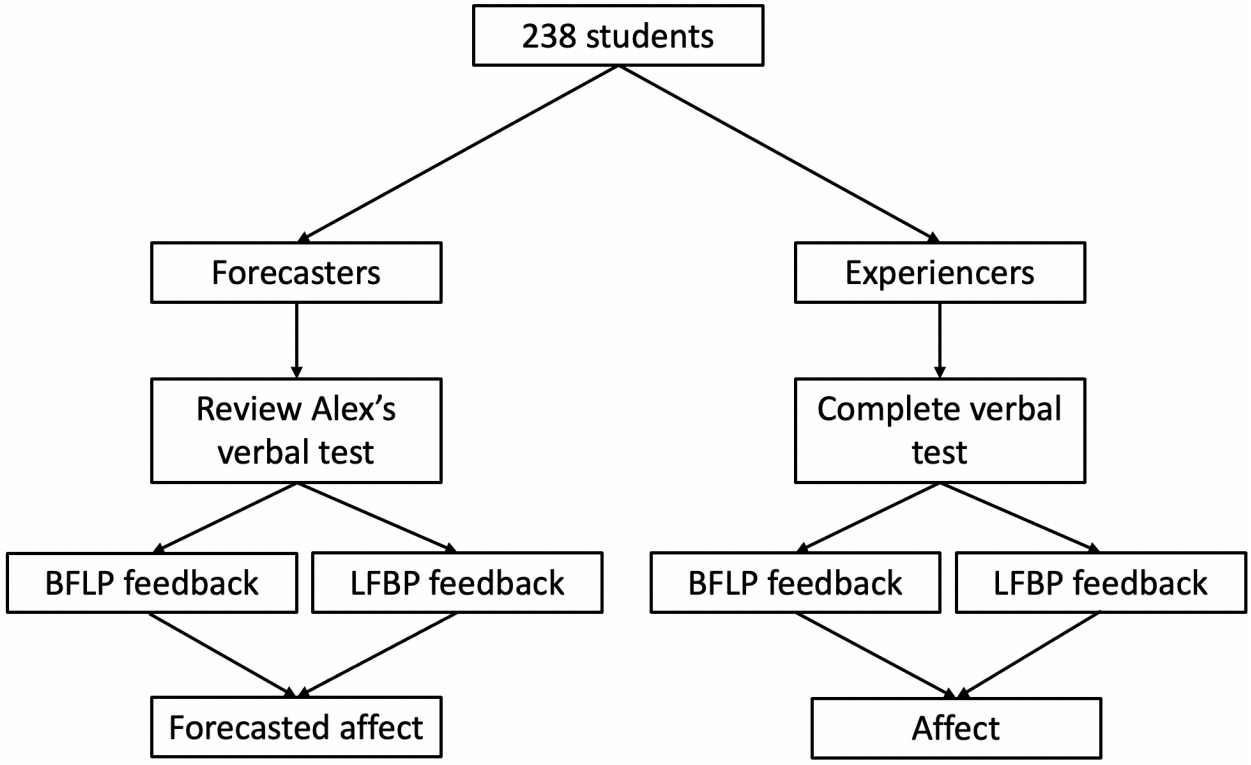
Note. BFLP = big fish in a little pond, LFBP = little fish in a big pond. Error bars represent ± 1 *SEM*.

Figure B-B3. Performance-Related Affect by Mindset Condition and Feedback Condition in Studies B1-B2



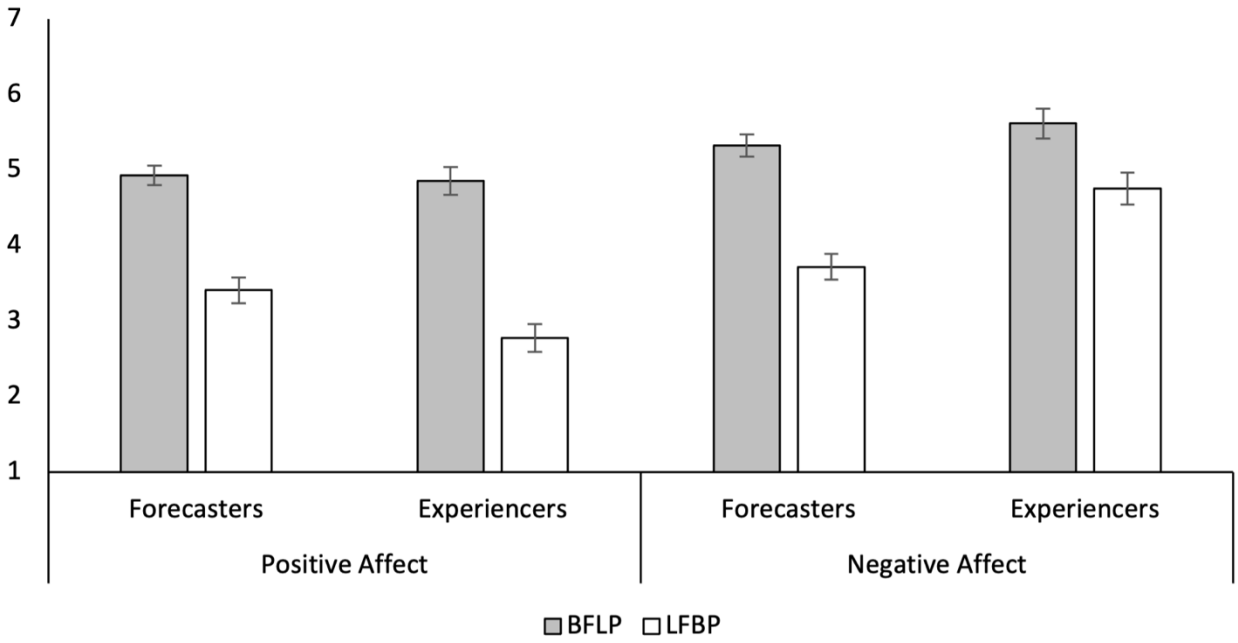
Note. BFLP = big fish in a little pond, LFBP = little fish in a big pond. Error bars represent ± 1 *SEM*.

Figure B-C1. Visual Depiction of Study C1 Procedures



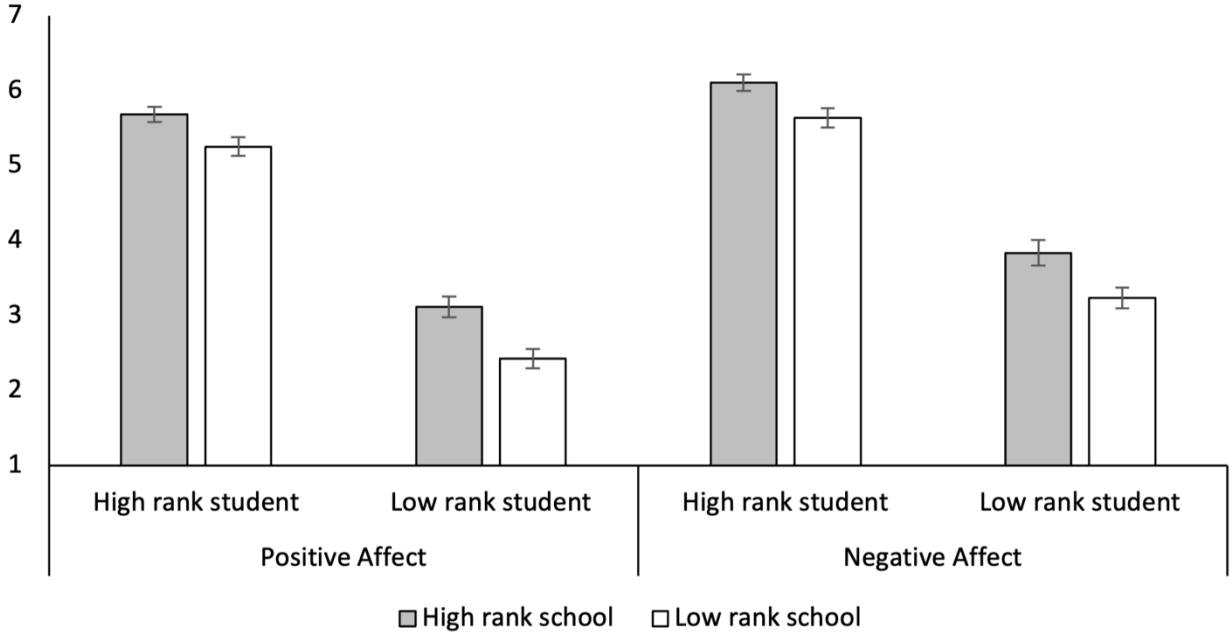
Note. BFLP = big fish in a little pond, LFBP = little fish in a big pond.

Figure B-C2. Performance-Related Affect by Affect Type, Feedback, and Perspective in Study C1



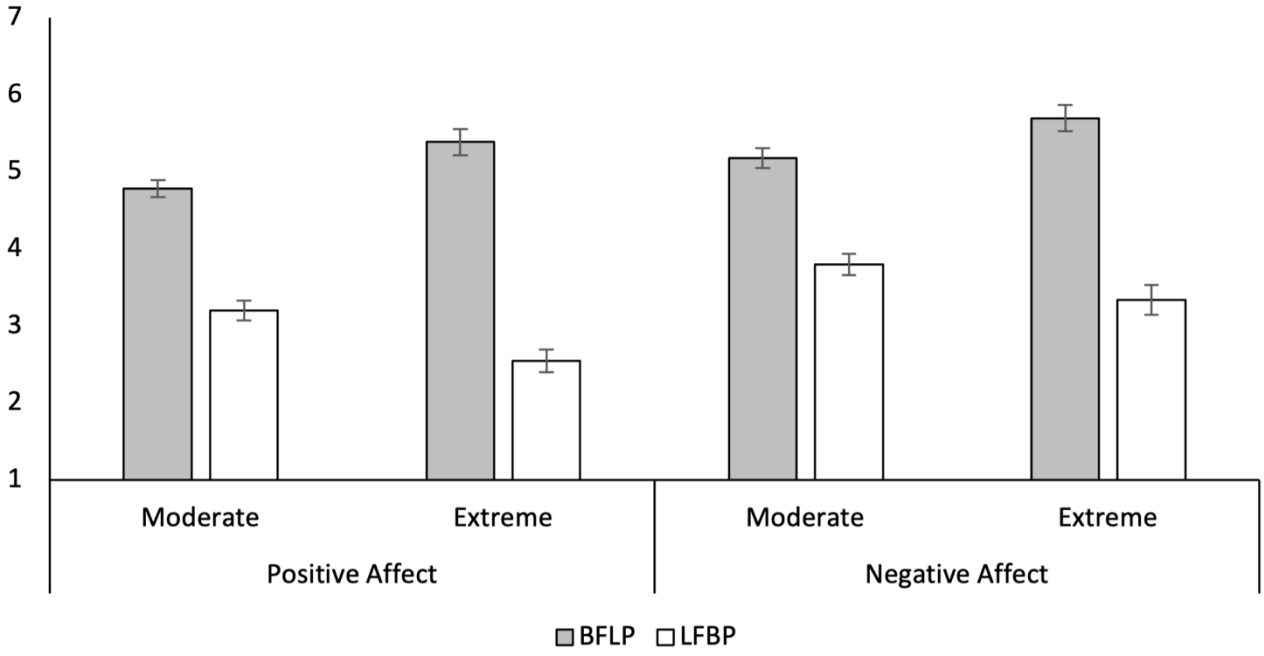
Note. Higher values mean more positive/less negative affect. BFLP = big fish in a little pond, LFBP = little fish in a big pond. Error bars represent ± 1 *SEM*.

Figure B-C3. Performance-Related Affect by Affect Type, School Rank, and Student Rank in Study C2



Note. Higher values mean more positive/less negative affect. Error bars represent ± 1 *SEM*.

Figure B-C4. Performance-Related Affect by Affect Type, Feedback Extremity, and Feedback Type in Study C3



Note. Higher values mean more positive/less negative affect. BFLP = big fish in a little pond, LFBP = little fish in a big pond. Error bars represent ± 1 *SEM*.