Big-Fish-Little-Pond Effect versus Positive Effect of Upward Comparisons in the Classroom: How does one reconcile contradictory results?

Julien Chanal, Philippe Sarrazin

To cite this version:
Julien Chanal, Philippe Sarrazin. Big-Fish-Little-Pond Effect versus Positive Effect of Upward Comparisons in the Classroom: How does one reconcile contradictory results?. Revue Internationale de Psychologie Sociale = International review of social psychology, 2007, 20 (1), pp.69-86. hal-00389735

HAL Id: hal-00389735
https://hal.science/hal-00389735
Submitted on 29 May 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Big-Fish-Little-Pond Effect versus Positive Effect of Upward Comparisons in the Classroom:

How does one reconcile contradictory results?

Julien P. Chanal, and Philippe G. Sarrazin
University J. Fourier of Grenoble, France

International Review of Social Psychology/ Revue Internationale de Psychologie Sociale

Julien Chanal and Philippe Sarrazin are with the Laboratory “Sport et Environnement Social”, Université J. Fourier, Grenoble, France.
Correspondence concerning this article should be addressed to Dr. Philippe Sarrazin, "Laboratoire Sport et Environnement Social" E.A. 3742, UFRAPS - Université J. Fourier, Grenoble I. BP 53 - 38041 Grenoble Cedex 9 - France. E-mail: philippe.sarrazin@ujf-grenoble.fr
Abstract

Studies investigating social comparison in the classroom have lead to contradictory results concerning upward comparison effects. Research demonstrates that they can lead either to enhancing student’s academic performance without influencing his/her self-concept or to decreasing student’s academic self-concept through Big-Fish-Little-Pond Effect (i.e., BFLPE). Our study tries to reconcile these results by further investigating social comparison effects relative to various frames of reference according to Stapel and Suls (2004) postulates. Effects of implicit or explicit social comparison on student’s self-concept and performance in Physical Education classes were thus considered simultaneously in the same study. Multilevel modelling analyses results demonstrate simultaneously positive and negative effects depending on the outcome and on the explicitness or implicitness of frame of reference considered. Counterbalancing the negative effects of the class-average level demonstrated in BFLPE studies, our results clearly support for assimilation effects of the small group comparison explicitly selected by the student on both self-concept and performance.

Key words: social comparison, Big-Fish-Little-Pond effect, academic self-concept educational context, upward comparison, group effect
Introduction

Social comparison processes have been pointed out to provide standards of comparison used by individuals in self-evaluations (e.g., Festinger, 1954) and in particular among key antecedents of academic self-concept (e.g., Skaalvik, 1997). In the educational context, different frames of reference can be used by students: the whole class, a small group of peers, or a particular classmate (Skaalvik & Skaalvik, 2002). In the literature, effects of these different frames of reference are quite divergent. As Wheeler and Suls (2004) recently emphasized, an apparent discrepancy exists between two set of results concerning social comparison effects in the classroom. A large number of studies (e.g., Marsh & Hau, 2003) investigated and supported the negative effects of class-average achievement level on individual student’s self-concept in the Big-Fish-Little-Pond Effect model (BFLPE, e.g., Marsh, 1987). By contrast, other researches (Blanton, Buunk, Gibbons, & Kuyper, 1999; Huguet, Dumas, Monteil, & Genestoux, 2001) recently showed positive effects of social comparison with one (or two) classmate(s) on student’s performance. This paper aims to reconcile these two seemingly divergent results and to further understand social comparison effects of various frames of reference on self-evaluations and performance in the classroom.

The Big-Fish-Little-Pond Effect: The group as frame of reference

BFLPE occurs when equally able students have lower self-concepts when they compare themselves to more able students and higher self-concepts when they compare themselves with less able students. For example, if an average-ability student is in a class of highly able students, this student’s academic abilities would be lower than the average of the other students in this class, and this discrepancy would lead to academic self-concepts that are below average. Conversely, if the student is in a class of less able students, then this student’s academic ability would be above the average of the other students in the class, and that difference would lead to academic self-concepts that are above average. According to the
BFLPE model, academic self-concept is positively correlated with individual achievement, but negatively related to class-average achievement. These hypotheses are illustrated in bold lines in Figure 1. A considerable number of research now exists in support of these predictions (e.g., Marsh & Craven, 2002, for review). These effects have been estimated in various cultures and domains (e.g., Marsh and Hau, 2003) including the physical activity domain (Chanal, Marsh, Sarrazin & Bois, 2005).

**Choices of comparison target: The individual as frame of reference**

Other researchers (Blanton et al., 1999; Huguet et al., 2001) have check that choosing to compare upward with a better performing target results in improved performance in naturalistic educational contexts. For example, Blanton et al (1999) ran a study investigating the effects of choices of comparison targets. Students were asked to name a classmate in seven disciplines with whom they compared their grades. It was revealed that the comparison target had slightly better grades than the student who had chosen him/her, indicating an upward comparison tendency. Grades obtained by the comparison target were also put in bond with student's subsequent own grades. Results demonstrated that they predicted obtaining higher grades for the subject, controlling for his/her previous grades. These results appeared for the first choice of comparison (Blanton et al., 1999) and were extended to the second ones (Huguet et al., 2001). These hypotheses are illustrated by the dotted lines in Figure 1.

How can one explain of divergent effects of upward comparison on self-evaluation?

BFLPE studies have clearly demonstrated that upward social comparison induced negative self-evaluations (e.g., Marsh & Hau, 2003). The central interest of Blanton et al. (1999) and Huguet et al. (2001) was concerning the student's performance (i.e., grades), however they did not completely ignore self-perceptions. They measured a similar variable called
comparative evaluation which corresponds to the evaluation of one’s ability compared to others. They found that this variable was not predicted by the grades of the comparison target but only by participants’ own grades. In other words, comparing upward did not lower comparative evaluation. As pointed out by Wheeler and Suls (2004), “there is a potential problem here because upward comparison should lead logically to lower comparative evaluation. If individuals are comparing with people better than themselves, they should be less likely to claim that they are better than their peers” (p. 572). Upward comparison seems to lead to lower self-concepts when the frame of reference is the class (i.e., BFLPE) and to higher grades without influencing student’s self-concept when the frame of reference is a target choice. How can we explain why upward comparison did not lead to the same negative effects on student’s self-concept and produce positive effects on student’s grades?

One of the causes of conflicting results could be found in the different methodologies used across studies. Stapel and Suls (2004) demonstrated that “method matters” when social comparison processes are investigated. Specifically, these authors carried out five studies to demonstrate effects of social comparison on activation, behaviour and self-views differ in function with the implicit or explicit nature of the social comparison. The protocol of this five-study design included presenting implicit and explicit social comparisons to the subjects before they assessed a similarity focus test (study 1), working on a lexical decision task (Study 2), and answering a general knowledge test (Study 3). An explicit comparison is said to occur “when a person is provided with a comparison target and a comparison is explicitly asked for” whereas an implicit comparison appears “when relevant comparison others are identified immediately in the absence of an explicit directive or force to compare” (Stapel & Suls, 2004, p. 861). In the explicit comparison conditions, participants were then asked to compare themselves (e.g., “Are you more or less intelligent than XX”) with specific others that have been presented to them (Britney Spears, Marie Curie, or Pamela Anderson Lee).
whereas in the implicit comparison conditions this direct comparison induction was absent.

Results demonstrated that implicit comparisons result in contrast (i.e., a negative effect) and explicit comparisons produce assimilation (i.e., a positive effect) whether it was for behaviour or self-views.

This distinction between social comparison processes might explain why previous results are contradictory. According to Stapel and Suls (2004), asking students to choose a comparison target might have activated trait priming and similarity focus with the selected target and so on exerted assimilation effect towards him/her. Thus, the students in both low- and high-ability classes can actively compare themselves with those who have slightly higher grades. By an assimilation process, this upward comparison can increase motivation or serve as a model of effective learning practices. As a result, they can improve achievement as shown by Blanton et al. (1999) and Huguet et al. (2001). At the same time, students in high-ability classes are also involuntarily exposed to higher performing classmates. By a contrast process, they may suffer a decline in academic self-concept (Wheeler & Suls, 2004).

The present investigation

The goal of the present study was to consider both of these paradigms simultaneously to further understand social comparison processes in the classroom. As underlined by Skaalvik and Skaavik (2002): “To provide a better understanding of how social comparisons affect students’ self-concept, researchers need to consider both group and individual comparisons” (p. 237) as both are indispensable to adequately capture the complexity of social comparisons.

To our knowledge, only one study has specifically treated these two sources of social comparison (Seaton, Marsh, Dumas, Huguet, et al., 2005, 2006). Specifically, this group of researchers collaborated in the reanalysis of the studies of Blanton et al. (1999) and Huguet et al. (2001). Using multilevel modelling, preliminary analyses replicated some original results:
T1 grades, comparative evaluation and target comparison level were significant predictors of T2 grades. More interesting, when comparative evaluation was used as a dependent variable, results showed a negative and large class-average ability effect, controlling for T1 grades. Moreover, a small positive effect of the target comparison level was found, but only in the reanalysis of Huguet et al. (2001) data. Whereas the effects of class-average ability on self-evaluation were consistently negative in both studies, the effects of choosing a more able comparison target on self-evaluation were never negative and even were sometimes positive.

The first objective of the present study was thus to replicate findings from Seaton et al. (2005) in the Physical Education (PE) domain and to respond to some limitations of this previous reanalysing study. As Seaton et al. (2005) underlined, impacts of social comparison effects on student's self-concept were not the original intention of Blanton et al. (1999) and Huguet et al. (2001) studies. Therefore, instead of using a self-concept scale with strong psychometric properties, Blanton et al. (1999) and Huguet et al. (2001) used a single item variable named comparative evaluation. Potential problems and inconsistent results described above concerning effects of target comparison level on self-evaluation might be due to the use of this scale. We therefore wanted to replicate these results using an appropriate self-concept measure.

The second objective of the present study was to test various implicit and explicit social comparisons simultaneously on student’s self-concept and performance. Specifically, we tested the effect of three frames of reference – an explicit target of comparison, an explicit small group of comparison and an implicit group of comparison (i.e., the class) – to further explain the discrepancy. In accordance with Seaton et al. (2005) and Wheeler and Suls (2004), we hypothesised that upward comparisons can simultaneously contribute to improved performance as shown in Blanton et al. (1999) and Huguet et al. (2001) studies, and lower self-evaluations as predicted by BFLPE studies. More precisely, following Stapel and Suls
(2004) we presumed that (1) explicit social comparison with a selected target but also with a
selected small group would lead to assimilation effect on both student’s self-concept and
performance. At the same time, we expected (2) that implicit social comparison with all the
classmates would lead to contrast (i.e., negative BFLPE on self-concept) on student’s self-
concept as previous studies on BFLPE demonstrated (e.g., Marsh & Hau, 2003). Models
tested in this study are summarized in Figure 2.

------------- Insert Figure 2 -------------

Method

Participants and procedure

Participants included 385 students (185 boys, 185 girls and 15 not identified) from 19
classes of all levels of French high schools during the 2004-2005 school year. The average
age was 13.35 years old (SD = 1.1). The study was conducted during scheduled physical
education lessons, compulsory part of the French educational system. Data were collected
after the staff meetings and the reception of the grade reports of the first trimester.

Participants completed a questionnaire assessing their self-concept in PE, sex, age and their
targets of comparison. This questionnaire was administered in the classroom to all students
present on that day. The anonymity of the answers was assured. It was indicated to the
participants that there were no good or bad answers to the questionnaire and that each
participant must answer individually and honestly. At the same time, a teacher and two
trainees in PE per class assessed the level of each student. The teacher and trainees were
different for each class. After data collection, grades of the second trimester were recovered in
order to be used as a dependent variable.

Measures

The PE self-concept was measured with 6 items (e.g., “I am at ease in the majority of
the physical and sporting activities”) assessing the Sports Competence factor resulting from
the French version of the Physical Self-Description Questionnaire (Guerin, Marsh, & Famose, 2004). The answers were related to a Likert-scale in 6 points with markers at the ends going from 1 (False) to 6 (True). In the present investigation, there is support for construct validity of responses to this scale in that responses were internally consistent ($\alpha = .91$). The mean of the items was calculated and used as an indication of students’ PE self-concepts.

*Targets of comparison.* As in previous studies (Blanton et al., 1999; Huguet et al., 2001), the participants were asked to nominate with whom they typically compared in the class. Instead of nominating one or two students, a table enabled them to indicate the names of up to five classmates with no number being imposed. A box left the possibility of answering “nobody”.

*PE achievement.* The PE teacher and two trainees in PE carrying out a teacher-training course in the class during the year, were requested to assess each student’s PE level, answering individually to the question: “according to you, what is the PE achievement level of this student?” on a scale from 1 (Very bad) to 7 (Very good). Answers were strongly consistent ($\alpha = .89$) and the mean of the three indicators was calculated.

*Data analysis*

*Student’s grade, PE achievement and group comparison level*

In order to allot to each student his/her trimester grade, we found the names and PE achievement of each pupil based on their birthdates, their sex and their class. However, this information could not be collected when the birthdates were erroneous or when the sex did not correspond to the data collected. Thus, grades and PE achievement could only be found for 357 participants. The small group comparison level was then obtained by averaging the individual PE achievement of students nominated when they were available.

*Class-average and modified PE achievement*
As Seaton et al. (2006) underlined: “Without special instruction, intervention, or moderation, teachers tend to “grade on a curve” such that there is not much variation between classes in terms of the level assigned even when there are substantial differences between classes in terms of the ability levels of students within classes” (p. 11). This observation raises a problem for purposes of evaluating the BFLPE, because it is critical to have a class-average level that reflects the differing ability levels of the classes. Unfortunately, we couldn’t select schools involved in our study so achievement levels across the schools were not comparable.

To overcome this difficulty, we followed the procedure used by Seaton et al. (2006). Information relative to schools from which our classes came from allowed us to categorize them (relative to scores they obtained in a national standardized exam at the end of the junior high school) in 3 different levels: low (below the mean level of our sample) medium (between the mean regional level and the mean level of our sample) and high (above the regional level).

We used these categorizations to rescale the average ability levels on a 0-to-3 scale, with low = 0, medium = 1.5 and high = 3. This constant value was then added to the (within-class) standardized PE achievement level of each student, depending on the student’s class. These values were then standardized so that the grand mean across all students was zero. Following these recommendations we standardized (z-scoring) all variables to have M = 0, SD = 1 across the entire sample (see also Marsh & Rowe, 1996). Class-average level was then determined by averaging individual student achievement in each class (but not restandardizing these scores so that individual student and class-average achievement scores were in the same metric).

Statistical analysis

Descriptive analyses concerning the number of comparison targets and the direction of social comparison were initially led similarly to Blanton et al. (1999) and Huguet et al. (2001). Models of multilevel multiple regressions were then built using MLWin software.
(Rasbah, Browne, Healy, Cameron, & Charlton, 2001). Two main dependent variables were used, student’s (1) PE self-concept and (2) grades. Four models were built with student’s self-concept as the dependent variable. In model 1, gender, age, a product between those terms and the student’s individual achievement were introduced. In model 2, we added class-average achievement level. Then, we added the individual comparison target achievement (Model 3) or the small group comparison target achievement (Model 4). Concerning student’s grades as the outcome, four models have been built. In the first one (Model 5), gender, age, a product between those terms, student’s self-concept and target comparison level were introduced. In model 6, the student’s achievement level was added. Models 7 and 8 were identical to models 5 and 6 but target comparison level was replaced by small group comparison level.

Results

Descriptive analysis

Among students for whom data were available, 81 participants out of 357 did not choose anybody as a target of comparison (23%). Approximately 70% of the participants who selected at least one individual with whom to compare (i.e., 192 out of 276) chose more than two comparison targets. The number of comparison targets is higher than 2 (M = 2.59) if we consider the entire sample and 3 (M = 3.36) if we consider only those who declare to compare. The choices of comparison partners are mainly students of the same sex (87%), with boys choosing more partners of the same sex (90%) that girls (85%) do.

Choice of comparison

Table 1 presents students and comparison targets means grade as well as paired t-tests. Results plead in favour of an upward comparison tendency for the first choice (mean = 13.60 vs. students mean = 12.75, \( t = 3.85, p < .001 \)), the second choice (mean = 13.35 vs. students mean = 12.83, \( t = 2.16, p < .05 \)), or for the small group (mean = 13.36 vs. students mean = 12.86, \( t = 2.90, p < .01 \)).
Links between various frames of reference and student’s PE self-concept

Table 2 presents the results of analyses intended to test the links between frames of reference and student’s PE self-concept. Model 1 showed that student’s gender and student’s achievement ($\beta = .56$, and $\beta = .66$, $p < .05$, respectively) were positively related to student’s PE self-concept. Boys have a higher self-concept than girls and better achievement leads to better self-perceptions in PE. Model 2 demonstrated BFLPE, class-average level was negatively related to student’s PE self-concept ($\beta = -.72$, $p < .05$) whereas gender and individual achievement remain significant predictors ($\beta = .52$, and $\beta = .77$, $p < .05$, respectively). In model 3, the target comparison level appeared not statistically related to student’s PE self-concept ($\beta = .04$, $p > .05$) whereas in model 4, the small group of comparison did ($\beta = .19$, $p < .05$). In these two models, class-average level still remain statistically significant and negative ($\beta = -.79$, et $\beta = -.73$, $p < .05$, respectively).

Links between various frames of reference and student’s grades

Table 3 presents results of analyses intended to test the links between frames of reference and student’s grades. Model 5 demonstrated that student’s self-concept and target comparison level were both positively related to student’s grade ($\beta = .44$, and $\beta = .15$, $p < .05$, respectively). However, in model 6, when the student’s achievement was added, student’s self-concept and target comparison level did not reach significance ($\beta = .08$, $p = .18$ and $\beta = .09$, $p = .07$, respectively). Model 7 demonstrated that student’s self-concept and small group comparison level were both positively related to student’s grade ($\beta = .46$, and $\beta = .29$, $p < .05$, respectively). Moreover, these effects remain statistically significant and positive ($\beta = .12$, and $\beta = .15$, $p < .05$, respectively) when student’s achievement was added in model 8.
Discussion

The purpose of this study was to reconcile two sets of contradictory results concerning social comparison effects of upward comparisons. Previous studies in naturalistic educational contexts have provided contradictory evidence of potential benefits on student’s performance and negative effects on self-perceptions but have not been able to say if these processes were exclusive or not.

Social comparison choices

Results are in conformity with previous studies (Blanton et al., 1999; Huguet et al., 2001) concerning student’s upward comparison tendency. Indeed, grades of the first two targets of social comparison are slightly higher than grades of students who chose them.

Analysis of the number of comparison targets seems to indicate that the choice of one or two targets of comparison is not the most frequent. Indeed, the general mean of the number of selected partners is higher than 2 if the entire sample is considered, and 3 if only those who declare having at least one target of comparison are taken. This result confirms the existence of small group comparison as posited by Skaalvik and Skaalvik (2002). The small group comparison targets seems to have slightly better grades than the student who choose them, also reflecting an upward comparison tendency.

Links between frames of reference and student’s PE self-concept and grades

Multilevel modelling enabled us to test simultaneously the impact of variables from individual level (i.e., the target comparison and the small group of comparison level) and from class level (i.e., the average-class level).

Links with student’s PE self-concept

Results reveal important information when comparing the frames of reference under consideration concerning student’s self-concept. First of all, one of the purposes of this study was to replicate BFLPE in the PE domain. Our results clearly show negative BFLPE across
models 2 to 4. They replicate previous results in PE classes during a gymnastics training program (Chanal, et al., 2005). In addition, we wanted to test the influence of selected targets on an appropriate self-concept measure. Results corroborate previous findings of the reanalysis of Blanton et al. (1999) (Seaton, et al., 2005, 2006) concerning the target comparison level. As model 3 showed (see Table 2), target comparison level was not related to student’s PE self-concept. However, model 4 brings additional indications concerning group comparison. Results showed that PE self-concept was positively related with the small group of comparison level (i.e., the selected group) and negatively with the class-average level (i.e., the non-selected group). This result demonstrates that various frames of reference could simultaneously impact student’s self-concept. It is also in accordance with Stapel and Suls (2004) predictions concerning effects of implicit or explicit social comparison. When the comparison with a group was explicit (i.e., the selected group of comparison), assimilation occurred whereas contrast appeared when the comparison was implicit (i.e., the class). This result needs to be replicated before drawing final conclusions but it is interesting to consider that a selected group can have consequences opposed to a non-selected group.

Nevertheless, we could have assumed that explicit comparison with an individual target might also have led to assimilation whereas in fact, it did not. Previous results showed that this result is inconsistent depending on the data set and the subject (Seaton et al., 2005, 2006). Our point of view is that an individual frame of reference might be less common in student’s self-views than the favourite group is. This could be particularly true in PE courses where small groups are frequently formed (probably more than in others academic domains). Analyses of the number of comparison targets demonstrated that the majority of students chose more than two classmates with whom they compare themselves. Few students (less than 30) chose only one classmate to compare with. We thus were not able to build multilevel models specific to these students. Our assumption was that these students’ self-concepts may
have been more influenced by this single target and that an assimilation effect might have
been found.

*Links with student’s grades*

First of all, our results replicate former studies of Blanton et al. (1999) and Huguet et
al. (2001) concerning independent and positive effects of upward comparison with a target
comparison and student’s self-concept on student’s grades. The better the target comparison
level and the higher the self-concept were, the higher the grades were (model 5). However,
en these effects disappear when individual achievement was added as in model 6. This result
suggests that this effect could be an artefact of students choosing comparison targets similar
to themselves. Nevertheless, if the student’s self-concept effect seems to be totally
confounded with the effect of student’s achievement, the target comparison level still remain
fairly significant ($p = .07$) indicating a potential additive effect of this variable. We also tested
the impact of a selected group on student’s grades. Results were in accordance with our
hypothesis. Specifically and as predicted, model 7 showed that the small group of comparison
level had a positive impact on student’s grades independently of student’s self-concept.
Moreover, these effects remain significant when student’s achievement level was added to the
model. In other words, for equally able students, those who chose to compare themselves with
a better comparison group obtained better grades (i.e., assimilation effect).

*Limitations*

Some limitations of this study have to be pointed out. BFLPE studies necessitate
standardized performance measures to provide a common metric for comparing students’
level from different classes or schools. Even if this was not the case in our study, we were
able to scale classes in relation to schools they attended. To take the differences between
schools into account, we used a macroscopic indicator (i.e., scores obtained in a national
standardized exam) because PE contributes to this indicator. Nevertheless, this indicator is not
the most reliable considering PE. It would have been interesting to use the regional rank of the
school in this discipline.

Moreover, our study used correlational data and a cross-sectional design. Contrary to
Blanton et al. (1999) and Huguet et al. (2001) studies, this design did not allow us to control
longitudinally for previous grades when student’s performance was considered. However,
although we did not control these effects longitudinally, we added to our model a measure of
the student’s achievement level likely to replace grades obtained in the first trimester.

Conclusion

The ambition of the study was to reconcile two paradigms related to social comparison
processes providing paradoxical results in natural educational context. It appeared that these
processes are not exclusive but complementary. Results of this study (1) demonstrated that
various frames of reference may have simultaneous effects on student’s self-concept. They (2)
confirmed previous studies on social comparison choices (Blanton et al., 1999; Huguet et al.,
2001) and BFLPE (e.g., Marsh & Haw, 2003) concerning upward social comparison
tendency, positive effects of target comparison level on student’s performance (i.e., grades),
negative effect of class-average level and a null effect of target comparison level on student’s
self-concept, extending previous results in PE classes. Lastly, the results (3) bring new
information related to group comparison effects whether the group is explicitly selected or
not. When social comparison is explicit, a group as a frame of reference might have potential
benefits for students’ self-concept or performance. When social comparison is implicit,
negative effects of the group (i.e., negative BFLPE) on self-concept occur.
References


Effect with the positive effects of upward comparisons. Paper presented at the 2005 annual meeting of the Australian Association of Research in Education


Julien Chanal and Philippe Sarrazin are part of the Laboratory ‘Sport et Environnement Social’ E.A. 3742, UFRAPS - Université J. Fourier, Grenoble.

Correspondence concerning this article should be addressed to Dr. Philippe Sarrazin, ‘Laboratoire Sport et Environnement Social’ E.A. 3742, UFRAPS - Université J. Fourier, Grenoble I. BP 53 - 38041 Grenoble Cedex 9 - France. E-mail: philippe.sarrazin@ujf-grenoble.fr or to Julien Chanal, ‘Laboratoire Sport et Environnement Social’ E.A. 3742, UFRAPS - Université J. Fourier, Grenoble I. BP 53 - 38041 Grenoble Cedex 9 - France. E-mail: julien.chanal@ujf-grenoble.fr

We would like to acknowledge Dr. Hans Kuyper, Bram Buunk, Pascal Huguet and two anonymous reviewers for their insightful comments on an earlier version of the manuscript.
Footnotes

1. More precision concerning the procedure can be found in Seaton et al. (2005, 2006)

2. As suggested by an anonymous reviewer, we tested (1) if students who did not choose any comparison target differ from others students in their self-concept or grades and (2) if they were less prone to the BFLPE. Results of these additive analyses showed that these students did not differ from others and that the size of the BFLPE did not vary significantly for these students.
Table 1

Means ratings, standard deviations and t-test between student's grades and student's choices of comparison

<table>
<thead>
<tr>
<th></th>
<th>Student's grade</th>
<th>First comparison choice</th>
<th>t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>N</td>
<td>178</td>
<td>12.75</td>
<td>2.82</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Student's grade</th>
<th>Second comparison choice</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>N</td>
<td>158</td>
<td>12.83</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Student's grade</th>
<th>Small group of comparison choice</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>N</td>
<td>208</td>
<td>12.86</td>
</tr>
</tbody>
</table>

Note: *p < .05, **p < .01, ***p < .001 (paired t-tests). Positives t values indicate an upward comparison tendency.
Table 2

*Multilevel models with student’s PE self-concept as a dependent variable*

<table>
<thead>
<tr>
<th>Variables</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Effect</td>
<td>SE</td>
<td>Effect</td>
<td>SE</td>
</tr>
<tr>
<td>Fixed effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-.33*</td>
<td>.13</td>
<td>-.27*</td>
<td>.07</td>
</tr>
<tr>
<td>Gender*</td>
<td>.56*</td>
<td>.09</td>
<td>.52*</td>
<td>.09</td>
</tr>
<tr>
<td>Age</td>
<td>.09</td>
<td>.09</td>
<td>-.01</td>
<td>.07</td>
</tr>
<tr>
<td>Gender × Age</td>
<td>.05</td>
<td>.09</td>
<td>-.01</td>
<td>.07</td>
</tr>
<tr>
<td>Student’s achievement</td>
<td>.66*</td>
<td>.07</td>
<td>.77*</td>
<td>.07</td>
</tr>
<tr>
<td>Target Comparison Level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small group comparison Level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class-average Level</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residual variance components</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level 2 class</td>
<td>.27</td>
<td>.10</td>
<td>.03</td>
<td>.02</td>
</tr>
<tr>
<td>Level 1 students</td>
<td>.58</td>
<td>.05</td>
<td>.58</td>
<td>.05</td>
</tr>
<tr>
<td>N</td>
<td>329</td>
<td>329</td>
<td>203</td>
<td>206</td>
</tr>
</tbody>
</table>

Note: All outcome and predictor variables at the individual student level were standardized ($M = 0, SD = 1$). All parameter estimates are statistically significant when they differ from zero by more than two standard errors (SEs). *A positive score indicates that boys have better self-concept than girls.*
Table 3

Multilevel models with student’s PE grades as a dependent variable

<table>
<thead>
<tr>
<th>Variables</th>
<th>Model 5</th>
<th></th>
<th>Model 6</th>
<th></th>
<th>Model 7</th>
<th></th>
<th>Model 8</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Effect</td>
<td>SE</td>
<td>Effect</td>
<td>SE</td>
<td>Effect</td>
<td>SE</td>
<td>Effect</td>
<td>SE</td>
</tr>
<tr>
<td><strong>Fixed effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>.09</td>
<td>.13</td>
<td>.15</td>
<td>.10</td>
<td>.27*</td>
<td>.09</td>
<td>.23*</td>
<td>.09</td>
</tr>
<tr>
<td>Gender</td>
<td>-.10</td>
<td>.15</td>
<td>-.23*</td>
<td>.11</td>
<td>-.41*</td>
<td>.14</td>
<td>-.34*</td>
<td>.11</td>
</tr>
<tr>
<td>Age</td>
<td>.02</td>
<td>.10</td>
<td>.10</td>
<td>.08</td>
<td>.13</td>
<td>.08</td>
<td>.12</td>
<td>.08</td>
</tr>
<tr>
<td>Gender × Age</td>
<td>-.12</td>
<td>.13</td>
<td>-.15</td>
<td>.10</td>
<td>-.17</td>
<td>.12</td>
<td>-.12</td>
<td>.10</td>
</tr>
<tr>
<td>PE self-concept</td>
<td>.44*</td>
<td>.07</td>
<td>.08</td>
<td>.06</td>
<td>.46*</td>
<td>.07</td>
<td>.12*</td>
<td>.06</td>
</tr>
<tr>
<td>Target Comparison Level</td>
<td>.15*</td>
<td>.07</td>
<td>.09</td>
<td>.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small group comparison Level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.29*</td>
<td>.07</td>
<td>.15*</td>
<td>.06</td>
</tr>
<tr>
<td>Student’s achievement</td>
<td>.72*</td>
<td>.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Residual variance components</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level 2 class</td>
<td>.10</td>
<td>.06</td>
<td>.09</td>
<td>.05</td>
<td>.00</td>
<td>.00</td>
<td>.06</td>
<td>.03</td>
</tr>
<tr>
<td>Level 1 students</td>
<td>.68</td>
<td>.08</td>
<td>.38</td>
<td>.04</td>
<td>.63</td>
<td>.07</td>
<td>.39</td>
<td>.04</td>
</tr>
<tr>
<td>N</td>
<td>178</td>
<td>178</td>
<td>179</td>
<td>179</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. All outcome and predictor variables at the individual student level were standardized ($M=0$, $SD=1$). All parameter estimates are statistically significant when they differ from zero by more than two standard errors (SEs). * A positive score indicates that boys have better grades than girls.
Figure Caption

Figure 1. Postulates of a BFLPE model in the classroom are in bold, postulates of the upward comparison studies with dotted line.

Figure 2. Models tested in our study. Dotted lines are those tested separately.
Upward comparisons in the classroom

Class-average level

Individual achievement level

Target comparison level

Student's self concept

Student's grades

+ - + + + + +