Lost in translation : congruency of teacher and student perceptions of assessment as a predictor of intrinsic motivation in ethnodiverse classrooms

Pat El, R.J.

Cover Page

Universiteit Leiden

The handle http://hdl.handle.net/1887/19934 holds various files of this Leiden University dissertation.

Author: Pat El, Ron Jonathan
Title: Lost in translation : congruency of teacher and student perceptions of assessment as a predictor of intrinsic motivation in ethnodiverse classrooms
Issue Date: 2012-10-09

3

Multilevel Predictors of Differing Perceptions of Assessment For Learning Practices Between Teachers and Students[2]

Abstract

Assessment for Learning (AfL), as a way to promote learning, requires a 'match', or shared focus between student and teacher to be effective. But students and teachers may differ in their perception of the purpose and process of classroom assessment meant to promote learning. Perceptions regarding AfL practices in their classroom were collected from both teachers and students. Teacher efficacy for instructional strategies, student engagement, classroom management, and student self-reported language proficiency were considered possible influencing factors. Multilevel analysis on self-report questionnaires on AfL practices administered to 650 students and 38 teachers revealed a substantial mismatch in perceptions between teachers and students. Congruency of teacher-student perceptions was highly homogeneous within classes. High teacher efficacy and low student language proficiency were associated with a incongruent AfL perceptions. Findings are interpreted using the self-verification theory.

3.1 Introduction

Incongruent perceptions of Assessment for Learning practice

Various authors (P. Black & Wiliam, 1998b) have advocated using assessment as an instructional approach to improve learning processes by utilizing its possibilities for scaffolding and monitoring student progress. This integration of assessments in the learning process has been called Assessment for Learning (AfL) and is contrasted against the traditional Assessment of Learning (P. Black & Wiliam, 1998a). Although Assessment for Learning may be part and parcel of classroom instruction; students and teachers may disagree on whether it is practiced (Broadfoot, 1998). A core principle of AfL is that it functions as a two-way learning process: learners are scaffolded into learning how to progress based on their current achievement, while teachers adapt their teaching according to the garnered assessment information of the current level. The

---

extent to which AfL effectively supports students' learning is facilitated by a close 'match' or agreement between student and teacher about how assessment is utilized; students need to recognize teachers' efforts in scaffolding their learning and monitoring their progress in order to advance their learning (Popham, 2008; Sadler, 2010).

Some studies cast doubts on the congruency between teacher and student perception of AfL-practice (Birenbaum et al., 2006; McMillan, 2007), referring to conflicting perceived purposes (Popham, 2008) or stating that students preempt a formative use of assessment. Only scarce quantitative results are to be found in regards to the degree to which students and teachers align in their perception of practiced AfL. Clarifying possible mismatches between student-teacher perceptions could, in our view, contribute to the implementation of AfL. The aim of this study, therefore, is to evaluate student and teacher perceptions of actual AfL practices and find factors that may explain possible individual variance between students and teachers.

Focus on Assessment for Learning

Assessment can provide a powerful tool in enhancing learning and promoting students' motivation to learn (James & Pedder, 2006; McMillan, 2007). AfL is a process of continual interaction between teachers and individual learners, in which feedback provision and its acceptance and utilization are key elements (P. Black & Wiliam, 2009; Struyven et al., 2005; Davis, 2006). A necessary condition for students to accept feedback is to recognize that feedback is being provided. In defining the construct of AfL, common principles in most studies are identified as: (1) rich classroom questioning, (2) facilitating self- and peer assessments, (3) clarifying goals and criteria, and (4) giving feedback/feedforward (P. Black & Wiliam, 1998a; Assessment Reform Group, 2002; James & Pedder, 2006; Whitelock, 2010). As suggested by Stiggins (2005) and Pat-El, Segers, Tillema and Vedder (2011) these four general principles can be further subsumed under two functions of AfL: a) monitoring to track student progress and b) scaffolding to show or help students recognize what areas need improvement. Monitoring refers to analyzing student learning progress to foster student self-monitoring with the intent to find challenges and opportunities for optimizing learning. Scaffolding refers to classroom interaction wherein learning goals and criteria are clarified through, and in addition to, classroom questioning. These two ingredients constitute ways of feedback provision (Wiliam, 2011). In order for AfL to be supportive of learning, it needs to be recognized by students, and congruent perceptions between students and teachers are needed to optimize the benefits of feedback provision. Incongruent perceptions on whether Assessments are embedded in the learning process between teachers and students may lead to misunderstanding and misinterpretation of the assessment information (Bartholomew et al., 2001; Norman, 1986). Several studies (MacLellan, 2001; Raviv, Raviv, & Reisel, 1990) found differing perceptions of the formative nature of assessments between students and their teachers, in which teachers tend to perceive their assessments as formative, whereas students report more summative assessment methods. Students evaluate the formative nature of assessments to be more implicit and 'hidden' than their teachers (Könings, 2007). Moreover, students often reported they did not convincingly perceive clear goals, room for personalization of learning, or

fascinating subject-contents, teachers reported they did perceive these aspects to be present in daily educational praxis (Könings, 2007).

Unclear are the factors that influence the congruency of perceived AfL. The current study looks at promising explanatory constructs: (1) teachers' self reflective barriers through teacher efficacy, which refers to teachers' beliefs in their capability to successfully accomplish teaching tasks, (Allinder, 1994) and (2) students' language proficiency as a proxy for the students' understanding and appreciating AfL information.

Teacher Efficacy

Teachers' beliefs, particularly about their personal effectiveness or efficacy, have been found to contribute to their effectiveness and goal attainment (Ashton & Webb, 1986). Teacher with high efficacy were oriented towards high students' achievement and success (Guskey, 1988). Perceived self-efficacy has been defined as personal judgments of one's competency to produce desired effects by one's actions (Bandura, 1997). Tschannen-Morann and Hoy (1998) defined three domains of teacher-specific competency beliefs: (1) efficacy for instructional strategies, (i.e., teachers' belief in being able to use varied assessment strategies, respond to difficult questions, or engage students at adequate levels of competence); (2) efficacy for classroom management, (i.e., maintaining classroom rules and order); and (3) efficacy for student engagement, (i.e., getting students motivated to learn). Teacher efficacy has been linked to effective classroom instruction (Eren, 2009), openness to new ideas (Allinder, 1994), and enthusiasm for classroom instruction (Hall, Burley, Villeme, & Brockmeijer, 1992). However, Gerges (2001) showed that teacher efficacy may block flexibility and variation in employing instructional strategies, such as explorative classroom questioning. High efficacious teachers seem to show more rigid use of teaching strategies and content coverage, (Wheatly, 2002). Mastery experiences contribute to beliefs of efficacy (Bandura, 1997), but mastery experiences themselves remain a subjective experience, wherein self-efficacious beliefs themselves can fuel a confirmation bias for mastery experiences, in turn boosting one's own self-efficacious beliefs (Mynatt, Doherty, & Tweney, 1977). This process of self-verification (Swann Jr., Chang-Schneider, & Angulo, 2007) describes how people are more apt to seek information that confirms one's self-beliefs, in order to avoid incongruency between their perceived self and their experienced self (Carver & Scheier, 2000). In this respect, teachers high on teacher efficacy beliefs would focus more on particular classroom signals that boost their efficacy beliefs, and focus less on information that is contrary to those. The prime focus on congruency with personal intentions might run the risk of missing student signals important for alignment in assessment perceptions. AfL requires teachers to actively adapt their teaching based on student information. Thus, teachers high on teacher efficacy beliefs might be at risk of missing out on cues signaling them to adapt their teaching. In line with the theory of self-verification we expect teachers with high teacher efficacy beliefs, mainly beliefs concerned with the successful implementation instructional strategies, to perceive more AfL practiced in their classroom than their students do.

Language proficiency mediating assessment for learning

Formative feedback, as key ingredient of AfL, needs to be understood by the student and recognized as supportive for learning (Bartholomew et al., 2001). Assessment practices in classrooms build on the students' ability to grasp meaning and purpose of the information provided to them by, in this case, the AfL strategies of monitoring (i.e., appreciation of strength and weaknesses) and scaffolding (action on learning routes) (Sadler, 2010). We assume that student language proficiency becomes an important determinant to understanding assessment information (Wertsch, 1997). Failing to grasp any nuances in feedback provision can become detrimental when students misunderstand their teacher's communication.

Student language proficiency has been found to be strongly related to learning success (Oortwijn, Boekaerts, & Vedder, 2008), to student approaches to learning (Biggs, 1990), and academic achievement (Collier, 1989; Van der Slik, Driessen, & De Bot, 2006). AfL, with its provision of formative feedback, is a linguistic endeavor. The importance of language proficiency in an academic context has been stressed as to apply to all students regardless of their cultural backgrounds. In the current study, we take up the important role of language for learning, to hypothesize that students, low on language proficiency will less appreciate AfL information offered to them. .

The current study

The current study investigates the alignment of students' and teachers' perceptions on AfL as practiced within their classroom. Our research question is: do students and teachers differ in their perceptions of AfL practice, and if so, can those differences be explained by high teacher efficacy and low student language proficiency?

3.2 Method

Sample

Thirty-one junior vocational high schools in the Netherlands were randomly approached, of which seven agreed to participate in this study. Students and teachers of these schools took part in this study. Questionnaires in this study were administered to 650 students ($N_{girls} = 323$, $N_{boys} = 326$; 1 missing) and 38 teachers. The participating teachers represented a broad domain of subjects. The average class size was 17.6 students ($SD = 4.88$; min = 11; max = 27). Teachers (20 females, 18 males) were on average 44.5 years old ($SD = 11.21$). and active for 17.8 years ($SD = 12.37$) (at their current school for 10.0 years ($SD = 10.32$). Female teachers ($M = 42.5$, $SD = 9.68$) and male teachers ($M = 46.6$, $SD = 12.62$) were about the same age, ($t(36) = -1.13$, $p = .264$). Students were on average 13.9 years old ($SD = 1.13$); girls ($M = 13.7$, $SD = 1.10$) were slightly younger than boys ($M = 14.1$, $SD = 1.13$), $t(643) = -4.17$, $p < .001$.
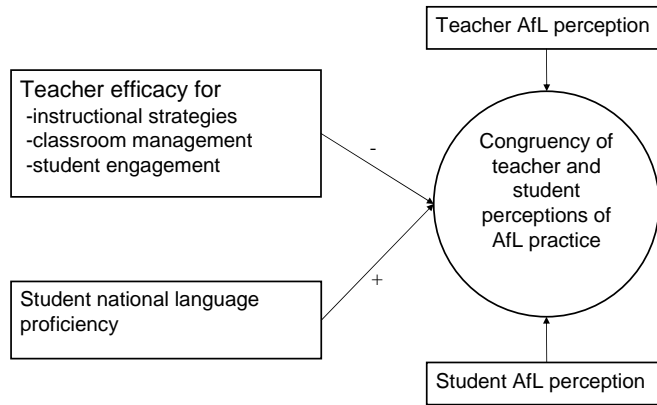
Figure 3.1: Theoretical model

Measures

Questionnaires were used to measure AfL perceptions in teachers and students in conjunction with teachers' efficacy for teaching, teachers' and students' self-esteem, students' language proficiency, as well as demographic variables such as age, and gender.

Dependent variable

Perception of AfL.

Perceptions of AfL practices were measured with the Assessment for Learning Questionnaires which consists of both a teacher and a student version (Pat-El, Tillema, et al., 2011). This questionnaire was validated to compare teacher and student perceptions of classroom AfL practice. The questionnaire consists of two subscales: Monitoring, which is comprised of items about feedback giving, and Scaffolding, which is comprised of items about clarifying learning goals and criteria. All items were scored on 5-point Likert scale items. Differences (i.e., alignment) were computed by subtracting the teachers' TAFLQ mean subscale scores from the student scores from the SAFL-Q. Negative difference scores therefore indicate higher teacher scores, while positive difference scores indicate higher student scores. Items in both the TAFL-Q and the SAFL-Q are worded similarly.The Monitoring subscale consisted of 12 items (Cronbach's α = .91). Sample items are "I encourage my students to reflect upon how they can improve their assignments", and "I discuss with my students how to utilize

their strengths to improve on their assignments". The Scaffolding subscale consisted of 16 items (Cronbach's $\alpha = .87$). Sample items are "I ensure that my students know what they can learn from their assignments", and "I adjust my instruction whenever I notice that my students do not understand a topic".

Independent variables

Teacher efficacy.

Teachers rated their efficacy for teaching on the Ohio State Teacher Efficacy Scale (OSTES) (Tschannen-Morann & Hoy, 1998). The 12-item OSTES measures teacher efficacy in three subscales: Efficacy for instructional strategies (EIS), efficacy for classroom management (ECM), and efficacy for student engagement (ESE). Each subscale was comprised of four self-report Likert items. To avoid different measurement scales in the questionnaire, the measures were adapted to five-point scales ranging from 1 (strongly disagree) to 5 (strongly agree). Sample items are "To what extent can you provide an alternative explanation or example when students are confused" (EIS); "How much can you do to control disruptive behavior in the classroom" (ECM); and "How much can you do to get students to believe they can do well in schoolwork" (ESE). Cronbach's αs in the present study were .77, .88 and .73 respectively. α for the total self-efficacy score was .85.

Language proficiency.

Student language proficiency was measured with a self-report scale from the ICSEY study (Berry, Phinney, Sam, & Vedder, 2006). This four-item self-report scale to determine language proficiency showed high correlations with evaluations of a person's language proficiency by others (Kirchmeyer, 1993). In the four-item scale students were asked to evaluate how well they were able to read, write, speak and understand the Dutch language. Scores ranged from 1 (not at all) to 5 (very well). Cronbach's α in the present study was .82.

Control measurement

Since male and female students and teachers generally differ in their evaluation of the learning environment (Van Grinsven & Tillema, 2006) this study controlled for student and teacher gender. Since Self-esteem of teachers and students is strongly related to teacher efficacy beliefs (Judge & Bono, 2001), and with students' evaluations of feedback (Nicol & MacFarlane-Dick, 2006), student and teacher self-esteem will be used as control variables as well.

Students and teachers rated their self-esteem with the Single Item Self-Esteem measure (Robins, Hendin, & Trzesniewski, 2001), The item is "I have high self-esteem" scored on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree).

Procedure

Schools were approached by email and telephone. Teachers and their students participated by informed consent. Students completed the questionnaires during regular school hours. Filling out the questionnaires took teachers and stu-

dents about 25 minutes. Respondents were assured that their contribution was anonymous. Students received a candy, as small reward when they returned their fully completed questionnaire.

Analysis

The design used was cross-sectional and correlational. Multilevel modelling (in MLwin version 2.02) was used to deal with the hierarchical nature of the data. In this way classroom effects can be controlled for when analyzing student effects, and classroom level variables can be introduced to the models (Kreft & De Leeuw, 1998). Differences in perceptions between students and teachers were first calculated into a difference score. This difference score was used as an outcome variable in a multilevel regression model. In the null-model, where only the constant was used in the equation, this procedure allows for the testing of mean differences between teachers and their specific students. Multilevel backwards selection was used to test both predictors and interaction between predictors in explaining the AfL perception alignment between teachers and students.

3.3 Results

A descriptive summary of the variables used in this study is provided in Tables 3.1 and 3.2. Student and teacher means and standard deviation on the perception of AfL practice indicated that teacher means are substantially higher than student means, with little variation within groups. Teachers evaluated themselves as highly self-efficacious.

Do teacher and student perceptions of AfL differ?

Multilevel modeling was used to determine matching in perceptions of AfL. The null-model, where only the constant was entered in the model, was used to evaluate the mean of perception differences. Teachers ($M = 3.8$; $SD = 0.61$) reported more use of monitoring than their students ($M = 2.8$; $SD = 0.76$), $\beta = -0.93$ (0.11), $p < .001$) and more use of scaffolding ($M = 4.2$; $SD = 0.49$) than their students ($M = 3.5$; $SD = 0.81$), $\beta = -0.70$ (0.11), $p < .001$). The intra-class correlation for differences in Monitoring was .51 and for differences in Scaffolding .48. The high intra-class correlations showed that average differences between students and teachers can vary greatly from teacher to teacher, and that alignment in perceptions is highly homogeneous within classrooms. This indicates that there is a strong relationship between AfL perceptions in a classroom, instead of it being a mostly individual mismatch. There was no indication of a random slopes model.

Predicting AfL perception differences

Predictors of AfL perception differences were tested with a multilevel backwards selection method (Twisk, 2006). The starting model has all variables and all necessary random intercepts and slopes included. In each subsequent step the predictor with the lowest p-value is deleted until only predictors with

Table 3.1: Means and standard deviations of the variables in the present study.

Variable	N	Mean	SD
Student level variables			
Student age	650	13.9	1.13
Monitoring	650	2.8	0.76
Scaffolding	650	3.5	0.81
Student self-esteem	643	3.5	1.17
Dutch Language proficiency	643	4.5	0.65
Teacher level variables			
Monitoring	38	3.8	0.61
Scaffolding	38	4.2	0.49
Teacher self-esteem	37	4.1	0.85
Years teaching	38	17.8	11.65
Years teaching in current school	37	9.8	9.71
Teacher age	38	44.5	11.21
Efficacy for instructional strategies	38	4.2	0.54
Efficacy for classroom management	38	4.2	0.70
Efficacy for student engagement	38	3.6	0.63
Perception differences			
Monitoring Student - Teacher	648	-1.0	.89
Scaffolding Student - Teacher	638	-0.7	.96

$p < .10$ remain. In Table 3.3 and Table 3.4 a summary of the starting model and the final model are given.

Monitoring.

After first entering control variables and predictors, backwards selections yielded the following results: low degree of alignment (large student- teacher difference on AfL per classroom) on Monitoring was significantly predicted by Student Language Proficiency, Efficacy for Instructional Strategies (EIS) and teacher gender. Only student self-esteem and student gender showed significant changes in deviance (Δ) when random slopes were removed, indicating that the average classroom-level differences of these two variables should remain in the model. (Table 3.3)

The final model indicates that higher Student language proficiency ($\beta = 0.07$, $p = .04$) corresponds to a slightly closer alignment between teachers' and students' perceptions of Monitoring, and higher teacher efficacy in instructional strategies corresponds to a wider gap ($\beta = -0.43$, $p = .005$). The relationship between EIS and monitoring alignment was inspected for curvilinearity, but the relationship was found to be linear. The apparent mismatch was found to be larger for female teachers than for males ($\beta = -0.41$, $p = .006$). Adding the interaction between teacher gender and efficacy for instructional strategies did not yield a significant improvement in fit ($\beta = -0.15$, $p = .31$, Δdeviance = -.29, $p = .65$). Efficacy for student engagement was a marginally significant predictor for mismatches in monitoring perceptions. Efficacy for classroom

Table 3.2: Correlation matrix of the variables in the present study.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	Student age	1.00														
2	Student-Monitoring	0.14	1.00													
3	Student-Scaffolding	0.12	0.70	1.00												
4	Student self-esteem	-0.02	0.10	0.13	1.00											
5	Dutch Language proficiency	-0.02	0.07	0.14	-0.06	1.00										
6	Teacher-Monitor	0.24	0.16	0.05	-0.02	-0.06	1.00									
7	Teacher-Scaffold	0.23	0.04	-0.02	-0.01	-0.11	0.73	1.00								
8	Teacher self-esteem	-0.09	-0.02	0.03	0.00	0.00	0.10	0.17	1.00							
9	Years teaching	0.39	0.01	-0.05	0.06	-0.02	0.22	0.11	0.08	1.00						
10	Years teaching in current school	0.29	0.00	0.00	0.04	0.05	0.24	0.20	0.04	0.58	1.00					
11	Teacher age	0.13	0.01	-0.03	0.07	0.01	0.01	-0.08	0.31	0.85	0.61	1.00				
12	Efficacy for instructional strategies	0.17	0.17	0.11	0.00	-0.05	0.62	0.53	0.18	0.03	0.03	-0.06	1.00			
13	Efficacy for classroom management	0.25	0.14	0.16	-0.01	-0.12	0.46	0.49	0.43	0.11	0.06	0.04	0.40	1.00		
14	Efficacy for student engagement	0.21	0.12	0.08	-0.05	-0.11	0.60	0.58	0.10	0.03	0.23	-0.08	0.58	0.37	1.00	
15	Monitoring Student - Teacher	-0.04	0.75	0.56	0.09	0.10	-0.53	-0.46	-0.13	-0.13	-0.16	0.01	-0.27	-0.19	-0.30	1.00
16	Scaffolding Student - Teacher	-0.02	0.56	0.85	0.11	0.17	-0.35	-0.55	-0.03	-0.10	-0.11	0.02	-0.20	-0.13	-0.24	0.72

management (ECM) was not found to be a significant predictor of differing perceptions on Monitoring .

Scaffolding.

After first entering all variables, both control and predictors, backwards selections yielded the following results: A high degree of difference was significantly predicted by Student self-esteem, Language proficiency, and Teacher Efficacy for instructional strategies (Table 3.4). Only student self-esteem and student gender showed significant changes in deviance when random slopes were removed.

The results indicate that high language proficiency scores correspond with a smaller difference between teachers' and students' perceptions of scaffolding ($\beta = 0.13$, $p < .0001$), while on the other hand teachers' high in efficacy in instructional strategies correspond with larger discrepancy scores ($\beta = $ -0.438, $p < .0001$). The relationship between EIS and scaffolding alignment was inspected for curvilinearity, but the relationship was found to be linear. High student self-esteem coincided with a slight reduction in discrepancy (i.e., reducing mismatch) ($\beta = 0.07$, $p = .006$). Efficacy for student engagement and efficacy for classroom management were not significant predictors of difference in perception on Scaffolding.

Table 3.3: Backwards selection of predictors for the difference scores of Monitoring (student-teacher).

	Null β(SE)	Start β(SE)	final β(SE)
Intercept	-0.93^{a}*(0.11)	1.34^{b}*(0.73)	1.47^{b}*(0.69)
Student Level			
Student gender (boy = 0)		0.08(0.09)	
Interaction Gender student*teacher		-0.14(0.13)	
Student self-esteem*		0.04(0.03)	0.05^{c}*(0.03)
Dutch Language proficiency		0.06(0.04)	0.07^{b}(0.04)
Teacher level			
Teacher gender (male = 0)		-0.42(0.18)	-0.41^{a}(0.16)
Teacher self-esteem		0.01(0.11)	
Efficacy for instructional strategies		-0.50(0.20)	-0.43^{b}(0.19)
Efficacy for classroom management		0.09(0.15)	
Efficacy for student engagement		-0.22(0.17)	-0.25^{c}(0.16)
		u1	0.15
		u2	<0.01 0.01
		e	0.36
deviance	1295.93	1190.96^{a}	1247.18^{a}

Note: $^{a} = p < .01$; $^{b} = p < .05$; $^{c} = p < .10$; * = random intercept

3.4 Discussion

The aim of this study was to examine the congruency of student and teacher perceptions of AfL practices and to examine whether possible teacher-student

Table 3.4: Backwards selection of predictors for the difference scores of Scaffolding (student-teacher).

	Null β(SE)	Start β(SE)	final βSE
Intercept*	-0.70[a]*(0.11)	0.26*(0.89)	0.36*(0.78)
Student Level			
Student gender (boy = 0)*		-0.13(0.08)	-0.13[c]*(0.08)
Interaction Gender student*teacher		0.01(0.12)	
Student self-esteem*		0.07(0.02)	0.07[a]*(0.03)
Dutch Language proficiency		0.14(0.04)	0.13[a](0.04)
Teacher level			
Teacher gender (male = 0)		-0.37(0.23)	
Teacher self-esteem		0.06(0.14)	
Efficacy for instructional strategies		-0.27(0.24)	-0.44[b](0.18)
Efficacy for classroom management		0.06(0.18)	
Efficacy for student engagement		-0.26(0.20)	
	u1	0.18	0.20
	u2	0.02 0.11	0.03 0.10
	u3	<0.01 -0.01 0.01	0.01 -0.01 0.01
	e	0.39	0.36
deviance	1395.70	1261.57[a]	1316.85[b]

Note: [a] = $p < .01$; [b] = $p < .05$; [c] = $p < .10$; * = random intercept

differences could be attributed to teacher efficacy and student language proficiency. Mismatching perceptions between students and teachers and its predictors have been rarely quantitatively researched but are of importance to understand the impact Assessment for Learning has on promoting student learning (James & Pedder, 2006; McMillan, 2007). Based on the theory of self-verification (Swann Jr. et al., 2007) it was expected that efficacious teachers focus on information that confirm their self-efficacious beliefs regarding instructional and management routes taken in instruction, which would predict a high end perception of their AfL practice in comparison to what their students perceive. On the student level, language proficiency was seen as a predicting factor in appreciating and understanding the supportive function of feedback.

This study's findings confirmed differences in student and teacher perceptions of AfL-practice within classrooms. Teachers in general perceived a higher level of AfL practice as present in their classrooms than their students did. The incongruencies found on Monitoring and Scaffolding varied greatly between classrooms and is highly homogeneous within classrooms, which suggests that there is a strong relation between a teacher's teaching and AfL practice on the one hand and student perception of it on the other. It is therefore likely that teaching-related variables could explain disagreements between students and teachers. The mismatch found in our study supports findings in earlier studies (Könings, 2007; MacLellan, 2001) in which teachers were found to overestimate their practice of AfL. Further study into mismatching perceptions on the practice of AfL between teachers and their students can help to detect why some instructional activities to foster learning sometimes have unexpected or contradictory effects (Olkinuora & Salonen, 1992).

In our study, teacher efficacy for instructional strategies and students' language proficiency predicted differences in AfL perceptions: high teacher efficacy was associated with more incongruency in AfL. In line with the theory of self-verification (Swann Jr. et al., 2007), efficacious beliefs about instructional design, and to a lesser degree about student engagement, predict a larger incongruency between teachers' and students' perception of AfL as it is practiced. The relationship we found between efficacy and AfL supports the argument that teachers who are confident in their ability to teach might focus more on the feedback they give, and subsequently may overestimate their students' evaluations of AfL.

Student language proficiency also was found to predict AfL perception differences. That is: students' higher levels of self-reported language proficiency correspond to more congruency. The relationship between language proficiency and AfL incongruency was stronger for Scaffolding than for Monitoring. These findings suggest that scaffolding is more language dependent than monitoring. For male students the mismatch was marginally bigger than for female students. An explanation may be that Scaffolding in AfL depends on understanding and following recommendations from feedback. Studies are available in which boys tend to appraise teaching and processes of evaluations more negatively (Worthington, 2002), and that girls are more language proficient than boys (Boyle, 1987). A more substantive explanation is that high instructional efficacious (i.e., task oriented) teachers initiate more negative interactions with male students than with female students (Jones & Dindia, 2004).

We found gender related differences also to be present for the teachers participating in this study. A greater mismatch in perceived monitoring was found especially for female teachers. The analysis suggests that this effect is not due to an interaction of teacher gender with efficacy for instructional strategies. A meta-analysis by Cornelius-White (2007) might help explain this finding. This meta-analysis suggests that female teachers are more likely than male teachers to emphasize the social-emotional aspects of classroom instruction (and assessment for that matter). Although it remains conjecture, because scaffolding is more associated with social-emotional aspects of teaching, through classroom dialogue and sharing understanding of personal goals to realize growth (Järvelä, Lehtinen, & Salonen, 2000), it could be that female teachers are more able to correctly recognize socio-emotional signals, but have more difficulties in task-related areas corresponding with monitoring activities.

The exploratory nature of the study was reflected by its cross-sectional correlational design. Theoretically we have support for the hypothesis that teacher efficacy, through a process of self-verification, can be an explanation for misalignments between students and teachers. This study's main findings have implications for classroom practices and research in AfL. Teacher efficacy seems a double edged sword; it can be a positive factor in valuing a mastery goal structure (Wolters & Daugherty, 2007), but also work out negatively by blocking variation and by directing teachers' focus away from (implicit) student learning signals.