THE OLD QUESTIONS ARE THE BEST: STRIVING AGAINST INVALIDITY IN QUALITATIVE RESEARCH

Betina da Silva Lopes, Helena Pedrosa-de-Jesus and Mike Watts

ABSTRACT

This chapter enters an old debate on the shape of validation processes in qualitative research. We discuss a reflective research validation framework related to teaching approaches and practices. The majority of investigations in this area draw mainly on indirect observation, semi-structured interviews or the application of questionnaires and inventories. To this extent, only “half-the-story” has been reported. The validation framework here develops a five-part three stage structure, conceptualized as an “iterative-interactive-process,” integrating a set of strategies aimed at the “minimization of invalidity.” The application of the framework is illustrated through a longitudinal study investigating the relationship between classroom questioning practices and teachers’ preferential teaching approaches. Fieldwork in this naturalistic-interpretative research was...
conducted during four academic years and entailed close collaboration
with a group of four university teachers lecturing biology to
undergraduates.

INTRODUCTION

The validity of qualitative research has been the subject of intense debate
over many decades. In earlier times conceptions of validity saw it as a sin-
gle issue, for example: Is a particular form of assessment, or analysis of
data, valid or not? More recently, definitions of validity have been pre-
sented as much more complex (see, e.g., Cunningham, 1998; Lincoln &
Guba, 1985; Miles & Huberman, 1994); however, few versions that relate
to qualitative research have gained significant purchase: there is still a mor-
ass of opinion. Our own approach has been to use “validation processes,”
a series of strategies, continuous procedures, rather than a single “test of
validity.” We envisage this as processes of “minimizing research invalidity,”
thereby, in our view, maximizing research quality. Thus, there are degrees
of validity rather than a claim to determine the work as either entirely valid
or not (Polkinghorne, 2007). The validation framework we use here has
been built from earlier work (Selvaruby, O’Sullivan, & Watts, 2008; Weir,
2005a, b), and has a five-part three-stage structure informing the work we
have conducted (e.g., da Silva Lopes, 2013; Kanhadilok & Watts, 2014;
Pedrosa-de-Jesus, Moreira, da Silva Lopes, & Watts, 2014).

In order to discuss the genesis and application of our “validation-as-
iterative-interactive-process” framework, this chapter begins by exploring
briefly different notions of validity, old and new. We also review research
into the conceptualizations and practices of university teachers, which is
the focus of the current (naturalistic, longitudinal) research case, and
through which we can present and illustrate these processes of validating.
The summary discusses key aspects we have presented and makes the case
for further research and debate.

VALIDITY IN QUALITATIVE RESEARCH

University academics across Europe (indeed the world) are exhorted to
display honesty, openness, reliability, and rigor in their work (European
Science Foundation (ESF, 2011). Some (e.g., Wood, 2012) have pointed to the under-determination of such descriptors and round on “rigor” as an inappropriate metaphor. While it is often used to mean “consequential,”

Etymologically speaking, it means “stiff,” as in the term “rigor mortis,” which describes the condition of someone who has been dead for more than three hours. (Wood, 2012, p. 3)

Wood continues to challenge the notion of rigor by pointing out that, in the current use of the term, it is seldom made clear if it applies to the methods used or the outcomes that are delivered. Our sense of rigor entails the kinds of openness and accessibility, fairness, responsibility, and duty of care urged by the ESF (2011): a state where as many of the parameters of a given research activity are made explicit and can, therefore, be seen in relation to one another. For us, a sense of “validation integrity” corresponds with a culture of academic “fittingness,” in which research is designed with as much consistency, congruence (Watts & Bentley, 1986), and transparency as possible.

The debate is old, dense and sometimes confusing, not simply because of the complexity of examining rigorous quality in qualitative research, but due also to the diversity of terms in use, sometimes divergent, others ambiguous or overlapping. Moreover, setting standards for validity in qualitative research is challenging because of the need to incorporate subjectivity and creativity as well as “quality” within the process (Gray, 2004; Johnson, 1999). Over time, there have been several responses to these needs: the first a form of denial, where issues of validity are merely ignored because they are seen to be associated with a quantitative approach and therefore impossible to achieve within qualitative inquiry (Lincoln & Guba, 1985). On the other hand, some argue that defining criteria for rigor/validity is indeed important in order to combat the reputation of qualitative researchers as “second class” (Field & Morse, 1985; Silverman, 2006). Inside this broad continuum we find two key divergent opinions. Some authors “borrow” the positivistic concepts (such as validity, fidelity, replicability, and generalizability) and work to adapt these. In these instances we also find researchers who make reference to “truth” (Silverman, 2006). Others rename the problem. Lincoln and Guba (1985), for example, generated labels they considered more appropriate in qualitative studies than traditional methods of validation: “Trustworthiness” rather than validity, which refers to the quality of an investigation as judged by three criteria: credibility, transferability, and dependability. Miles and Huberman (1994, p. 110) used plausibility, sturdiness, and
confirmability. Beyond these come various writers with pragmatic solutions, not least through “triangulation,” commonly positioned as a panacea for all such research issues (Blaikie, 1991; Hammersley, 2008).

Our position is that more important than entering into renaming is to give sound conceptual instruments (models) that help qualitative researchers to minimize vulnerability and maximize confidence in their research. For us, any quest for consensus on a singular, absolute validity is best replaced by the development of multiple “defensible knowledge claims” (Kvale, 1996) or a transparent “decision chain” (Long & Johnson, 2000). As Sireci (2013) notes, “validity is not a property of a test, but rather it refers to the use of a test for a particular purpose” (p. 99). It is the researcher’s responsibility to theorize and evaluate the conceptual framework of the work, to make continuous “corrective” checks on data for credibility and plausibility, to test for false statements, analyze sources for potential biases, to question and ask what, why, and how? As Schutt (2013) points out, it is the researcher’s responsibility to maintain professional integrity, accuracy, and probity in working to elucidate their research. The purpose of validation in this sense is not to provide a single answer (“Yes, this is valid”), so much as demonstrate resilient attempts to minimize invalidity. It refers to the degree to which evidence exists to support or not explicit uses and, to this extent, we hark back to Lather (1986) who suggests that the role of validation is less to support and confirm interpretations than to explore what might be wrong with them.

The Five-Part Three-Stage Validation Framework

Ours, then, is a broad definition of validation (Kane, 2006). The framework we discuss here is originally based in Weir’s (2005a, b) language work. Weir proposed a socio-cognitive validation framework for language testing that, he argues, can form the basis of any test development and validation project. To examine the validity of a test, he says, requires both explicit theory and technique to guide the validation approach, and thus a validation framework to operationalize validity in its various manifestations (Weir, 2005a, b, p. 39). We have used our adaptation of his work (Selvaruby et al., 2008), for example, for the validation of national 16+ testing and the use of school-based assessment. It is an evidence-based framework for validity, developed to enhance transparent dialogue and debate about the outcomes that are generated before, during, and after the research.
The model itself is intended as a framework of conceptual tools to help shape understanding and actions. It integrates five processes of validation, namely context validation, theory-based validation, response validation, criterion-related validation, and consequential validation. Fig. 1 illustrates the arrangement of each aspect in the timeline of conducting a research project and also the mutual influences between them.

Context Validation

This process is concerned with the extent to which the broad remit of the research, the initial choice of direction, and data-gathering tasks are germane to the anticipated audiences. In a speech two decades ago to the (then) UK’s Teacher Training Agency, Hargreaves (1996) described educational research as “irrelevant to practice … uncoordinated with any preceding or follow up research … virtually nobody reads.” Much more recently, Bennett (2013) said:

I am astounded by the amount of [education] research I come across that is either (a) demonstrably untrue or (b) patently obvious… If that sounds like a sad indictment of educational research, it is. (p57)

As MacIntyre (1998) has pointed out, educational research must work hard to be relevant, not to become overly narrow and specialized, and have contextual validity. So, are the features of the research task(s), its
questions, structures and administration appropriate to the sample of participants and respondents involved? Do the research questions clearly grow out of the relevant context? Is the purpose of the research of value, of interest? Are the data sets being proposed appropriate? It is not uncommon for research to be dismissed because the initial context was ill-conceived — the research is fatally wounded even before it begins. There is no doubt that the research needs to be well-bedded: stakeholders, respondents, and “end-users” in the research can be asked to comment on these issues before the research fully begins.

Theory-Based Validation

This process is concerned with how the theoretical framework of the research informs the fit with the data-gathering methods. Are the processes involved in the research congruent with its overlying philosophy and underlying principles? Here we place importance on what the researchers and respondents will actually be doing: how does the performance of the respondents, the ways in which the data is actually gathered, relate to the broad or specific theoretical models used in the research? Besides methodological theory, this aspect of validation also entails theories about the research object itself. Grounded theory (Glaser & Strauss, 1967) may be a way forward, based as it is in the “reverse” of usual approaches to theory-driven research, but this is notoriously difficult to do in any pure form (Thomas & James, 2006).

Considering these first two, context and theory validation, it becomes clear that validity of the research project is an issue even before the research project has fully started. In the same way as within quantitative studies, qualitative studies should entail strategies of theory selection and participant selection (Yin, 1993). Careful and detailed planning through the development of a “logic train” is essential: what is the logical chain between the research question, the context, the theory, and the conceptualization of the research? Can the congruence be seen between theory and data collection? As answers to such questions arrive, so the “logic train” will be adjusted throughout the research. Maintaining a dialogue with the “field” is important, compiling a research diary to register doubts and decisions will help further steps in the investigation.

Response Validation

This relates to the means of gathering data, responses by the respondents, and the interpretation of these by the researcher. The researcher needs to
account for the ever-changing context within which research occurs; it is important to describe the changes that do occur in the setting and how these then affect the way the responses are generated. So, how far can one depend on the scores or performances in the research method? To what extent do the interview questions, questionnaire items, tests, and observation schedules (etc.) achieve what they set out to do? How can these be related to the “categories of response” derived by the researcher? How are “untidy” or inappropriate responses to be dealt with? This is the more usual version of “content validity,” a systematic approach to validation criteria but further expanded to cover the coding and categorization by the researcher. So, the sense here is of questioning how accurately the accounts represent respondents’ realities of social phenomena: the testing of validity relates not to the data but the inferences drawn from them.

There are a number of strategies for enabling response validity as described here. The researcher can document the procedures for checking and rechecking the data throughout the study. Another colleague can take the role of “critical friend” with respect to the outcomes of analysis of the responses, and this process can be documented. The researcher can actively search for and describe negative instances that contradict previous observations. After the study, one can conduct a data audit that examines the data collection and analysis procedures, and makes judgments about the potential for bias or distortion. In order to verify construct validity and coding fidelity it might be important to proceed with “peer-debriefing,” “member checks,” and “inter-judge agreements” (Selvaruby et al., 2008). Again, maintaining a “decision trail” during the process of data gathering might be relevant for future decisions.

Criterion-Related Validation

This process considers the relationship of the outcomes of the research to other evidence in the field, to the interpretative frame of reference. What external evidence is there that, outside of the categories of response themselves, the outcomes of the research are appropriate? To what extent can this (possibly relatively untested) approach compare with another for which the validity has been well established? Are the interpretations consistent with other research that measures the same (or similar) constructs? Studying the literature of the area under study enables the outcomes to be embedded within extant research, about what the findings mean, and enables the researcher to be sensitized to broader, developed concepts.
Consequential Validation

This process looks at the broad effects and impacts the research outcomes have on its various stakeholders. This is about how research adversely affects or benefits the situation of the research, the extent to which research has the potential to play an effective role in some form of educational change (Massey & Barreras, 2013). For example, teachers might have a beneficial “backwash validity” if the students’ perceptions of their teaching are clearly pointed out; the teachers would be more successful because they will be focused on what is being demanded. That would be for both, students and teacher; they must know what the test asks in order to be prepared for this. So, how well do respondents recognize or identify with the outcomes of the data they produced? What is the effect on learners, teachers, and others in the frame of reference? Responses to these questions should not be forgotten, should be addressed and integrated into the research results.

We have been making the case that validating knowledge claims is not a mechanical process but, instead, is an argumentative practice. The purpose of these processes is to convince audiences of the likelihood that support for the claims being made is sufficiently strong that the claim can serve as a basis for understanding and action. Considering these last two processes, criterion, and consequential validation, it becomes clear that validity of the research project continues to be an issue even after the main part of the research project has finished. Careful and detailed planning through the development of “critical friends,” respondent feedback, inter-judge comment is essential: what is the logical chain between the research question, the context, the theory, and the conceptualization of the research? Maintaining the research diary to register doubts and decisions will be helpful in reporting the investigation.

AN EXEMPLAR STUDY: ACADEMICS’ TEACHING CONCEPTIONS AND PRACTICES

In order to discuss an application of this validation framework, we describe a recent longitudinal study conducted in the context of university biology teaching. In the last decades, particularly since 1980, research considering the investigation of academics’ conceptualizations of their teaching and teaching practices have grown substantially (Kane, Sandretto, & Heath,
The emerging research has been justified in the context of university transformation toward innovation. It can also be seen as a natural consequence of the efforts to extend knowledge already constructed through research on: (i) teaching styles and approaches involving primary and secondary teachers (Pajares, 1992) and (ii) the relationship between learning conceptions, learning styles or approaches, and learning outcomes of university students (Biggs, 1999; Ramsden, 1992).

In this sense three major research areas can be identified: (i) the study of preferential teaching approaches, a construct developed by Prosser and Trigwell (Trigwell, Prosser, & Taylor, 1994); (ii) the study of orientations to teaching, initiated by the work of Kember (Kember & Gow, 1994), and finally (iii) the research on lecturing styles (Brown & Pluske, 2007; Felder & Spurlin, 2005; Heimlich & Norland, 2002). We develop a more detailed literature review in Pedrosa-de-Jesus and da Silva Lopes (2011). Here we are interested in highlighting the main research convergences and divergences that are summarized in Tables 1 and 2.

It is principally the research gaps indicated in Table 2 that lead to the interrogations of Devlin (2006) and Eley (2006). Both these authors explicitly critique the assumptions that teachers act in ways, when questioned, that they say they do, assuming that the conceptions of teaching described during interviews are merely post-hoc reflections and have no functional role, and do not necessarily influence everyday teaching routines.

In our research, four Portuguese university teachers lecturing in biology to undergraduates, were observed during two consecutive academic years. The research was directed at investigating how they commonly use

**Table 1.** Literature Review on Teaching Conceptions and Practices of University Teachers (Main Convergences).

<table>
<thead>
<tr>
<th>Key-Aspect</th>
</tr>
</thead>
<tbody>
<tr>
<td>• University teachers have different forms of “thinking” and “doing” teaching. The same content might be taught in very different forms (Prosser et al., 2005).</td>
</tr>
<tr>
<td>• There are several conceptual models that aim to sustain the interpretation, comprehension of these diverse ways of thinking and doing teaching. All models, in essence assume two broad distinct modes: some teachers are more focused on the content while other teachers are more focused on developing learning processes (Devlin, 2006; Postareff et al., 2008).</td>
</tr>
<tr>
<td>• Empirical evidence indicates that it is possible to differentiate teachers by considering their ways of thinking and doing teaching, based on the analysis of their responses to specific questions (questionnaires, inventories, and interviews) (Brown &amp; Bakhtiar, 1988; Kane et al., 2002; Trigwell &amp; Prosser, 2004).</td>
</tr>
</tbody>
</table>
questions during lectures (both students’ and their own), and how they managed to implement some student-centered strategies suggested by a group of science education researchers. Data was gathered by participant and non-participant observation of the teachers’ professional activity during their lectures, through semi-structured interviews and also the application of a translated and validated version of the revised Approaches to Teaching Inventory (Trigwell, Prosser, & Ginns, 2005).

The main research aim was to contribute to a deeper understanding of teachers’ use of questions in their classroom and how they promote questioning throughout didactical interactions. In this sense it was important to describe teachers’ conceptions, motivations related to teaching, to questioning, and also to describe their adopted practices. Understanding the connection between teaching conceptions, teaching intentions, and teaching practices is crucial to the design and implementation of successful strategies envisaging quality teaching and, consequently, learning at higher education (Weimer, 1997). For more detailed descriptions please see: Selvaruby et al. (2008), Pedrosa-de-Jesus, da Silva Lopes, and Watts (2008), Pedrosa-de-Jesus, Watts, and da Silva Lopes (2009), Pedrosa-de-Jesus and da Silva Lopes (2011), Pedrosa-de-Jesus and da Silva Lopes (2012), and Pedrosa-de-Jesus, da Silva Lopes, Moreira, and Watts (2012).

**Table 2.** Literature Review on Teaching Conceptions and Practices of University Teachers (Main Divergences/Research Gaps).

<table>
<thead>
<tr>
<th>Key-Aspect</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Some investigators consider that teaching conceptualizations are context dependent, others not. Therefore, the distinctions between conceptions and intentions of teaching and their relationship to practices of teaching are still unclear (Devlin, 2006).</td>
</tr>
<tr>
<td>• The previous aspect may be related to the assumption of many studies that espoused theories of action and theories in action are equivalent (Marton &amp; Booth, 1997; Prosser et al., 2005). Actually there are very few studies that cross indirect data obtained through questionnaires, inventories, or interviews with data gathered through direct observation.</td>
</tr>
</tbody>
</table>

**THE “VALIDITY” OF THIS RESEARCH CASE-STUDY OF UNIVERSITY TEACHING**

Our belief is that the more comprehensive the approach to validation, the less invalidity can be aimed at discrediting the overall task. This is not to
swamp the research endeavor with innumerable “truth tests,” but to ques-
tion each stage of the work. Understandably, each of these different aspects
of validation influences the others, as illustrated in Fig. 1.

Context-Based Validation

Our claim is that the overall context of this research is important, and our
validation through context comes in two parts. First, Portuguese higher
education institutions are undertaking a challenging process of innovation
through the Bologna Process (Veiga & Amaral, 2009). Within this, the pro-
cess of “transforming the pedagogy of universities” has prompted teachers’
efforts to reflect on and adapt their teaching practices (Prosser, Martin,
Trigwell, Ramsden, & Luevkenhausen, 2005). So, our research aimed to
explore these individual and institutional needs, not least in overcoming
difficulties related to the operation of the Bologna philosophy toward
daily class activity. In this sense, the research fits within an overall
national — even European-wide — context (Crosier, Purser, & Smidt,
2007), relates closely to aims and ambitions of the university itself, and to
those of the relevant departments of the university. The clear aims were dis-
cussed with the main stakeholders of the project, principally key members
of the university, the university departments, and the individual teachers,
and met with their full approval.

Second, the form of the research must be congruent with this broad con-
text. The research project should be conducted, as far as possible, in a nat-
uralistic and suitable milieu for exploring conceptions of, and approaches
to, teaching. It was intended to explore real-life conditions so that research
findings are deemed as fully contextually appropriate as possible.
Therefore, it was decided to adopt a “case-study” research design with eth-
nographic dimensions (Cohen, Manion, & Morrison, 2003; Gray, 2004;
Tuckman, 1990), implying close and long-time collaboration between tea-
chers and the researcher.

Theory-Based Validation

The theoretical components of our study also come in two parts. The first,
the investigation of teaching conceptions and practices, has already been
discussed. The extensive literature review, combined with the research aims
of our project, highlighted the identification of these teachers’ conceptions
about teaching. This led us to select the Approaches to Teaching Inventory (ATI), developed by Trigwell et al. (1994). The ATI explores the ways in which academics undertake teaching and has identified two “extreme” teaching approaches, namely ITTF – “information transmission teacher focused” and CCSF – “conceptual change student focused.” The most recent version of this instrument (Trigwell et al., 2005) has 22 sentences describing intentions (closely related to teaching conceptions) and specific teaching strategies. Teachers are asked to place themselves on a Likert-type scale from 1 to 5, and the results are based on the mean score of the numeric response for each item in both scales.

The main reasons for selecting this instrument were: (i) the inventory was, like our study, developed in the context of higher education, and that it was short and concise made it straightforward to be answered by busy university teachers; (ii) the process of developing the inventory was described broadly in the literature (Trigwell & Prosser, 2004), as well as its use by other research groups (see, e.g., Eley, 2006; Hendry, Lyon, & Henderson-Smart, 2007), enhancing our confidence in its utility. Furthermore, (iii) it includes topics on teacher-student interaction (through classroom questioning), the second dimension of research of our project. Besides that, the idea of “approaches to teaching” was particularly apposite in conducting our research since it integrates the teaching practices and teaching theories, while Kember’s (1997) construct of “teaching orientations,” for example, looks (only) at teaching conceptions. On the other hand, the majority of studies that discuss lecturing styles are mainly descriptions of the teaching strategies that academics adopt (without integrating teachers’ conceptualizations).

The second theoretical component of the research is related to questioning processes within teaching and learning contexts. Confronted within the impossibility of studying every factor that integrates the complex dynamic of the teaching-learning processes, it was decided to focus on this one specific dimension, the use of questions. These are considered to be powerful pedagogical tool boxes to promote quality learning (Chin & Osborne, 2008; Dillon, 1991; Pedrosa-de-Jesus et al., 2012), and also provide a clear theoretical direction for the research.

Response Validation

This is scrutiny of our technical competence in operating the research dimensions and conducting the research. Three illustrative aspects of the research process will be highlighted.
Translation of the Original ATI from English into Portuguese

After selecting Trigwell et al.’s instrument we asked the authors, via e-mail, for permission to translate this into Portuguese. We explained our research aims and methodology to the authors, who then agreed to the translation, in return asking for a copy of the final version of the instrument. The translation was conducted using the process of “back translation”:

1. Two independent Portuguese translations were undertaken. Translation A was made by one of us (BdSL) and translation B was made by an English-Portuguese teacher external to the research team.
2. Both translations were compared by the research group. Group discussion led to a third, improved version (translation C).
3. The third version was then back-translated by another person external to the research project.
4. The original English version and the back-translated into English version were compared. The research group agreed that the essence of the inventory was maintained.

Since we did not find any other published Portuguese version of the inventory, it was decided to pre-test the instrument. An e-mail was sent to all teachers of the university (N=890) where the project was being implemented, asking them to respond to the inventory. We obtained a response of 12% (n=102 teachers). The obtained internal consistency values were considered adequate (Cronbach Alpha above 0.75 for both dimensions — ITTF and CCSF).

Use of the Portuguese ATI

As previously described, we have worked with four academics during four consecutive academic years. Non-participant lecture observations were conducted. The four teachers responded to the inventory, two showing a preferential ITTF approach, the other two a preferential CCSF approach (see Table 3). These results were “confronted and confirmed” during an interview with each teacher. During the interview no comparisons between teachers were made, since the aim of the project was not to compare teachers’ performance or to establish rankings of “better” teachers. Each teacher recognized themselves in their descriptions. Naturally some divergent perspectives also emerged, and were analytically explored and are discussed in Pedrosa-de-Jesus and da Silva Lopes (2011).
Development and Use of a Teacher Questioning Categorization System

Since the research aim was to explore the relationship between teachers’ conceptions of teaching and the way they used questions during didactical interaction, it was necessary to categorize teachers’ questions. For this, we adopted a categorization system previously developed (Pedrosa-de-Jesus et al., 2008). The research findings were considered to be insufficient to describe the convergences and divergences of the way the four teachers used questions. It was therefore decided to develop a new categorization system. An extensive literature review led to the integration of the question categories with a form of discourse analysis (Aguiar, Mortimer, & Scott, 2010). The resulting coding system integrates two dimensions: the observed teachers’ behavior and the intention beneath that behavior. The behavior dimension is composed of three levels of analysis: (i) micro (the questions per se, frequency and cognitive level of the questions); (ii) meso (the dialogic or non-dialogic nature of the teachers’ reaction to a student intervention, and also the teachers reaction to a non-student answer), and (iii) macro (interaction extension, this is number of “moves” of each teacher-student dialogue). We describe this categorization system further elsewhere (da Silva Lopes, 2013; Pedrosa-de-Jesus & da Silva Lopes, 2011), in this chapter we highlight only the process of “validating” the system for categorizing teachers’ questioning, which was undertaken from two perspectives.

Discussion of the Question-Categorization System with a Panel of Seven Judges

A panel of seven researchers with different backgrounds and research experience (including two PhD students, two junior researchers, and three senior researchers) was constituted in order to discuss the clarity and consequent efficiency of the questioning categorization frame. For this purpose, a written document was prepared, composed of three parts: (a) brief description of the aims and methodology of the research project and some key aspects of the literature; (b) detailed description of the categorization

Table 3. Preferential Teaching Approaches of the Four Biology Undergraduate Teachers.

<table>
<thead>
<tr>
<th>Teacher</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITTF SCALE</td>
<td>4</td>
<td>4</td>
<td>3.5</td>
<td>3.7</td>
</tr>
<tr>
<td>CCSF SCALE</td>
<td>3.7</td>
<td>3.4</td>
<td>4.2</td>
<td>4.3</td>
</tr>
</tbody>
</table>
model, with two illustrative examples; (c) four dialogues to be categorized by each member of the panel. Each judge read the document and categorized the examples that were given. After this task, a discussion was undertaken about major difficulties in interpreting the categories and applying the coding system. The main mentor of the categorization system took several descriptive notes in order to improve some minor aspects. To reinforce the confidence in the model, and also to have more sustained arguments for the public when communicating this categorization system to a broader audience, it was afterward decided to gauge the level of agreement between each researcher of the panel and the mentor for both dimensions of the categorization system, namely the observed teachers’ behavior and the intention beneath that behavior. The agreement was considered satisfactory. The lowest levels were obtained with the categories that integrated the dimension “intention.” Considering that the classification of intentions is naturally more subjective than the classification of, say, “observed behaviors,” the agreement obtained was considered to be a natural and understandable consequence. Within this the confidence in the categorization system was increased.

Discussion of the Categorization System with Four Teachers Who Constitute the Research Case of the Investigation (Main Informants)

The final aspect of response validation took a slightly different approach. The model was considered adequate and innovative by our panel of judges. However, considering the detail of the analytical approach, several members of the panel emphasized that this could be a handicap for the validation by the academic teachers, since they are, naturally, not necessarily familiar with this type of coding. So there was a risk that percentage agreements might be low(er) due to “coding errors,” and eventually lead to disagreement. In line with this recommendation, and considering the research aims of the project, we decided to “validate” the model through a “task-based” interview (Koichu & Harel, 2007), where the teachers were asked to “think aloud” while coding various episodes, and express their doubts to the researchers. The interviewer (researcher) maximized efforts at maintaining a neutral position. Audio-records of the interviews were transcribed verbatim and subjected to qualitative content analysis, which is described in detail in Pedrosa-de-Jesus et al. (2012). Again, in order to facilitate discussion of the categorization system with a broader audience we decided to gauge the levels of agreement between each teacher and the mentor of the model. The lowest percentage that was obtained was 82%, and considered to be very acceptable.
Criterion-Related Validation

This process considers the relationship of the research to the interpretative frame of reference. By combining direct and indirect observation (interviews and the inventory) it was possible to confirm a strong internal relationship between teaching conceptions and the adopted teaching practices, in this case questioning, reinforcing the theoretical assumption that “teaching in action” and “theories of teaching” are complementary phenomena (Postareff, Kaajavuo, Lindblom-Yläne, & Trigwell, 2008). Indeed, the selected inventory was able to distinguish the four teachers, considering their preferential teaching approaches. It was possible to verify that the differentiation of teachers by their ATI responses goes beyond the post-reflection levels, as argued by Devlin (2006) and Eley (2006). In this specific research case, teachers identified as having different preferential teaching approaches actually do behave in different “ways” during lectures, while teachers identified as having the same approach behave similarly, at least in respects that concern questioning practices. It is in this sense that we can argue that the teachers’ questioning practices can be a useful indicator of their main teaching and learning conceptions. The relationship between these two dimensions is explored in more detail in Pedrosa-de-Jesus and da Silva Lopes (2011), as well as in da Silva Lopes (2013).

The research case we present here provides evidence that modifications to teachers’ practice (in this case teacher questioning) do not necessarily imply a preferential teaching approach change, since the identified teaching approach of each university teacher has proved to be relatively constant across time (namely, over four academic years). This may indicate that external factors can potentially induce a change in teachers’ behavior (questioning) without implying any changes in their more deep-seated conceptions of teaching and learning. Considering the stability of the approach of these university teachers, it can be argued that our findings indicate that Trigwell and co-workers’ concept of preferential teaching approaches is close to Vermunt’s (1998) concept of “orientation” and to Meyer’s (2000) concept of “orchestration,” which are considered to be allied to intrinsic motivation. For a broader discussion please see Pedrosa-de-Jesus et al. (2008).

Consequential Validation

Besides extending the preferential teaching approaches framework (Trigwell & Prosser, 2004), which has more recently also been used to
understand regulation processes of teaching (Lindblom-Ylänné, Nevgi, & Trigwell, 2011) and teacher’s emotion in teaching (Trigwell, 2012), the work was highly valued by our group of teachers and by the cohorts of undergraduate students involved. It also highlighted issues considered to be useful for the design of effective strategies of academic development (Table 4), which also has been a recent object of research within the preferential teaching approaches framework (Trigwell, Caballero Rodriguez, & Han, 2012).

**CONCLUSION**

The “validity” of qualitative research is not a settled subject. The validation framework we present here enters this old debate with a fresh approach. Through the exploration of a specific research case involving four Portuguese university teachers, we emphasize how the researcher is responsible for showing that his/her interpretations, decisions, and actions are not simply “invented” or capricious, but have been the product of conscious construction and scrutiny. Our validation arguments are intended to provide an overall evaluation of the evidence for and

---

**Table 4.** Relevance of the Project for Its Stakeholders (Academics and Educational Researchers).

<table>
<thead>
<tr>
<th>Key-Aspect</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Identification of alternative strategies for promoting teacher reflection:</strong> All teachers stressed the novelty of the “task-based” interview experience, namely reflecting through the use of “concrete” examples of their own lectures. The use of real data from the lectures of each teacher in an organized way revealed to be an efficient strategy to enhance teachers’ positive motivation toward reflection.</td>
</tr>
<tr>
<td><strong>Use of the inventory as an intervention instrument:</strong> One of the teachers suggested the use of the inventory as a way to promote group discussion in order to confront different perspectives.</td>
</tr>
<tr>
<td><strong>Highlighting the importance of diversity and flexibility toward academic development:</strong> Strategies envisaging teaching improvement through reflection have to take into account the complex relationship between theory and action. The way each teacher manages to adopt or adapt a particular strategy is influenced by his personal motivation and ultimately his conceptions. The same strategy suggested to a group of teachers might have different outputs, since the personal theory drives our motivations and intentions, and by implications influence our perception of the “outside world,” including the suggestion of particular teaching-learning-assessment strategies.</td>
</tr>
</tbody>
</table>
against the proposed interpretation/use (i.e., for and against the interpretive argument). In this way, a statement or knowledge claim is not intrinsically valid; rather, its validity is a function of the context of use and the inter-subjective judgment. A statement’s validity rests on a consensus within a community of researchers and respondents, producers and end-users. The validation process takes place in the realm of interpretative interaction, and validity judgments make use of a kind of communicative rationality.

Considering the specific research area of academics conceptualizations and practice, the recommendation taken from our validation framework is that future research efforts should consider the integration of data gathered through direct observation in order to fully understand the complex relationship between what teachers’ believe, intend to do, and actually do during classes. If this is not possible, researchers should at least explicitly acknowledge why they did not take these dimensions into account and reflect on the limitations that it might bring into their conclusions. We believe that the considerations discussed here are particularly relevant within investigations that follow a naturalistic-interpretative paradigm since the interpretation of qualitative data tends to be more exposed to criticism considering possible biases or subjectivity in comparison to quantitative data.

Finally, the presented model may constitute a useful instrument to support researchers (in particular PhD students) in their effort of “striving against invalidity.”

ACKNOWLEDGMENTS

The authors acknowledge the financial support of Portuguese Fundação para a Ciência e a Tecnologia (SFRH/BD/44611/2008; PTDC/CPE-CED/117516/2010). We also thank the four professors, and all students involved in the project.

REFERENCES


European Science Foundation (ESF)/All European Academies (ALLEA). (2011). The European Code of Conduct for Research Integrity, European Science Foundation, Strasbourg, France.


Entwistle and Walker (2000)
Dear Author,

During the preparation of your manuscript for typesetting, some questions may have arisen. These are listed below. Please check your typeset proof carefully and mark any corrections in the margin of the proof or compile them as a separate list.

**Disk use**

Sometimes we are unable to process the electronic file of your article and/or artwork. If this is the case, we have proceeded by:

- Scanning (parts of) your article
- Rekeying (parts of) your article
- Scanning the artwork

**Bibliography**

If discrepancies were noted between the literature list and the text references, the following may apply:

- The references listed below were noted in the text but appear to be missing from your literature list. Please complete the list or remove the references from the text.

- **UNCITED REFERENCES:** This section comprises references that occur in the reference list but not in the body of the text. Please position each reference in the text or delete it. Any reference not dealt with will be retained in this section.

**Queries and/or remarks**

<table>
<thead>
<tr>
<th>Location in Article</th>
<th>Query / remark</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>AU:1</td>
<td>The right running head is amended to fit the space. Please check the amendment for correctness.</td>
<td></td>
</tr>
<tr>
<td>AU:2</td>
<td>Please confirm whether the order of the section headings set is correct.</td>
<td></td>
</tr>
<tr>
<td>AU:3</td>
<td>The reference “Entwistle and Walker (2000)” is not cited in the text. Please cite the reference; if no citation is available then the reference needs to be removed from the list.</td>
<td></td>
</tr>
</tbody>
</table>