Evaluation Practices of Doctoral Examination Committees: Boundary-Work Under Pressure

Maja Elmgren[1] · Åsa Lindberg-Sand[2] · Anders Sonesson[2]

Abstract The doctorate forms the basis for academic careers and the regeneration of academia, and has increasingly become important for other sectors of society. The latter is reflected in efforts on institutional, national as well as supranational levels to change and adapt the doctoral degree to new expectations. As doctoral education is embedded in research, changes in governance and funding of research further affect the doctorate. The evaluation of the doctoral thesis appears, however, to have remained true to the academic tradition: an examination committee exercising their gatekeeping in a ceremonial setting. This study sets out to explore doctoral examination committees' evaluation practices. Insights were gained through six focus group interviews with experienced examination committee members at three large research-intensive universities in Sweden. Of particular interest is how the object of evaluation is formed, the nature of the boundary-work conducted, and variations in examination practices related to different and changing conditions for research and doctoral education. Our results show how the object of evaluation emerges through a gradual interpretation of the thesis and defence, becoming more complex and nuanced as the process of evaluation progresses from its initial stages to the final closed discussions of the committee. The finalised object of evaluation, only fully present at the conclusion of the closed meeting and hence transient in nature, encompasses the research contribution, educational achievement, and academic competence of the candidate. Furthermore, the boundary-work conducted in this process often transcends the object of evaluation to include also supervision and the local context for doctoral education and research, and hence contributes to upholding, and potential changing, norms in research fields, educational contexts,

✉ Anders Sonesson
anders.sonesson@ahu.lu.se

1   Ångström Laboratory, Department of Chemistry, Uppsala University, Uppsala, Sweden

2   Division for Higher Education Development, Department of Educational Sciences, Lund University, Lund, Sweden

 Springer

and academia at large. This extended boundary-work intensified as problems and inconsistencies were discovered during the evaluation process. The ceremonial staging underscored the gravity of the decision and the extended boundary-work. Despite changing conditions for the doctorate, our findings highlight the importance of the practice of evaluation committees, and the disciplinary communities to which they belong, for upholding and negotiating norms in academia.

Introduction

The doctoral degree dates back to medieval times as an intra-academic qualification coupled to positions, privileges, and responsibilities. Accepted across borders, the doctorate provides an early example of the global convergence of disciplines and degrees. Its long history and importance are visible in the often highly ceremonial staging of the evaluation of the doctoral thesis and the awarding of the degree. The doctoral examination committee is at the heart of this process, which can be understood as a form of boundary-work (Gieryn 1983), and their evaluation practice forms a strong link to the universities' long histories even in our present time of rapid change and reorientation of universities and their role in society. This simultaneity of tradition and change highlight a potential tension in the evaluation committees' work that our study sets out to explore; in what ways do the committees approach their task amidst changing conditions for doctoral education?

Many of these changes are consequences of a growing interest in doctoral education among policymakers, funding bodies, authorities, and national as well as supranational organisations and networks. This interest relates to the perceived strategic importance of the doctorate for national and regional development and competitiveness, as well as for dealing with the wicked problems of the world. This, in turn, has resulted in a proliferation of national rules and regulations as well as international policies and guidelines, further underpinned by global drivers such as massification, professionalisation, and quality assurance agendas (Andres et al. 2015). From once having been exclusively the concern of the professoriate, the doctorate today finds itself at an intersection – and sometimes a tug-of-war – between different policy fields, notably research, education, and innovation (Ruano-Borbalan 2022; Elmgren et al. 2016; Sonesson et al. 2023). In Europe, the doctorate was included in the Qualifications Framework for the European Higher Education Area (QF-EHEA 2005), leading to revised national regulations and degree descriptors. Other efforts, not unique to Europe, include funding schemes, structured graduate schools, formalised curricula, learning outcomes, and quality assurance mechanisms (Byrne, Jørgensen, and Loukkola 2013; Kehm 2020; Taylor, Kiley, and Holley 2021). Furthermore, new forms of middle-management in doctoral education have emerged worldwide (Byrne, Jørgensen, and Loukkola 2013; Kehm 2020; Taylor, Kiley, and Holley 2021).

Notwithstanding these changes, the assessment of doctoral candidates appears to be largely in line with tradition: an examination committee's evaluation of the thesis – and commonly also of the candidate's defence. This raises several questions: How does the practice of examination committees reflect the changing conditions for doctoral education? How do, in turn, examination committees contribute to the shaping of doctoral education? It is therefore important to investigate the inner conditions for the evaluation practice of examination committees: How do they go about making their decisions?

Purpose of the Study

The purpose of our study is to explore the evaluation practices of examining committees in doctoral education, and by that gain insights into assessment of doctoral theses and candidates, in times of change and influence from competing interests. Of particular interest is what committees are looking for and how the object of evaluation is formed in the evaluation process, as well as the nature of the boundary-work in which committee members engage. Furthermore, we are interested in variations in evaluation practices and how these can be related to different disciplinary traditions and current conditions for doctoral education as well as for research.

Roots, Commonalities, and Variations in Doctoral Examinations

The roots of doctoral examination are found in the *disputatio,* a formalised debate in the medieval universities of Europe. During such events the respondent was expected to show the ability to orally defend the doctrinal canon – to embody the truth – while the opponent's role was to point out errors and show that the respondent was wrong (Sère 2020; Chang 2004). The focus was on spoken arguments and the theses, i.e. claims, to be defended were typically not the respondent's own. The *disputatio* was also used to qualify candidates for privileges and tasks related to teaching; the title *doctor* is derived from the Latin verb *docēre,* "to teach". Upon passing the *disputatio inauguralis* candidates were judged worthy of joining the republic of scholars and becoming academic citizens (Chang 2004). During the Enlightenment the medieval scholastic ideal was gradually replaced with the one centred on the pursuit of knowledge (Ruano-Borbalan 2022). The prior collection of separate theses was replaced by a coherent and self-contained manuscript or book, similar to a modern thesis. With the ideal of *Bildung* and the Humboldtian model of higher education, new ideals for the doctoral thesis followed – it was now to be written by the candidate and display originality (Chang 2004).

Thus, the ideal has shifted from the medieval orator, the teacher of truths, to the researcher of today, the creator of knowledge. This also entails a shift in what is understood as valid evidence of knowledge: from the arguments put forward in the room to the discoveries presented in writing. Subsequently, the thesis has become the common denominator for the examination of doctoral candidates worldwide. But the specific evaluation practices for the doctoral degree differ. Often a defence of

the thesis is included but differences can be seen in how and when such defence is arranged – at a public *disputatio*, in a closed *viva voce,* or publicly after the thesis has been passed by the examination committee – as well as in relation to what should be defended, which always includes the thesis but often also general disciplinary understanding. Theses may also vary in format, with monographs dominating in many countries and disciplines, while compilation theses are more common in others.

Notwithstanding these differences, most academics are likely to conceive of the examination for a doctoral degree as safeguarding similar qualifications. This is due to the long tradition of the degree, its importance for academia, and the process of global convergence that began already in medieval times and has been further reinforced by modern-day transnational academic mobility and the increasing exchange of examiners in doctoral education. To understand these processes of convergence, evaluation practices must be studied in their national, disciplinary, and historical contexts (Hamann 2019). Research on the assessment of the doctoral degree has mainly been performed in the anglosphere. It developed early in the UK and the US (e.g. Tinkler and Jackson 2000; Morley, Leonard, and David 2002; Carter and Whittaker 2009), and in Australia (e.g. Johnston 1997; Holbrook et al. 2004). It has later been complemented with perspectives from countries such as New Zealand (e.g. Carter 2008), Finland (Aittola 2008), Canada (Chen 2014), South Africa (e.g. Sharmini et al. 2015), and Norway (Kyvik and Thune 2015; Kyvik 2014).

In this article, we contribute with the Swedish case in order to complement this growing research on the particularities and communalities of the doctoral assessment. Furthermore, our research object, doctoral evaluation practices, needs to be analysed and interpreted within a particular national and historical context. In contrast to many other countries a Swedish thesis is formally published before the defence and the evaluation. Thus, the examination committee's assessment is a summative one. This means that the assessment is not utilised to revise or amend the thesis and the verdict is hence either a pass or fail. This procedure has similarities with the assessment of applicants for academic positions.

Regulations for the doctoral assessment are found mainly at the institutional level. Normally, a committee consisting of either three or five members are appointed by the faculty dean upon suggestion from the candidate's department. The committee members are typically expected to have a degree of scientific expertise equivalent to associate professor ("docent" in Swedish) and at least one should belong to another institution.

The Swedish defence is a public, highly ceremonial and frequently attended event, in which academics, students, family, and friends are present. An opponent from another university, often one abroad, leads the academic discussion, and the defence may begin with the opponent and/or the candidate giving prepared presentations related to the thesis. The members of the examining committee, as well as any other member of the audience, have the right to ask complementary questions. After this public act, the examination committee withdraws to a private room to discuss the thesis and the defence. They are normally joined by the opponent and the main supervisor, neither of whom may take part in the decision. Furthermore, these two persons may be asked to leave before the final decision is made. Although

the committee has the real power to do so, the failing of a thesis on the day of the defence is extremely rare (Stigmar 2019).

A firm tradition, earlier reinforced by legislation, holds that nothing from the committee's discussion should be communicated outside of the room, except in the very rare event of a failed thesis. In such a situation, a written justification should be produced.

The Swedish doctoral defence is sparingly regulated on the national level. Instead, it is shaped by tradition and older, now-revoked laws which were more elaborate. Of note, however, is the introduction of national learning outcomes in 2007, closely mirroring the level descriptors in QF-EHEA (2005). Nine of these outcomes describe the competence of a candidate, but only one concerns the quality of the thesis: "For the Degree of Doctor the third-cycle student shall […] demonstrate through a dissertation the ability to make a significant contribution to the formation of knowledge through his or her own research." (Swedish Higher Education Ordinance 1993).

Assessment of Doctoral Candidates

Most available research on the assessment of doctoral candidates is based on analyses of formal evaluation reports. These studies have largely focussed on characteristics of passed and failed theses. Johnston (1997) identifies two types of comments in such reports. The first was directed at intellectual endeavours and focussed on significance, scientific contribution, rigour, and originality. The second type was directed at communicative aspects and focussed on literary presentation, style, and accuracy, indicating that examiners appreciate logical, focussed, and concise texts which show the path taken by their authors. Others have reported similar results, although the importance of theoretical frameworks, critical reviews of earlier research, and connections between the literature and findings have been more emphasised (Holbrook et al. 2007; Aittola 2008). The demonstration of an ideological position and a clear doctoral voice were also valued qualities (Mckenna, Quinn, and Vorster 2018). Examiners' reports were indicative of holistic approaches to the assessment, meaning that examiners saw more than the sum of the parts (Holbrook, Bourke, and Fairbairn 2015). These results were verified in an interview study by Lovitts (2007) in which examiners were asked to make implicit criteria explicit.

Inconsistencies and variations in examiners reports have been discussed in several studies (e.g. Johnston 1997). Aittola (2008) describes assessment practices with internal discrepancies and lacking criteria for assessing doctoral theses. Furthermore, correlations between the recommendations made by examiners have by some been found to be low, indicating low reliability in the initial assessment process (Kemp and Mcguigan 2009), while others have noticed a high level of consistency (Holbrook et al. 2008). Written guidelines with specified requirements on the report and clear criteria for judgement have been suggested to reduce observed inconsistencies (e.g. Johnston 1997), but in an interview study, only a third of the examiners used such criteria when they were present (Mullins and Kiley 2002). Carter and Whittaker (2009) found that conversations between examiners in meetings before

and after the actual viva contributed to a shared culture and agreement among examiners.

Experienced examiners' views on the purpose of doctoral studies have been found to lean towards either the production of a thesis with certain characteristics, or the development of the skills and attitudes necessary for independent research, or a combination of both (Mullins and Kiley 2002). Aittola (2008) recognised that the focus was often directed towards the doctoral thesis as a printed document; the evaluation focussed on the quality of research rather than the competence of a candidate.

Studies concerning the oral defence are therefore interesting, as these may provide further insight into how a candidate's competence is assessed. There is, however, no consensus in the literature regarding the function of the defence. Several motives have been identified, including checking the candidate's understanding and ability to produce and present research, clarification of weaknesses in the thesis, ensuring authenticity, allowing the examiners to further develop the candidate's ideas and to provide advice on publication, checking whether the candidate can defend the thesis, testing the candidate's knowledge of the broader literature, testing oral skills, contributing to the final decision in borderline cases, and acting as a rite of passage (Tinkler and Jackson 2000). Despite this, the contribution of the oral defence to the overall assessment is unclear (Bourke and Holbrook 2013). Carter (2008) recognised two opposite opinions among academics, wherein some deemed the oral defence as entirely benign, while others "insisted that, although candidates ought to be fine once they proceed to the oral, it was still an examination, and they should not have the champagne on ice" (ibid p. 371). Kyvik (2014) argues that the practice of pre-evaluation of theses, in order to avoid embarrassing failures at the examination, increases the risk that the approval of the doctoral degree is implied before the defence.

Although Powell and Green (2007) anticipate a trend towards public rather than closed oral defences, Kyvik (2014) discerns mixed opinions among examiners. Those in favour of a public defence saw it as fair to a doctoral candidate, a good opportunity for the candidate to present and defend their research, and a ceremonial occasion which emphasised the rite of passage. Those who opposed it – of which many shared the UK context and the tradition of closed vivas – argued that the practice made it even harder to fail, and that the ceremonial situation impeded examination of important details, making the defence more a formality than a real test.

Variation in the examination practices can also be seen in relation to the thesis format. The traditional format for theses is the *monograph*, a doctoral candidate normally being its sole author. Although supervisors and local communities for research and education contribute to its quality, the monograph is traditionally perceived as a creation of the doctoral candidate – from the original idea and research design to the analysis, interpretation, and conclusions. It is, however, increasingly common to write *compilation theses* (with published research articles and sometimes unpublished manuscripts as well as a synthesising introduction), especially in STEM (Aittola 2008; Kyvik 2014; Mason and Merga 2018) and in Scandinavia also in disciplines where the monograph until recently has been the norm (Powell and Green 2007; Elmgren et al. 2016). Sharmini et al. (2015) describe how many examiners found compilation theses easier to examine, since these were partly peer-reviewed already, while others found it harder,

because of potential incoherence of the articles and the uncertainty of the intellectual input of a candidate when co-authored papers were included. It has been argued that compilation theses not only reflect the reality of research better than traditional theses but also are more transparent, since the traditional format obscures the often-substantial aid from supervisors (Mason and Merga 2018). With regard to difficulties in determining the student's contribution to the thesis, an oral examination has been suggested as a remedy (Sharmini et al. 2015).

The research on evaluation of doctoral theses does not normally include reflections and meta-discussions on the evaluation practice as such. We think that is a weakness. Some scholars, however, have reflected on their own experiences as examiners in doctoral education, pointing out problematic areas and offering suggestions for development. Examples are Grabbe (2003) and Alexander and Davis (2019) who stress the pivotal role of the external examiners as gatekeepers, discuss how examiners can promote excellence, and also emphasise the essentiality of the oral defence. Moreover, they underscored the indispensable local responsibility for upholding quality. Experiences and opinions of examiners have also been captured through surveys (Tinkler and Jackson 2000; Kyvik 2014; Sharmini et al. 2015), an informal panel discussion (Carter 2008), and interviews with opponents (Aittola 2008).

By specifically studying committee members' experiences and reflections on their practice and collegial deliberations we will contribute insights into the assessment of doctoral theses and candidates. As Lamont (2009: 7) argues: "Debating plays a crucial role in creating trust: fair decisions emerge from a dialogue among various types of experts, a dialogue that leaves room for discretion, uncertainty, and the weighing of a range of factors and competing forms of excellence. It also leaves room for flexibility and for groups to develop their own shared sense of what defines excellence — their own group style, including speech norms and implicit group boundaries" (ibid.: 7). Therefore, it is important to investigate what happens in the room with the examiners.

In our case, following the Swedish tradition, that room is closed, and the deliberations are not documented. That is one of the reasons why we have chosen to arrange focus group interviews with experienced committee members to hear their description and understanding of the nature of their evaluative work. More important is that focus groups give insights not only into their actions and thinking in specific cases, but also into their reflections on the evaluation practice in general, founded in their long experiences and rich in examples and illustrative cases. The use of focus group discussions with experienced committee members provides insights into a complex and exclusive practice that the researcher otherwise would have difficulty to access and make sense of.

Theoretical Perspective: Evaluation Practices as Boundary-Work

To evaluate is to find and value differences among phenomena or entities regarded to be similar enough to be compared (Zerubavel 1996). To sort things out – in relation to perceived standards – is a fundamental human activity (Bowker and Star

1999). Whether spontaneous or organised, the categorisation of differences involves the application of *symbolic boundaries*, i.e. 'conceptual distinctions made by social actors to categorise objects, people, practices, and even time and space by which individuals and groups struggle over and come to agree upon definitions of reality' (Lamont and Molnár 2002: 168). Symbolic boundaries are 'the lines that include and define some people, groups, and things while excluding others' (Lamont, Pendergrass, and Pachucki 2015: 850). Hence, they are of importance in 'the creation of inequality and the exercise of power' (ibid.: 850) and in evaluative practices (Lamont 2012). Such conceptual distinctions frame and structure practices, but are also constitutive of a wider landscape where different communities of practice are connected and related to each other by the boundaries they share (Wenger 1998). To participants of such communities, symbolic boundaries form aspects of the taken-for-granted norms and infrastructure of practice, which in part tends to become almost invisible (Bowker and Star 1999). However, there are also in every society examples of symbolic boundaries which when applied are painfully present and contested.

Academic life is characterised by an abundance of symbolic boundaries, expressed both in everyday work and in the structures that have developed over time and that frame and shape disciplines, departments, academic practices, and hierarchies. The doctoral degree is one of the most important, as it separates those who aspire for higher positions in research and teaching from other employees in academia. Examination committees are set up to realise this symbolic boundary in action.

In doing so, committees engage in boundary-work, which is the effort through which symbolic boundaries are upheld, justified, or changed in social situations. Gieryn (1983) describes how academics guard, maintain, and strengthen the field of science by finding situationally valid arguments to clarify demarcations between science and non-science. Bowker and Star (1999: 296) conclude that 'scientists are especially good at documenting and publicly arguing about the boundaries of categories'. The peer review process is a ubiquitous example in which academics, on the basis of professional abilities developed throughout their careers, apply both implicit and explicit notions of quality and appropriateness (Lamont 2009).

Where boundary-work creates and maintains symbolic boundaries, a *boundary object,* in a theoretical sense, does the opposite. A boundary object is a material or symbolic artefact that may harbour and serve several different categorisations and standards simultaneously (Star and Griesemer 1989). Bowker and Star (1999) note the important role played by boundary objects as interfaces between different social worlds or communities of practice by being robust enough to maintain coherence while at the same time being plastic enough to serve different needs. Wenger (1998) describes boundary objects as artefacts or social practices that are shared by different social agents but not necessarily for the same purposes or instilled with the same meaning. We view the doctoral thesis as an important boundary object.

Boundary objects are situated at the crossroads of different systems of categorisations. In a formal evaluation, in which the outcome of an assessment may have tangible consequences, the encounter between different systems of categorisations applied in the evaluation process can lead to *tensions, paradoxes*, and *torques*

(Bowker and Star 1999). Tensions arise when different categorisations lead to different judgements, while paradoxes may appear when different systems of categorisations are perceived as being incommensurable. Torques are acts of classification that lead to severe consequences for individuals and their biographical trajectories. An example of torque is the (rare) case of a failed thesis. In the context of examination committees' evaluations, all these outcomes are possible. The ceremonial staging of the examination event probably supports the maintenance of the thesis as a boundary object in fulfilling different expectations.

Research Approach, Data Collection, and Analysis

To explore the experiences of committee members and their reflections on the evaluation practices, data was collected using focus group interviews. Six focus groups were interviewed, two at each of three large research-intensive Swedish universities, over the course of five and a half years. Each focus group was composed of between five and nine experienced committee members with similar disciplinary affiliation, who therefore shared a disciplinary understanding. The disciplinary areas were the humanities (Hum), social sciences (SoS), biology (Bio), mathematics (Math), medicine (Med), and educational sciences (EdS).

All participants were informed about the study and gave consent to participation, were guaranteed confidentiality in taking part in the focus group, and anonymity in relation to how the findings of the project would be presented.

The conversations were rich and often intense, and could be characterised as qualified professional exchanges, in which potential utterances might be either chosen or avoided due to the impression the interviewees wanted to convey in relation to the other participants. Collectively the interviewees had broad experience from committee work at many different universities both in Sweden and abroad. They also referred to a broad variety of cases from several disciplines and subdisciplines, as well as from interdisciplinary fields.

The guiding questions we posed were primarily focused on exploring the foundation for the decision of the committee: What do you assess? What qualities are you looking for? We also asked the members to describe in what ways other contributions, beside the published thesis, were regarded in the evaluation. Other questions were: How does the committee make its decision? What is included in the deliberation? What role does different formalities play? Finally, we asked them how they viewed the task of the evaluation committee and its future role.

Each interview lasted approximately an hour and a half. They were recorded and later transcribed verbatim (almost 100,000 words). Quotes used in this article were translated by us. The analysis was conducted by reading each full interview (initially while listening) and making an initial thematic content analysis. This was done separately by two of the authors for all six interviews. After comparing and linking the preliminary themes from each of the interviews, the process included a comparison of the themes across all six interviews, which enriched and consolidated their scope and sub-thematic variations. This was done by all three authors together.

Table 1 Condensed subthemes and themes

CONDENSED SUBTHEMES	THEMES
Expected and observed qualities	**The gradual emergence of the object of evaluation**
Initial components of the evaluation object	
The function of the defence	**The situational logic of the day of the defence**
Theses format: compilation theses and monographs	
Formalised criteria and regulations	
Disciplinary and interdisciplinary contexts and local cultures	**The extended boundary-work of the examining committee**
Nature of the deliberation and possibilities for feedback	
Integrity and fairness	
The responsibility of the department	
International and historical comparisons	

In the initial reading of the full interviews two overarching themes emerged that characterised the evaluation process, initially at a superficial level but later deepened as the analysis continued. The first theme encompassed the committee members' descriptions of how and in what ways their perception of the thesis and the competence of the candidate developed from the initial and individual reading of the thesis to the final decision: The object of evaluation as a gradually emerging and generated entity. The second theme related to the committee members' descriptions of increased intensity and internal tensions in the process of generating the object of evaluation and leading up to the final decision, strongly linked to the expectation to pass but with the real possibility to fail. We labelled this theme as "the situational logic". These themes spurred our search for a theoretical framework that would make it possible to delineate the evaluation object as a symbolic rather than factual entity and for which the boundaries were negotiated in the collegial deliberation. Our theoretical framework was thus developed in parallel with the analysis.

In the next step, a bottom-up content thematisation resulted in a large number of subthemes (including compilations of citations), which were then summarised and condensed into 10, which both informed the two overarching themes and underpinned a new, third overarching theme (the extended boundary-work of the examining committee) and thus contributed to the structuring and descriptions of the findings at the integrated level (Table 1).

Fig. 1 Three aspects of the object of evaluation in a doctoral examination
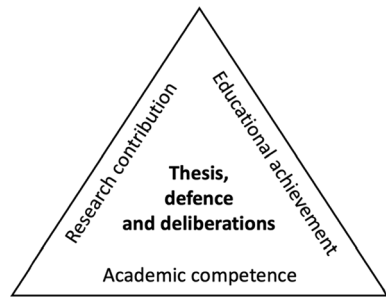
Findings

The evaluation practice was described by the members of our focus groups as a complex and gradual process, through which the committee members gained a more elaborate understanding of the qualities of the thesis, but also of the circumstances and conditions of its creation. In this process, an object of evaluation was constructed with which the committee could engage in their boundary-work. In the often-intense deliberations, further fuelled by the strong expectation of a passing grade, the boundary-work frequently expanded from the thesis and the candidate's competence to encompass also the supervision as well as the organisational and social context for the candidate's research and education.

The Gradual Emergence of the Object of Evaluation

With the thesis as a starting point and stable artefact, the construction of the object of evaluation began when the individual committee members received the thesis and ended with the final decision, arrived at after the public defence and subsequent closed meeting of the committee. The transient evaluation object encompasses the research contribution, educational achievement, and academic competence of a candidate (see Fig. 1). The doctoral candidate is thus simultaneously assessed as a junior researcher, an advanced student, and a developing academic.

Reading the Thesis – Individual Assessment of Research Quality

Throughout the evaluation process, our informants looked for many aspects of intellectual endeavours, while communicative aspects were typically mentioned only in passing. The gate-keeping function was seen as the most important. Many had their own checklists in order to find, for example, severe methodological errors, inconsistency, or deficient comprehension of the field. This practice was challenged by some, who valued brave and promising candidates who had produced "an interesting failure" [Hum] more than candidates who had produced formally correct but uninteresting theses. This was also reflected in discussions

about the potential need for formal criteria, which some informants expected to raise quality, although others feared would result in less creativity and more trivial theses, while also diminishing the space for expert evaluation, leading to "collective stupidity" [SoS].

All groups repeatedly stressed the importance of independence and a scientific contribution. A good candidate should have made a "take-off from the base" [Bio] presenting his or her own ideas and showing critical reflexivity. Characteristics such as "originality and to some extent courage which is not foolhardy" [Hum] were appreciated. A grasp of the field was considered fundamental in order to be able to make an important contribution; awareness of earlier research, with both appreciation and a critical eye, were sought after.

A "conscious structure" [Hum] with clear connections between research questions, material, methods, theoretical perspectives, conclusions, and reflections, was deemed vital – although some informants argued that consistency within individual articles (in a compilation thesis) was enough, and that a variety of methods or theoretical perspectives could enrich a thesis, if a candidate could "tie together" [EdS] the parts. Research questions should be answered and theoretical frameworks used; high ambitions that were not followed-through were seen as a lack of consistency: "It shouldn't be some kind of theory examination" [Hum]. Some mentioned that it was sufficient if someone had "developed one of two important sides of their work" [EdS], either through theoretical or empirical contributions. It was held by some that it could be regarded as more important to bring new perspectives than new results.

In areas where compilation theses were the norm, the previously published work was seen as a sign of quality and also of having made a scientific contribution. Furthermore, several aspects of knowledge, understanding, and skills were recognised as inherent in the published work: "Isn't that what we boil down to international publishability?" [Bio]. Consequently, if a paper was not yet published, its potential to be published needed to be assessed.

It became clear that the evaluation practices described by most focus groups rested primarily on traditions within the specific research fields and disciplines rather than on educational regulations:

> Practice has developed differently in different parts of biology. And we still live with that, we work according to patterns that our predecessors founded. [Bio]

When the national degree learning outcomes were shown in the end of the interviews, more aspects (like research ethics and the ability to discuss research authoritatively in international contexts) were discussed. These learning outcomes were seen as relevant, but they were not normally used in the deliberations and evaluation. Although these outcomes are intended to be fulfilled, some saw them rather as something to strive for, and others pointed out that the task of the committee is to focus on the research contribution presented in the thesis, and that the assessment of all the other learning outcomes should be handled by the department and the educational programme to which the student belongs.

Attending the Defence – the Assessment Widens

Even though the thesis was the focal point for the evaluation, the defence contributed to the emergence of the evaluation object. The defence thus fulfilled many important functions in the assessment, especially concerning the competence of a candidate, but because it takes place in the present and is not documented, its contribution was also problematised.

The defence allowed the candidate to demonstrate his or her understanding of the thesis and the field. Moreover, it showed "how the doctoral student acts under pressure – and despite that still manage a public discussion about the work" [Med]. Some even had experienced that it was "seriously discussed that this doctoral student should be rejected – because of the defence" [Math]. On the other hand, a candidate's "shyness" or "nervousness" could obscure their competence. The difficulty of deducing a candidate's competence from the defence, combined with the general reluctance to fail a thesis, was problematised and the defence was in several interviews discussed as a "play to the gallery" and a suggestion was to make it "less of a party, and more of a test". [Bio]

The ceremonial situation and the public format were, however, also brought to the fore as a quality enhancer:

> Compared to the British system, there are quality enhancers in the public event in that you will eventually have to defend something publicly. It is in itself quality enhancing – no one wants to defend something that [interrupted by another interviewee who agrees] [Bio]

Furthermore, the defence was taken seriously since it contributed to the reputation of the doctor, positively or negatively:

> But above all, I think you can do well, you know, a good defence is furthermore something that a large audience sees. So, then there are a lot of mouths that can pass it on further and say that she had, she defends, brilliant defence, you know, it was a joy and, to see how skilfully she handled eh, eh, eh, eh the opponent's questions, and, and, and excelled up there. And it's something that people like to talk about and that, it can of course affect a career. [Math]

> You take comfort in the fact that this person has no future in this business. [Bio]

The pivotal role of the opponent for the academic discussion was summarised as "it takes two to tango" [Bio]. A bad opponent, e.g. one who either "thinks that you should, sort of, bust the doctoral student [or] thinks that 'now we are going to do something really fun for, for mom and dad here' – and nothing else" [Math], could make it difficult for a respondent to shine. In such cases, the committee members stressed the importance for them to compensate by asking more questions "from a list [from which you] delete as the opponent ticks some of them off" [Math].

The possibility of using the defence to make up for a poor thesis was further problematised in relation to the committee members' own academic standing, since the defence was not documented:

I could very well imagine that […] the respondent shows […] abilities that are not visible in the thesis, […] but in the future I must be able to defend why I approved that thesis. [Math]

After the disputatio the defence lives on only in the memory of those present. The thesis remains.

Meeting with the Other Committee Members – New Perspectives and Joint Assessment

The closed deliberations following the defence marked the beginning of the collective phase of the assessment process, with more perspectives contributing to the understanding of the qualities of the thesis and the candidate, and hence to the construction of the evaluation object.

The choice and composition of committees were normally considered satisfactory. For the individual committee member, input from other members was described as vital since previously unnoticed aspects could be revealed. This, in addition to the more outspoken comments that the opponent may provide in a closed room, and the possibility of asking questions of the supervisor, contributed to a richer view on both the thesis and the competence of a candidate. Furthermore, the presence of other committee members made it possible for an individual to also grasp the character of the evaluation object from other perspectives than their own. Thus, the composition of the committee and the choice of opponent were crucial to reassure members in facing their individual responsibilities for their joint decision.

Committee members were not necessarily experts on a candidate's particular strand of research and were often chosen to have complementary competences; they could be chosen for their broad disciplinary knowledge, insights in a particular field, belonging to other disciplines in order to assess the thesis' broader relevance, or expertise in an important subsection of the thesis (in our interviews often exemplified with statistics). This expert role was by some regarded as less demanding, since the expectations on them for the evaluation were demarcated, while others thought it was "easier when you have better understanding of the field" [Math]. The presence of an opponent, for instance from another country, allowed a form of benchmarking as "one of the questions that is always posed to the opponent is 'would this dissertation be approved in your system?'" [Bio].

Which role a specific committee member would have was usually not made explicit beforehand and sometimes remained unclear in the deliberations:

Am I here because I represent the discipline in general or am I an expert on the issue of poker or whatever it is about? Or am I the one, the cheap member from home? [laughs] Or what is really my, my role in this context? [Hum]

In rare cases the composition was not deemed appropriate, such as when nobody in the committee had a proper overview of the field:

I had an idea about why they had asked me to sit on the grading committee. But it turned out when I got there that none of the others on the grading com-

mittee were in that area. And that was problematic. And the opponent was very critical. And then the question is: What do you do? So, it's also a thing like this to check out who they are? Ha, ha. Who are the members of the grading committee? Is there anyone who can actually make a, an assessment? [EdS]

The composition of the committee and choice of individual members was further discussed in relation to potential biases. The appointment of committee members and opponent was often made based on suggestions from the supervisor, the rationale being that the careful composition of the committee was vital for the deliberations and assessment, and dependent on thorough knowledge of the field. The supervisor's involvement could however lead to the defence being "rigged" and committee members suggested "not on grounds of their competence but on their expected loyalty" [SoS] and from "the supervisor's network" [EdS]. Another way the committee could be biased was by choosing committee members who lacked sufficient disciplinary insight to be able to argue for and insist on a fail. As a consequence, some informants said they informed themselves about the other members invited to the committee to be able to "check who they are" [EdS].

Deliberating Qualities and Disentangling Contributions

Through our interviews it became apparent that a candidate's academic competence and educational achievement were assessed mainly through the evaluation of the candidate's research contribution, which was therefore the dominant aspect of the emerging evaluation object. In describing this evaluation our informants stressed the importance of disentangling the candidate's research contribution from that of supervisors, peers, and co-authors. The necessity, and difficulty, of this disentanglement was a recurrent theme in our interviews, irrespective of whether our interviewees discussed monographs or compilation theses.

Although the monograph was primarily perceived as the creation of the doctoral candidate, committee members were attentive to potential contributions from others. A monograph had for example in the closed meeting been called an "Ellinor thesis" [Math], alluding to a Swedish song that includes a line from a child: "This I've written almost only on my own", to illustrate that the thesis perhaps had been partly written by the supervisor.

For compilation theses, which normally include articles with several other authors, including the supervisor, and an introduction called a *kappa*[1] summarising and/or synthesising the compiled articles, the process of disentangling a candidate's contribution was found to be especially delicate. This was done through analysis of the relation between the articles and the kappa, written statements on contributions to the papers, the candidate's defence, and testimonies from the supervisor. However, the perception of the function and importance of the kappa varied substantially both within and between focus groups. In the humanities group, the kappa was seen *as* the thesis, the main contribution to the field, and the articles viewed as supporting

---

[1] Swedish for coat.

material. In other groups the kappa was instead seen as the part of the thesis that made the particular contribution of the doctoral candidate visible. It could also be considered primarily as a formal summary of minor importance.

In the process of disentangling the research contribution, a candidate's authorship could in rare cases be contested altogether. Different reasons for this were discussed in our interviews, some of which also cast a shadow on the research environment. For example, the need among supervisors to publish was mentioned:

> Several of the supervisors are suddenly taking part, and you get a feeling that there is a lot at stake for them as well […] you can sometimes suspect that it is the supervisor who writes [the papers] because you notice some doctoral students who suddenly have huge problems with writing a kappa. [EdS]

The importance of rapidly publishing results, perhaps at the expense of building a candidate's competence, could also be linked to the overall research culture in which "you can´t forget that research has become somewhat of an industry, which should generate money […] and then you don't want a working force fiddling with words that hold no significance for your own research." [Med].

Compilation theses were considered by some easier and by others more complex to evaluate. Some found it "easier to read a compilation thesis since then the different parts are broken down into articles." [Math]. Another advantage was that "every part of the thesis has gone through this assessment before." [SoS]. Progression could also be discerned when not all parts of the text were published at the same time: "you can see when the articles, the first articles, what they look like, and how the development has been towards the end." [EdS]. On the other hand, and in addition to the difficulties in disentangling the contribution of a candidate, the critique directed towards published articles also spilled over to the other authors concerned. Furthermore, it resulted in the committee "opposing the quality of the journals involved." [SoS].

If, in the ongoing deliberations, the research contribution of a doctoral candidate began to appear insufficient, the discussion changed to also include academic competence and educational achievement. In the contextualisation of these aspects, reasons for the insufficiencies were discussed also in relation to external factors. This was reinforced by the logic of the situation.

The Situational Logic of the Day of the Defence

As described earlier, Swedish evaluation committees traditionally do not grade theses or present a written statement of their assessment (unless the thesis is failed). Furthermore, the thesis is published in a final version and with the firm expectation that it should pass, which is underscored by the ceremonial staging with friends and family attending and a reception prepared, often just outside the lecture hall. The "failing" of a thesis, or prescribing its revision, in the Swedish tradition is expected to take place before the defence. When committee members receive a thesis, a minimum of three weeks before the defence, they described this as an opportunity to pull the "emergency brake" [SoS] should the thesis be obviously fraught with errors or in other ways not up to their expectations. Pulling the emergency brake meant

contacting the organisers who typically would arrange a meeting in order to assess and handle the situation, usually by postponing the defence until the thesis had been sufficiently improved. A unanimous opinion was that any doubts about a pass should be raised early, and that "it should almost be considered indecent if one did not do so" [Bio]. At the same time, it demanded a large degree of certainty about severe shortcomings of the thesis, given the consequences. Furthermore, it necessitated that the thesis was read as soon as possible, and not too close to the defence. If the brake was not pulled, the doctoral defence would take place as planned, which most likely meant that the committee would pass the thesis, even if the evaluation process revealed new and previously unnoticed problems. All this gave rise to a growing pressure, to which the ceremonial framing also contributed:

> The family is waiting for the party, the flowers are on the table, the grading committee is invited to the party and the opponent is invited. Can you fail? No, even though the mistakes, so to speak, come creeping up. [SoS]

Due to the gradual emergence of the evaluation object, the extent of potential shortcomings of the thesis and/or a candidate's competence sometimes could not be fully realised until after the defence, at the closed meeting, from which the committee members were expected to return with a verdict of pass. Being compelled by the situational logic to pass a poor thesis and/or defence violated the committee members' standards and created a strong sense of unease, which remained many years later.

> It was a very painful defence, and we sat for a long time in the examination committee afterwards and it caused us extensive anguish before [the decision]. And I'm still in pain because we approved this… [Bio]

Some interviewees expressed a feeling of sometimes being almost held "hostage" [Hum] to the situation, and one member exclaimed: "So, it goes against every fibre in one's being, you know, but yet one [passes]!" [Bio]. In such situations, members found themselves in a dilemma where damage would be done irrespective of whether the thesis was passed or failed.

Expanded Boundary-Work

The deliberations of the committees often included discussions of the local conditions for research and doctoral education, as circumstances in the disciplinary and departmental environment shed light on both shortcomings and strengths in the thesis and the defence. Such discussions were also valuable as opportunities for benchmarking of expectations and norms for research and doctoral education within the discipline or field in general. Thus, the boundary-work expanded.

> The task of the grading committee is much, much, much, much wider than that. Because it is almost always the case in practice that the theses are approved because otherwise people tend to withdraw, but on the other hand it is such that, all the activities, the conversations that take place in the examination committee, after all, three persons are gathered, and the supervisor and the opponent, at the same time, who to a reasonable extent have read the same

Fig. 2 The object of evaluation expanding to include the local context: research practices, educational provisions, and academic environment
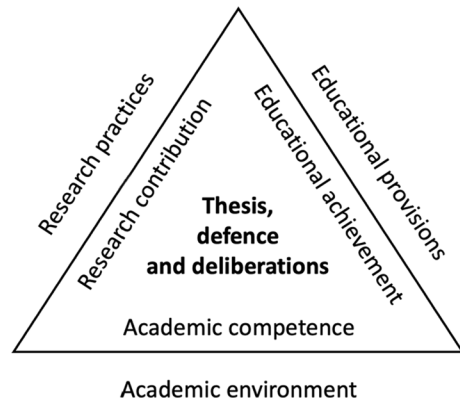
dissertation, in addition to having seen the same presentation and the same defence. And, and this is a great opportunity to, to make general, a general calibration of levels even though the outcome is, from the beginning, is given in most cases. [Math]

The collegially created object of evaluation thus worked as a case for normative discussions on research practices and educational provisions as well as on the academic environment to which a candidate belonged (see Fig. 2).

Identified shortcomings would further spur discussions on supervision and the local context for research and doctoral education.

So you have a dissertation that is very weak, and you have to think a little about why it is so weak? And then, what happens is that you sit and discuss the supervisors' efforts. And that was quite clear, I think in this case. That it, it's a supervisor failure. Definitely so. [Math]

The committee members were adamant that a weak thesis and defence should not go unnoticed, even though the thesis was eventually passed. The responsibility of the supervisor and local environment for doctoral education and gate-keeping was brought to the fore and critique was "directed towards the department and the faculty." [EdS]. Thus, committee work was part of a "scientific norm formation and socialisation" [SoS] in which "the most important aspect of the committees' work is to send signals to the department" [Bio], with the hope that a "painful discussion in the committee has an impact on the department." [Bio]. The tradition of not giving feedback to a candidate further enabled frank discussions and expanded boundary-work.

The respondent does not really get any feedback, and in some way that is one of the strengths with the meeting in the examining committee, that the arguments and discussions are treated as a secret […] and that such a critique, if you put it forward in the committee, I think is regarded more seriously, since you may state it more bluntly, I think so anyway, that the committee discussions could really be quite harsh and the message aimed at the supervisors or

at the professor in charge of the department, and that is possibly a strength with the system we have. I think that the feedback to the department is a very good thing, and the best part of our system, really. [Bio]

In particular the supervisor, and to some extent the local committee member(s), were the designated recipients of the conclusions of such discussions.

It is a very hidden but extremely instructive process both for the dissertation department and for the home department to which you later go back. [Bio]

In general, committee members agreed that the closed meetings and absence of written statements were prerequisites for serious discussions. Had they been public, they would have had to be restricted to the thesis and the respondent, and also be bolstered out of consideration for the candidate, friends, and family: "If you were to start documenting them, I suspect that this quality would disappear." [SoS].

Committee members described substantial intellectual rewards from such discussions, which in the long-term could benefit their own research and educational context. Each participant at the closed meeting would leave with a yardstick to be used in the next committee assignment or for their own practice as supervisor and researcher. In this benchmarking a broad representation was considered important:

In cases where you have a slightly larger context represented, with members from other countries and such, you have a great opportunity to have a very important discussion. [Math]

The consensus about the importance of the normative function of the discussions in the closed meetings was large.

Disciplinary Belonging and Variations in Boundary-Work

The expanded boundary-work of the committees was described in all of our six focus groups, but in different ways both within and among groups. This was most apparent in the way the committee members described their involvement in the boundary-work leading up to the vote for a pass of a weak thesis. The composition of the committee and the members' relative closeness or distance to the discipline of the thesis could affect the extent of the expanded boundary-work. The normative discussions were a form of negotiation in which the local circumstances, traditions, and challenges had to be considered in relation to the norms and developments of the discipline. It was seen as important to be observant of disciplinary drift and systemic quality issues within doctoral education as well as in the research field.

Because it is true, as long as it is only an internal question to let a thesis through. Then it is very hard then to come and say something other than that it must be approved, then it will be up to that department to do it – the actual quality test. As we get more and more universities, there is a risk that we start to have different ideas about what counts as good political science, for example. [SoS]

Between sub-disciplines there could also be both different practices regarding assessment and different views on quality. Something regarded as innovative in one area wasn't necessarily deemed as such elsewhere, or the research was just not considered to have been performed in a good enough way. This made it difficult for candidates in the borderland: In this "interface between disciplines, you find problems" [Bio].

Such discussions required both tact and moderation. Committee members farther from the discipline had an important but delicate task since they may lack knowledge on local circumstances and disciplinary particularities.

> When you sit there, you are often not at your own department and maybe not even in your own discipline. Then you have to somehow get a feeling for what is passable and what you can say as an outsider. I cannot come and totally diss a thesis that they might think is quite good. And then it becomes a bit of a diplomatic question. I would like to express what I think but at the same time you cannot be awkward. [SoS]

Despite these inherent challenges, it was considered important to compare related sub-disciplines, since the discussions served as a form of benchmarking, leading to "some kind of consensus" [Bio]. Norms taken for granted in a sub-discipline could also be contested in a useful way, as when someone "asks why the emperor has no clothes" [Bio] (alluding to the tale *The Emperor's New Clothes* written by H. C. Andersen).

The consensus about the importance of the normative function and expanded boundary-work of the discussions in the closed meetings was great and seemed related to committee members' sense of disciplinary belonging and responsibility. Particularly in the focus groups in biology, the humanities, and mathematics the expanded boundary-work was described with intensity, underlining the importance these participants placed on clarifying the relationships between the character and qualities of the evaluation object, and the local academic environment and the larger field of research.

However, in the focus group from medicine, where the disciplinary belonging was least pronounced, poor theses and defences were discussed in an almost neutral manner. Instead of describing animated discussions on supervision and the state of the discipline, the interviewees described looking for reasons, and resolutions, for candidate failures that lay outside the research community, in the administration or in the control of the educational process. The expanded boundary-work was hence less intense and tended to focus on formalities:

> There are these instructions available. And must there be at least four works, must there be one where you are the main author? And like going through the whole process and submission and resubmission and, eh, and that there should be at least two published and two as first author. So like, the formalities you usually check off and stuff like that so they're fulfilled. [Med]

Despite the commonly held positive opinion about the normative discussions in the closed meeting, a few interviewees expressed concern that these discussions

had relatively limited consequences for the field and the local practices, since the committees' conclusions would not necessarily reach key persons who could influence and change practices.

> So somehow it is a pity that you do not document those conversations, because they are very important when you're there with the others and when you retell it […] but then you forget it, so there is no memory left, so to speak, which is documented. [SoS]

Thus, the closed doors were seen as a prerequisite for the outspoken and formative discussions, i.e. the extended boundary-work, but were also problematised since there were no guarantees that these discussions would transcend the room, and thereby serious quality issues could prevail. Then, the collegially created evaluation object would dissolve as soon as the doors opened and the committee announced the simple message of a pass.

Discussion

Our focus-group interviews revealed boundary-work of great breadth and complexity, far exceeding the scope of more regular forms of educational assessment in present-day higher education. Instead, the evaluation practices described in our results reflect the disputatio's long history and traditional core function within academia and its disciplines. Hence, we can see how the disputatio, as an institution, is important for realising, upholding, and changing boundaries for the doctorate over time, across national borders, among subdisciplines, and at interdisciplinary junctions. In this capacity it can, potentially, safeguard or challenge disciplinary norms as well as encourage or constrain creativity and inter- and transdisciplinary excursions.

Examination Committees Value Similar Qualities in the Doctorate

We can confirm that there is an international common ground for the evaluation of the doctorate, upholding the symbolic boundaries within academia. A candidate's contribution to research, as visible in a thesis, was the dominant facet of the object of evaluation. The qualities our interviewees were looking for were similar to what has been reported earlier: scientific rigour and originality, the thesis' significance and contribution to the research field (Johnston 1997; Holbrook et al. 2007; Aittola 2008; Lovitts 2007; Trafford 2003), literary merits, and the accuracy of presentation (Johnston 1997), as well as a clear doctoral voice (Mckenna, Quinn, and Vorster 2018). Also confirming previous research was our committee members' understanding of the defence as an important rite of passage (Tinkler and Jackson 2000; Kyvik 2014) and as an opportunity for assessing a candidate's independence and contribution to the thesis (Sharmini et al. 2015). However, the relative emphasis on qualities assessed varied between our six focus groups, which was likely a result of differences in epistemology and academic cultures (Becher and Trowler 2001; Lamont 2009).

In contrast to previous research, comments on technical errors, weak language, or specific issues with research methods, were largely absent in our study. This can partially be explained by the fact that Swedish committees do not propose revisions. Furthermore, our methodology – focus groups – leads to wide-reaching reflections on a more abstract level, which cannot be found in evaluation reports.

Boundary-Work Extending from the Individual to the Collective

The committees' discussions contained many of those elements of uncertainty, discretion, and compromise that are characteristic of academic evaluations in general (Lamont 2009). Because those present at the closed meeting represented different institutions, and sometimes different disciplines or subdisciplines, these discussions also constituted a disciplinary or interdisciplinary negotiation and a form of benchmarking that over time may lead to a development of consensus among examiners, as previously reported by Carter and Whittaker (2009) in the context of the British viva.

Although some of our interviewees problematised the Swedish tradition of publishing the thesis before the disputatio, on the grounds that it therefore could not be revised and improved, they also saw how this *ex post facto* framing of their evaluation prompted them to discuss norms and expectations for the doctorate. When the deliberations revealed that a thesis had shortcomings and/or a candidate was lacking in competence, the discussion often turned to the supervision or the local context for research and doctoral education as a source of explanation. After all, both the supervisors and the local educational leadership had deemed the thesis and the candidate ready for the defence. This reorientation of the discussion meant that the evaluation expanded from the candidate's academic competence, research contribution, and educational achievement, to also include the academic environment and the quality of local research and educational provisions (see Fig. 2). Hence, the individual is evaluated in light of the local circumstances, and the collective is evaluated in light of the candidate's accomplishments. The boundary-work of Swedish evaluation committees thus involves more than gatekeeping (through failing or passing theses) and setting disciplinary expectations and norms for the doctorate. It often also includes normative discussions about the local context. Albeit no written statements or formal procedures resulted from this extended boundary-work, several of our interviewees described how important these discussions were for themselves and for their disciplines. Hence, doctoral examination as an institution is important for upholding and developing standards and norms within research, doctoral education, and, ultimately, academia (see Fig. 3).

The situational logic further prompted such expanded boundary-work by framing the committee's evaluation with two potential torques (Bowker and Star 1999). The first would be actualised if a thesis and a candidate, contrary to expectations, were to be failed. In such a case, the decision would preserve the committee members' sense of defending academic standards but be catastrophic to the doctoral student and those closely involved. The second torque would be actualised if a poor thesis instead was passed. In such a case, the decision would prevent the first torque but
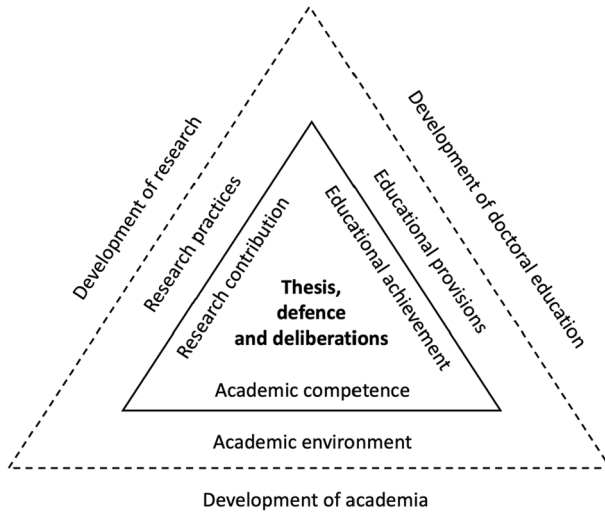
Fig. 3 Potential reach of boundary-work related to committees' evaluation of a thesis and defence, developing standards and norms within research, doctoral education, and academia

instead bring harm to the committee members themselves and to their disciplines. Committees normally chose not to fail, and are therefore more likely to have to deal with and alleviate the second torque – by looking for explanations and placing responsibility for a weak thesis on the collective level, with the supervision, local leadership, and the local environment for education and research.

The seriousness, emotions, pressures, and complex boundary-work our interviewees described underline the importance of the disputatio for academia and its disciplines. The present-day Swedish doctoral evaluation, as presented here, reflects the times when the disputatio was used in conjunction with academic promotions and appointments, as an intra-academic qualification. The three aspects of the evaluation of candidates have, as we have seen, emerged at different times in the history of the doctorate. The evaluation of academic competence was present already at the medieval disputatio inauguralis. With the Humboldtian ideal of original knowledge followed the second aspect, represented mainly by a candidate's thesis. Finally, with the advent of doctoral education the third aspect, educational achievement, emerged.

The situational logic of the day of the defence thus reflects the historical roots of the disputatio as a ceremonial and public display of a new competent member of academia – ultimately a testament to the rigour of a discipline itself. The deliberations on whether a respondent was competent (or promising or independent) enough to join, and represent, the discipline, and the question to foreign opponents whether the respondent would have passed in their national contexts, should be seen in light of the importance of such matters to the discipline. Hence a poor thesis and defence, albeit reluctantly passed, was a failure not only of the student but also of the supervisor and local environment, and ultimately of the discipline. This was illustrated by interviewees who shared how painful it had been, and still was, to pass a candidate whose competence was understood to be insufficient. Tellingly, the frustration

in such cases was not directed towards the candidate but towards the supervisor and local research community – towards other members of the discipline.

Although, as we have argued above, the particular procedures and framing of the Swedish disputatio are conducive to such extended boundary-work, it is highly likely to be an important aspect of the doctoral examination also in other countries and can explain the great importance attributed to the doctoral examination within academia in general (see for example Grabbe 2003; Alexander and Davis 2019). To our knowledge, however, this particular aspect of the committees' boundary-work has not previously been reported. The reason, we assume, is that it is unlikely to be captured by studying examiners' reports, which is the predominant empirical basis for research on the doctoral examination.

Whilst the closed meeting and absence of written statements were seen by our informants as prerequisites for serious and wide-reaching discussions, the Swedish arrangements might also lead to the normative discussions remaining inside the closed room, with lost opportunities for further discussions within the discipline or academia.

Coherence of the Evaluation Object

Committee members' disciplinary belonging appeared to be important for how they understood and related to the emerging evaluation object. This, in turn, seemed to be correlated with whether a thesis with problems, or a poor defence, would lead to expanded boundary-work or not. In groups where the disciplinary belonging was strong, interviewees described how a candidate's research was conceived of as a contribution to the discipline, and how the candidate's academic competence became visible mainly through this contribution and was often cast in terms of committee members' expectations on a future colleague, and thus important. Educational achievement, in turn, was essentially understood to be reflected in the previous qualities. Hence, the three aspects of the evaluation object, the research contribution, academic competence, and educational achievement of a candidate, were aligned and represented a coherent whole. Unsurprisingly, these committee members also described being frequently engaged in extended boundary-work and expressed both pride and anguish when sharing experiences of evaluating strong and weak candidates. These findings illustrate the doctorate as a decisive symbolic boundary in academia.

The coherence of the evaluation object, as well as the extended boundary-work it enabled, was founded on the thesis' function as a boundary object and hence its ability to connect different social practices, as pointed out by Bowker and Star (1999). As such it made it possible for the committee members to perceive the three aspects of the evaluation object as connected, and to imbue the defence and the deliberations with meaning and importance. The richness and coherence of the object thus demonstrated that something important was at stake, not only for the candidate and the department but also for the committee members themselves, as 'stewards' of their discipline (Golde 2006).

However, a lesser sense of disciplinary belonging seemed to be correlated with a partially incoherent, simplified, or even fractured evaluation object. In such cases, the research contribution, academic competence, and educational achievement were assessed partly in isolation from each other, which we interpret as a sign of the thesis losing its function as a boundary object. In the focus group where the disciplinary belonging was least pronounced, we were struck by the relative calmness and distance with which experiences of poor defences were discussed by some interviewees. Furthermore, an incoherent object of evaluation appeared to be associated with lower motivation for extended boundary-work in cases of problematic theses or weak candidates. Instead of animated discussions on supervision and the state of the discipline, interviewees described looking for reasons, and resolutions, for candidate failures outside the research community, in formal administrative routines, or in the management of the educational process. This also suggests that in some academic cultures, a thesis, apart from being viewed as a research contribution, may be understood more as a proxy for educational achievement than for elucidating the academic competence and the potential of a candidate. A fracturing of the evaluation object would undermine the thesis' function as a boundary object and hence weaken the links between norms and expectations in research, academia, and education. It also challenges the doctorate as a symbolic boundary in academia. Such a development can be explained by a combination of factors, such as the formalisation and expansion of doctoral education, the projectification of research, and the dissolution of disciplines.

The Doctorate in the Centre of a Field of Tensions

The last decennia's changes in research organisation and governance have resulted in the emergence of new problem- and innovation-oriented fields but also in the partial dissolution of many traditional disciplines and the replacement of traditional subject-based departments with multidisciplinary centres. Furthermore, the 'projectification' (Ylijoki 2016) of research, which is accompanied by increased dependence on external funding, has led to the research group, often functioning as a 'quasifirm' (Etzkowitz 2003) for the production of research in return for research grants, having become the dominant social context for many researchers and their doctoral students (see, for example, Fochler 2016; Fochler et al. 2016; Sigl 2016). These tendencies have potentially weakened the sense of being part of the broader phenomenon of academia, and may also have affected how the doctoral examination and the boundary-work of the committee are perceived. Even so, the consequences of any new regimes for research are likely to be moderated by cultural, epistemological, and contextual differences between different research areas. In both biology and medicine in Sweden, for example, externally funded projects normally form the basis for doctoral education, which takes place largely within supervisors' research groups. Despite these similarities the reasoning of our focus group in biology was in many respects more akin to the other groups', and notably also to the humanities, which belong to a research area where research groups are rare and external project funding is less common.

While present day research governance leads to a strong focus on the research production of doctoral students, educational formalisation instead emphasises the development of a wide range of doctoral competencies (Elmgren et al. 2016). This formalisation was partly welcomed by some of our informants who saw the value of such competencies for other sectors in society, as well as for their disciplines and for academia at large. It was however also a source of tensions in doctoral education, especially in conjuncture with research projectification, to which it nonetheless could provide a counterbalance. Furthermore, it could also counteract more vicious aspects of a strong discipline, such as the subordination of doctoral students to the will of the professoriate.

Even though doctoral education today comes with degree descriptors, syllabi, and formal regulative and administrative procedures, the presence of these elements in the committees' practices in our study was small. Our findings instead show that the complex character of the evaluation object generated in the evaluation process requires sophisticated judgment in a complex crossroads of different classifications and concerns. This was illustrated by those interviewees who expressed that formalisation through criteria, even if developed within academia, was a threat to profound core values and holistic judgement, and would lead to a transfer of power and influence from the discipline to the local educational management, at the expense of academic values and expanded boundary-work within the committee.

Our results underline the fact that the norms and criteria for the doctorate are (still) carried, and changed, by disciplinary communities. This conclusion is particularly interesting in light of recent decades' efforts on supranational, national, and institutional levels to reform doctoral education. Although many authors describe a global convergence towards a more ordered and universal degree (Byrne, Jørgensen, and Loukkola 2013; Kehm 2020; Taylor, Kiley, and Holley 2021), our results show how important the disciplines, and their evaluation committees, are in shaping the doctorate. The legislative changes, accreditation schemes, guidelines, and frameworks, as well as the increased managerial control and monitoring that constitute the formalisation of doctoral education (Andres et al. 2015; Byrne, Jørgensen, and Loukkola 2013; Elmgren et al. 2016; Willison and O'regan 2007; Ruano-Borbalan 2022), operate partly in other strata than where we find the committee members and other leading proponents of the disciplines. As a consequence, the doctorate is as likely to be (re)shaped by concerns and circumstances within disciplinary research as by educational reforms and societal expectations.

Concluding Remarks

Within academia, it is common to hold the view that the real assessment of doctoral theses and candidates' competence takes place *before* the defence, through final seminars, peer-review processes of academic journals, evaluation processes, and constant formative actions by the research community. It might therefore appear as though the final evaluation of a doctoral candidate lacks importance. To the contrary, we argue that evaluation committees, through their far-reaching and complex boundary-work, are fundamental for local research cultures. When a candidate is

accepted for defence, supervisors and the local community know that they themselves will be assessed, and in that risk their own reputations. The boundary-work of examination committees reflects the historical function of the doctoral examination for upholding the symbolic value of the doctorate within disciplines and ultimately within academia itself.

That said, the disputatio could be argued to have lost some of its traditional functions of upholding norms and expectations. The evaluation and norm-setting of research have in many areas gravitated towards the peer review processes of academic journals and funding agencies, which may entail an externalisation of the locus of control. Furthermore, a doctorate is normally not enough to get a permanent academic position, meaning that the real gatekeeping is postponed to a later stage (and to other types of evaluation committees). If the issue of research quality is entrusted to academic journals and the question of academic competence is put on hold, perhaps the thesis and defence will be evaluated more against educational frameworks than disciplinary norms. Paradoxically, such changes within research and academia, rather than educational formalisation, might lead to an instrumentalisation of the evaluation, reducing the symbolic value of the doctorate.

Whilst these developments might be beneficial for society (and for some research projects), our findings urge us to consider what they could entail for academia. The formation of scholars, who can contribute to solving future challenges, as well as renewing academia, is a complex mission, which calls for a strong sense of disciplinary honour and responsibility. How the doctorate will develop, in light of present challenges and tensions, is to a large extent dependent on the boundary-work of the examination committees. The ubiquitously expressed seriousness of our committee members in relation to the evaluation and the extended boundary-work is promising in this respect. Their important and far-reaching boundary-work needs to be understood, valued, and discussed also outside the examination committees, among academics, educational managers, and higher education policymakers.

Declarations

Conflict of interest   The authors have no relevant financial or non-financial interests to disclose.

directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Aittola, Helena. 2008. Doctoral Education and Doctoral Theses — Changing Assessment Practices. In *Cultural Perspectives on Higher Education*, eds. Jussi Välimaa and Oili-Helena Ylijoki, 161–177. Dordrecht: Springer, Netherlands. https://doi.org/10.1007/978-1-4020-6604-7_11.

Alexander, David E., and Ian R. Davis. 2019. The PhD system under pressure: An examiner's viewpoint. *Quality Assurance in Education* 27(1): 2–12. https://doi.org/10.1108/QAE-04-2018-0033.

Andres, Lesley, Søren S. E. Bengtsen, Liliana del Pilar Gallego Castaño, Barbara Crossouard, Jeffrey M. Keefer, and Kirsi Pyhältö. 2015. Drivers and Interpretations of Doctoral Education Today: National Comparisons. *Frontline Learning Research* 3(3): 5–22. https://doi.org/10.14786/flr.v3i3.177.

Becher, Tony, and Paul R. Trowler. 2001. *Academic tribes and territories: Intellectual enquiry and the culture of disciplines*, 2nd ed. Non-fiction: Open University Press.

Bourke, Sid, and Allyson Holbrook. 2013. Examining PhD and research masters theses. *Assessment and Evaluation in Higher Education* 38(4): 407–416. https://doi.org/10.1080/02602938.2011.638738.

Bowker, Geoffrey C., and Susan Leigh Star. 1999. *Sorting things out: Classification and its consequences.Inside technology, 99–1592600-6*. Cambridge, MA: MIT Press.

Byrne, Joanne, Thomas Jørgensen, and Tia Loukkola. 2013. *Quality Assurance in Doctoral Education – results of the ARDE project*. The European University Association.

Carter, Bernie, and Karen Whittaker. 2009. Examining the British PhD viva: Opening new doors or scarring for life? *Contemporary Nurse* 32(1–2): 169–178. https://doi.org/10.5172/conu.32.1-2.169.

Carter, Susan. 2008. Examining the doctoral thesis: A discussion. *Innovations in Education and Teaching International* 45(4): 365–374. https://doi.org/10.1080/14703290802377208.

Chang, Ku-ming. 2004. From Oral Disputation to Written Text: The Transformation of the Dissertation in Early Modern Europe. In *History of Universities: Volume XIX/2*, ed. Mordechai Feingold. Oxford University Press.

Chen, Shuhua. 2014. Balancing knowing and not-knowing: An exploration of doctoral candidates' performance of researcher selves in the dissertation defence. *Assessment and Evaluation in Higher Education* 39(3): 364–379. https://doi.org/10.1080/02602938.2013.834876.

Elmgren, Maja, Eva Forsberg, Åsa Lindberg-Sand, and Anders Sonesson. 2016. *The formation of doctoral education*. Lund: Lund University.

Etzkowitz, Henry. 2003. Research groups as 'quasi-firms': The invention of the entrepreneurial university. *Research Policy* 32(1): 109–121. https://doi.org/10.1016/S0048-7333(02)00009-4.

Fochler, Maximilian. 2016. Variants of Epistemic Capitalism: Knowledge Production and the Accumulation of Worth in Commercial Biotechnology and the Academic Life Sciences. *Science, Technology, & Human Values* 41(5): 922–948. https://doi.org/10.1177/0162243916652224.

Fochler, Maximilian, Ulrike Felt, and Ruth Müller. 2016. Unsustainable Growth, Hyper-Competition, and Worth in Life Science Research: Narrowing Evaluative Repertoires in Doctoral and Postdoctoral Scientists' Work and Lives. *Minerva* 54(2): 175–200. https://doi.org/10.1007/s11024-016-9292-y.

Gieryn, Thomas F. 1983. Boundary-Work and the Demarcation of Science from Non-Science: Strains and Interests in Professional Ideologies of Scientists. *American Sociological Review* 48(6): 781–795.

Golde, Chris M. 2006. Preparing Stewards of the Discipline. In *Envisioning the Future of Doctoral Education: Preparing Stewards of the Discipline - Carnegie Essays on the Doctorate*, ed. Chris M. Golde and George E. Walker, 3–20. Hoboken, NJ: Jossey-Bass.

Grabbe, Lester L. 2003. The trials of being a PhD external examiner. *Quality Assurance in Education* 11(2): 128–133. https://doi.org/10.1108/09684880310471560.

Hamann, Julian. 2019. The making of professors: Assessment and recognition in academic recruitment. *Social Studies of Science* 49(6): 919–941. https://doi.org/10.1177/0306312719880017.

Holbrook, Allyson, Sid Bourke, and Hedy Fairbairn. 2015. Examiner reference to theory in PhD theses. *Innovations in Education and Teaching International* 52(1): 75–85. https://doi.org/10.1080/14703297.2014.981842.

Holbrook, Allyson, Sid Bourke, Hedy Fairbairn, and Terry Lovat. 2007. Examiner comment on the literature review in Ph.D. theses. *Studies in Higher Education* 32(3): 337–356. https://doi.org/10.1080/03075070701346899.

Holbrook, Allyson, Sid Bourke, Terence Lovat, and Kerry Dally. 2004. Investigating PhD thesis examination reports. *International Journal of Educational Research* 41(2): 98–120. https://doi.org/10.1016/j.ijer.2005.04.008.

Holbrook, Allyson, Sid Bourke, Terry Lovat, and Hedy Fairbairn. 2008. Consistency and inconsistency in PhD thesis examination. *Australian Journal of Education (ACER Press)* 52(1): 36–48. https://doi.org/10.1177/000494410805200103.

Johnston, Sue. 1997. Examining the Examiners: An analysis of examiners' reports on doctoral theses. *Studies in Higher Education* 22(3): 333–347. https://doi.org/10.1080/03075079712331380936.

Kehm, Barbara M. 2020. Reforms of Doctoral Education in Europe and Diversification of Types. In *Structural and Institutional Transformations in Doctoral Education*, eds. Sónia Cardoso, Orlanda Tavares, Cristina Sin, Teresa Carvalho, In Issues in Higher Education, 85–104. Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-38046-5_4.

Kemp, Simon, and Courtney McGuigan. 2009. Do PhD examiners agree on their recommendations? *New Zealand Journal of Educational Studies* 44(1): 47–57.

Kyvik, Svein. 2014. Assessment procedures of Norwegian PhD theses as viewed by examiners from the USA, the UK and Sweden. *Assessment and Evaluation in Higher Education* 39(2): 140–153. https://doi.org/10.1080/02602938.2013.798395.

Kyvik, Svein, and Taran Thune. 2015. Assessing the quality of PhD dissertations. A survey of external committee members. *Assessment and Evaluation in Higher Education* 40(5): 768–782. https://doi.org/10.1080/02602938.2014.956283.

Lamont, Michèle. 2009. *How professors think: Inside the curious world of academic judgment*. Harvard University Press.

Lamont, Michèle. 2012. Toward a Comparative Sociology of Valuation and Evaluation. *Annual Review of Sociology* 38: 201–221. https://doi.org/10.1146/annurev-soc-070308-120022.

Lamont, Michèle, and Virág Molnár. 2002. The Study of Boundaries in the Social Sciences. *Annual Review of Sociology* 28: 167–195.

Lamont, Michèle, Sabrina Pendergrass, and Mark Pachucki. 2015. Symbolic Boundaries. In *International Encyclopedia of Social and Behavioral Sciences*, ed. James Wright, 850–855. Oxford: Elsevier.

Lovitts, Barbara E. 2007. *Making the implicit explicit: Creating performance expectations for the dissertation*. Stirling, Virginia: Stylus.

Mason, Shannon, and Margaret Merga. 2018. A Current View of the Thesis by Publication in the Humanities and Social Sciences. *International Journal of Doctoral Studies* 13: 139–154. https://doi.org/10.28945/3983.

McKenna, Sioux, Lynn Quinn, and Jo-Anne Vorster. 2018. Mapping the field of Higher Education Research using PhD examination reports. *Higher Education Research and Development* 37(3): 579–592. https://doi.org/10.1080/07294360.2018.1428178.

Morley, Louise, Diana Leonard, and Miriam David. 2002. Variations in vivas: Quality and equality in British PhD assessments. *Studies in Higher Education* 27(3): 263–273. https://doi.org/10.1080/03075070220000653.

Mullins, Gerry, and Margaret Kiley. 2002. "It's a PhD, not a Nobel prize": How experienced examiners assess research theses. *Studies in Higher Education* 27(4): 369–386. https://doi.org/10.1080/03075070220011507.

Powell, Stuart, and Howard Green. 2007. *The doctorate worldwid*. Maidenhead: McGraw Hill.

QF-EHEA. 2005. *A Framework for Qualifications in the European Higher Education Area*. Copenhagen: Bologna Working Group for Qualification Frameworks. European Higher Education Area.

Ruano-Borbalan, J.C. 2022. Doctoral education from its medieval foundations to today's globalisation and standardisation. *European Journal of Education* 57(3): 367–380. https://doi.org/10.1111/ejed.12522.

Sère, Bénédicte. 2020. Between disputatio and Polemics: Dialectics as Production of Knowledge in the Middle Ages. *Medieval Worlds* 12: 80–95. https://doi.org/10.1553/medievalworlds_no12_2020s80.

Sharmini, Sharon, Rachel Spronken-Smith, Clinton Golding, and Tony Harland. 2015. Assessing the doctoral thesis when it includes published work. *Assessment and Evaluation in Higher Education* 40(1): 89–102. https://doi.org/10.1080/02602938.2014.888535.

Sigl, Lisa. 2016. On the Tacit Governance of Research by Uncertainty: How Early Stage Researchers Contribute to the Governance of Life Science Research. *Science, Technology, & Human Values* 41(3): 347–374. https://doi.org/10.1177/0162243915599069.

Sonesson, Anders, Lena Stenson, and Gudrun Edgren. 2023. Research and education form competing activity systems in externally funded doctoral education. *Nordic Journal of Studies in Educational Policy* 9(2): 173–190. https://doi.org/10.1080/20020317.2023.2222440.

Star, Susan Leigh, and James R. Griesemer. 1989. Institutional Ecology, "Translations" and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907–39. *Social Studies of Science* 19(3): 387–420.

Stigmar, Martin. 2019. Learning from reasons given for rejected doctorates: Drawing on some Swedish cases from 1984 to 2017. *Higher Education* 77(6): 1031–1045. https://doi.org/10.1007/s10734-018-0318-2.

Swedish Higher Education Ordinance. 1993. *SFS 1993:100 Högskoleförordningen*. Utbildningsdepartementet [Sweden's Ministry of Education and Research].

Taylor, Stan, Margaret Kiley, and Karri A. Holley, eds. 2021. *The Making of Doctoral Supervisors: International Case Studies of Practice*. Routledge.

Tinkler, Penny, and Carolyn Jackson. 2000. Examining the Doctorate: Institutional policy and the PhD examination process in Britain. *Studies in Higher Education* 25(2): 167–180. https://doi.org/10.1080/713696136.

Trafford, Vernon. 2003. Questions in doctoral vivas: Views from the inside. *Quality Assurance in Education*. 11: 114–122. https://doi.org/10.1108/09684880310471542.

Wenger, Etienne. 1998. *Communities of practice*. Cambridge: Cambridge University Press.

Willison, John, and Kerry O'Regan. 2007. Commonly known, commonly not known, totally unknown: A framework for students becoming researchers. *Higher Education Research & Development* 26(4): 393–409. https://doi.org/10.1080/07294360701658609.

Ylijoki, Oili-Helena. 2016. Projectification and conflicting temporalities in academic knowledge production. *Theory of Science* 38(1): 7–26. https://doi.org/10.46938/tv.2016.331.

Zerubavel, Eviatar. 1996. Lumping and Splitting: Notes on Social Classification. *Sociological Forum* 11(3): 421–433.