LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

INAUGURAL-DISSERTATION

Faculty and Teacher Emotions in Different Contexts:

An Investigation of Emotions for Grading and Emotions for Teaching in Online Compared to Face-to-Face Settings

Inaugural-Dissertation

zur Erlangung des Doktorgrades der Philosophie

an der Ludwig-Maximilians-Universität München

vorgelegt von

Carolin Schwab

aus

Rosenheim

München, 2022

Erstbetreuerin (Referentin):

Prof. Dr. Anne Frenzel

Department Psychologie

Ludwig-Maximilians-Universität München

Zweitgutachter (Koreferent):

Prof. Dr. Moritz Heene

Department Psychologie

Ludwig-Maximilians-Universität München

Drittes Mitglied der Prüfungskommission:

Prof. Dr. Robert Stupnisky

University of North Dakota

Abgabe: 25. Juni 2021

Tag der mündlichen Prüfung: 01. Oktober 2021

Acknowledgements

First of all, I would like to thank my first supervisor, Prof. Dr. Anne Frenzel, for her constant support, encouragement, and collaboration during the last years. Anne, without your offer to be my supervisor, your high expectations, and inspiring enthusiasm, I would not have come this far.

Special thanks also go to Prof. Dr. Robert Stupnisky, Jordan Jaeger, and Ph.D. Allison BrckaLorenz, who contributed substantially to the development and implementation of my second study on faculty grading emotions. Without you, I would not have been able to realize such a comprehensive cross-context and cross-country comparison project.

Next, I want to thank my second reviewer Prof. Dr. Moritz Heene. Moritz, thank you for switching to R and igniting the spark of R also in me. Re-attending the stats classes to learn R was definitely worth it and you can be proud to be such a good stats lecturer.

I also want to say “thank you” to my great colleagues at the Learning Sciences research unit who always made me feel welcome and part of a team. Sylvia, Julia, Sara, Daniel, Anton, Manuel: thanks for the great lunch breaks, chats over coffee (or chai latte), and of course the enriching scientific exchange. Special thanks go to Sara, Julia, and Anton, who provided feedback on parts of my dissertation and thereby helped me to improve it.

Of course, I cannot thank my family enough, because you provided me with the opportunity to find my way, finally study at an university, and always supported me in this endeavor as much as you could! Without you, I’d never have gotten to where I am now.

Last but not least I want to thank my partner Stefan who always supported me in every possible way during the past years and especially in the last months, weeks, and days of writing on this piece of work. Stefan, thanks for being there for me, being patient, and loving me despite the thesis writing!

Zusammenfassung

Die Forschung zu Lehrenden- und Dozierendenemotionen hat in den letzten Jahren zugenommen. Es gibt jedoch viele Aufgaben in beiden Berufen, die besonders in Bezug auf emotionale Erfahrungen zu wenig erforscht sind. Forschung zu den Emotionen von Lehrkräften und Dozierenden fehlt in Bezug auf a) eine Aufgabe außerhalb des Klassenzimmers, die erstaunlich wenig Beachtung von Seiten der Forschung erhalten hat, obwohl sie einen großen Teil der Zeit von Lehrkräften und Dozierenden in Anspruch nimmt und für die künftige Karriere von Schüler*innen und Studierenden sehr relevant ist, nämlich das Korrigieren von Studierendenarbeiten und b) eine Form der Lehre, die sich vor allem in der Hochschulbildung immer weiter verbreitet hat, nämlich die Online-Lehre. Um Einblicke in die Emotionen von Lehrkräften und Dozierenden in diesen vernachlässigten Forschungsbereichen zu gewinnen, wurden drei Studien durchgeführt, um 1) Emotionen von Lehrkräften, die durch einen aufgabeninhärenten Stimulus eines Aufsatzes (Qualität der Handschrift) ausgelöst werden, beim Korrigieren und deren Auswirkungen auf die Notengebung zu untersuchen und 2) Emotionen von Dozierenden beim Korrigieren im Vergleich zu Emotionen bei der Lehre und in der Forschung zu untersuchen sowie Prädiktoren von Emotionen beim Korrigieren aus einer Kontroll-Wert-Perspektive in zwei Ländern (USA und Deutschland) zu vergleichen und 3) Emotionen von Dozierenden in der synchronen Online-Lehre im Vergleich zur Präsenzlehre zu untersuchen.

In unserer ersten Studie sollte getestet werden, ob natürlich hervorgerufene Lehrendenemotionen das Korrigieren beeinflussen; eine wenig erforschte Aufgabe, die einen erheblichen Teil des Berufslebens von Lehrern einnimmt (OECD, 2014). Frühere Befunde zu den Auswirkungen von extern induzierter Stimmung auf das Korrigieren waren gemischt (Brackett et al., 2013; Townsend et al., 1989). Um diesen Umstand zu erklären, wurde in der ersten Studie versucht, diskrete positive (Freude) und negative (Ärger und Langeweile)

Aktivitätsemotionen in einer Korrektursituation mittels eines natürlich vorkommenden, materialinhärenten Stimulus (Qualität der Handschrift) zu induzieren, um die unterschiedlichen Effekte der drei diskreten Emotionen auf die Noten zu untersuchen. An unserem Experiment nahmen 73 Lehramtsstudierende (62 weiblich) teil, die zwei Aufsätze von ähnlicher inhaltlicher Qualität in unterschiedlicher Handschriftqualität korrigierten. Within-participant Mediationsanalysen bei zwei Bedingungen (Montoya & Hayes, 2017) zeigten, dass der Effekt von Handschrift auf Noten (0.32, *SE* = 0.15, CI [0.03, 0.61]; entspricht einem Drittel einer Notenstufe) durch Freude (0.26, *SE* = 0.09, CI [0.06, 0.43]) und Ärger (0.44, *SE* = 0.12, CI [0.21, 0.67]), nicht aber durch Langeweile (0.02, *SE* = 0.03, CI [−0.04, 0.10]) mediiert wurde. Das bedeutet, dass die bessere im Vergleich zur schlechteren Handschriftqualität ein höheres Maß an Freude und ein geringeres Maß an Ärger induzierte, was wiederum zur Vergabe von besseren bzw. schlechteren Noten führte. Die Ergebnisse deuten darauf hin, dass der Effekt von Ärger auf Noten stärker ist als der Effekt von Freude. Das bedeutet, dass Ärger eine größere Gefahr für die Reliabilität von Noten darstellt als Freude. Die Rolle von Langeweile bei der Korrektur muss weiter untersucht werden, da die Handschriftmanipulation keine Langeweile induzierte. Die differenzierten Ergebnisse rechtfertigen weitere Forschung, um die spezifischen Prädiktoren und Effekte verschiedener diskreter Emotionen zu identifizieren, die in Korrektursituationen natürlich auftreten. Schließlich sollten Wege gefunden werden, um das Korrigieren zu einer weniger aversiven Aufgabe zu machen. Auf diese Weise könnten negative Folgen für das Wohlbefinden der Lehrkräfte einerseits und für die künftige Karriere der Schüler*innen durch verzerrte Noten andererseits verhindert werden. Insgesamt ist Korrigieren nicht nur eine relevante und wenig erforschte Aufgabe des Lehrenden-, sondern auch des Dozierendenberufs. Daher wollten wir in einem nächsten Schritt die Emotionen bei der Benotung in einer Dozierendenstichprobe genauer untersuchen.

Die Forschung zu Emotionen von Dozierenden ist spärlich, insbesondere Studien zu Emotionen von Dozierenden bei unterschiedlichen Aufgaben ist stark limitiert. Frühere Studien betrachteten hauptsächlich Emotionen während der Lehre (vgl. Stupnisky et al., 2019b), neuere Arbeiten begannen, emotionale Erfahrungen zwischen den Kontexten Lehre und Forschung zu vergleichen (z.B. Stupnisky et al., 2016). Korrigieren wurde jedoch bisher nicht als eigenständige und bedeutsame Aufgabe untersucht. Daher zielte die zweite Studie darauf ab, sechs diskrete Emotionen (Freude, Stolz, Langeweile, Angst, Ärger, Frustration) zu untersuchen, die Dozierende während des Korrigierend erleben könnten. Studie 2a verglich die Emotionen von Dozierenden beim Korrigieren mit Emotionen in Forschung und Lehre (US-Stichprobe, n = 1226). Die Mittelwertvergleiche zeigten, dass das Korrigieren im Allgemeinen weniger positive (vor allem weniger Freude, aber auch Stolz) und mehr negative Emotionen (vor allem mehr Langeweile, aber auch Frustration und in gewissem Maße Ärger) auslöste als Forschung und Lehre. Eine Ausnahme bildete Angst, die beim Korrigieren weniger häufig erlebt wurde als bei den beiden anderen Aufgaben, was möglicherweise auf das Fehlen direkter negativer Konsequenzen bei "Misserfolg" zurückzuführen ist. Studie 2b untersuchte die Emotionen von Dozierenden beim Korrigieren aus Sicht der Kontroll-Wert-Theorie, indem emotionsspezifische Bewertungsmuster in zwei Ländern (USA, n = 245 und Deutschland, n = 201) identifiziert wurden. Die länderübergreifenden Mittelwertvergleiche zeigten, dass Dozierende aus den USA beim Korrigieren im Allgemeinen ein höheres Maß an Kontrolle, Kompetenz und positivem Wert erlebten als deutsche Dozierende, die beiden Gruppen aber ein ähnliches Maß an negativem Wert (Kosten) beim Korrigieren wahrnahmen. Darüber hinaus berichteten Dozierende aus den USA wesentlich häufiger Stolz und Angst und weniger häufig Ärger, jedoch ähnlich häufig Freude, Langeweile und Frustration, im Vergleich zu den deutschen Dozierenden. Die länderspezifischen Unterschiede sind vermutlich auf die unterschiedlichen Umstände, unter

denen das Korrigieren in den Stichproben der beiden Länder stattfand, zurückzuführen. In Bezug auf mögliche Prädiktoren von Emotionen beim Korrigieren von Dozierenden zeigten multiple lineare Regressionen, dass der wichtigste Prädiktor für Emotionen beim Korrigieren in beiden Stichproben die Kosten (negativer Wert), also das Ausmaß, in dem Dozierende das Korrigieren als undankbare Aufgabe wahrnahmen, die sie von sinnvolleren Aufgaben abhielt, waren. Die prädiktiven Muster variierten zwischen Emotionen und Ländern. In der deutschen Stichprobe war die diagnostische Kompetenz ein weiterer sehr wichtiger Prädiktor über alle Emotionen hinweg, der konsistent alle positiven und negativen Emotionen in der erwarteten Richtung vorhersagte. In der US-Stichprobe war sozialer Wert eine weitere wichtige Wertedimension über den negativen Wert hinaus, die vier Emotionen (Stolz, Angst, Ärger, Frustration) in der erwarteten Weise beeinflusste, in der deutschen Stichprobe jedoch nur Angst. Gründe für die unterschiedlichen Muster zwischen dem US-amerikanischen und dem deutschen Kontext wurden diskutiert. Insgesamt weisen die Ergebnisse darauf hin, dass Universitäten versuchen sollten, die Umstände beim Korrigieren zu verbessern, um das emotionale Erleben der Dozierenden zu optimieren. Dies könnte möglicherweise durch die Unterstützung der Entwicklung von Korrekturkompetenzen oder eine substanziellere Anerkennung des Korrigierens erfolgen. Diese Studie trug zum Verständnis von Emotionen von Dozierenden beim Korrigieren bei, aber sie erfasste die Emotionen in der Lehre nur im Rahmen eines Kontextvergleichs (im Vergleich zu Korrektur und Forschung). Da die Lehre ein integraler Bestandteil der Aufgaben von Dozierenden ist, sind die dabei erlebten Emotionen von Beduetung. Und obwohl die Emotionen von Dozierenden in der Lehre vergleichsweise gut erforscht sind, gibt es eine Form der Lehre, die bisher in Bezug auf Emotionen wenig untersucht wurde: die Online-Lehre. Der Ausbruch der COVID-19 Pandemie brachte weitreichende Aufmerksamkeit für diese konstant wichtiger werdende Form der Hochschulbildung.

Daher haben wir versucht, die Frage zu beantworten, wie sich die durch die COVID-19 Pandemie bedingte Verlagerung des Lehrens und Lernens an Hochschulen in online-gestützte Umgebungen auf die Unterrichtserfahrungen von Dozierenden auswirkte. Dafür bezogen wir Daten aus einer Studie von Daumiller et al. (2019) und replizierten das Studiendesign, um die Erfahrungen von Lehrkräften die vor der Pandemie Präsenzveranstaltungen unterrichteten (Stichprobe 1) mit den Erfahrungen von Lehrkräften, die in Zeiten der Pandemie synchrone Online-Kurse unterrichteten (Stichprobe 2; Daten wurden von den Autoren während der Pandemie erhoben), vergleichen zu können. Die Dozierenden beantworteten (a) einen Basisfragebogen und (b) mehrere sitzungsspezifische Tagebücher (nach 1-10 Sitzungen). Analysen des Basisfragebogens ergaben, dass sich die Dozierenden in Stichprobe 1 (n = 101) und Stichprobe 2 (n = 71) in ihrer wahrgenommenen Autonomie, Kompetenz, Verbundenheit und Selbstwirksamkeit vor der Pandemie sowie in ihrem Stress bei der Arbeit zum Zeitpunkt der Datenerhebung nicht unterschieden. Gemessen an den eigenen Vorerfahrungen nahmen die Lehrkräfte ihre Verbundenheit (d = 1.34, BF = $1.2e^{12}$), die Zufriedenheit mit der Lehre (d = 0.32, BF = 2.18) und die positive Valenz der Lehre (d = 0.36, BF = 5.13) während der Online-Lehre im Vergleich zur Präsenzlehre als reduziert wahr, nicht aber ihr Autonomie- und Kompetenzerleben. Analysen der aggregierten sitzungsspezifischen Tagebuchdaten ergaben, dass in der Online-Lehre die Befriedigung der Grundbedürfnisse nach Autonomie (d = 0.56, BF = 56.10), Kompetenz (d = 0.51, BF = 22.40) und Verbundenheit (d = 0.61, BF = 184.98) geringer und das emotionale Erleben ungünstiger (weniger Freude, d = 0.68, BF = 823.03; mehr Ärger, d = 0.56, BF = 59.85; tendenziell mehr Scham, d = 0.39, BF = 3.05; ähnliche Werte für Stolz, Langeweile und Angst) war als im Präsenzunterricht. Je mehr die digitale Lehrumgebung Quasi-Live-Erfahrungen ermöglichte (d.h. je mehr Studierende durch die Kamera zu sehen waren), desto mehr fühlten sich die Dozierenden tendenziell mit ihren Studierenden verbunden (β = .26, p =

.03, BF = 2.04). Die Ergebnisse implizieren, dass es wichtig ist, digitale Lehrkompetenzen zu fördern, um die Befriedigung der Bedürfnisse nach Autonomie und Kompetenz zu verbessern und digitale Lehr-Lern-Umgebungen zu schaffen, die den Aufbau von Beziehungen zu den Studierenden ermöglichen, um die emotionalen Erfahrungen von Dozierenden und Studierenden Lehren und Lernen im virtuellen Raum zu optimieren. Die Präsenzlehre bleibt ein wichtiges Element der universitären Ausbildung, das nach der COVID-Krise durch Online-Angebote ergänzt, aber nicht vollständig ersetzt werden kann.

Insgesamt kommen wir zu dem Schluss, dass die Emotionen von Lehrkräften und Dozierenden und ihre Prädiktoren und Auswirkungen in bestimmten Kontexten noch zu wenig erforscht sind und eingehender untersucht werden müssen, um schließlich das emotionale Erleben von Lehrkräften und Dozierenden bei weniger angenehmen Aufgaben ihres Berufs zu verbessern, um in der Folge das Wohlbefinden zu steigern und die hohen Burnout-Raten zu reduzieren.

Summary

Research on teacher and faculty emotions increased over the last years. There are, however, many tasks in both professions that are under-researched in general and with respect to emotional experiences in particular. Research on teacher and faculty emotions is missing with respect to a) one task outside the classroom that received surprisingly little research attention, although it takes up a good portion of teachers' and faculty members' time and is highly relevant for students' future careers, namely grading student work, and b) one form of teaching that started to become increasingly wide-spread especially in higher education, namely online teaching. To gain some insights into teacher and faculty emotions in these neglected research areas, we conducted three studies to investigate 1) teacher emotions for grading as elicited by a task-inherent cue of an essay (handwriting quality) and the emotions' effects on grades, 2) faculty emotions for grading as compared to emotions for teaching and research as well as antecedents of faculty grading emotions from a control-value perspective in two countries (U.S. and Germany), and 3) faculty emotions for online teaching in synchronous lessons as compared to face-to-face teaching.

In our first study we aimed to test if naturally elicited teacher emotions influence grading, an under-researched task that takes up a considerable portion of teachers' professional lives (OECD, 2014). Previous research showed that the effects of externally induced mood on grades were mixed (Brackett et al., 2013; Townsend et al., 1989). To shed light on this issue, the present study sought to induce discrete positive (enjoyment) and negative (anger and boredom) activity emotions in a grading situation by means of a naturally-occurring, material inherent cue (handwriting quality) to explore the differentiated effects of the three discrete emotions on grades. Our experiment involved 73 student teachers (62 female) grading two essays of similar content quality in varied handwriting quality. Two-condition within-participant mediation analyses (Montoya & Hayes, 2017) showed that the

effect of handwriting on grades (0.32, *SE* = 0.15, CI [0.03, 0.61]; equals one third of a letter grade) was mediated by enjoyment (0.26, *SE* = 0.09, CI [0.06, 0.43]) and anger (0.44, *SE* = 0.12, CI [0.21, 0.67]), but not by boredom (0.02, *SE* = 0.03, CI [−0.04, 0.10]). That is, the better compared to the worse handwriting quality induced higher levels of enjoyment and lower levels of anger, which in turn led to the assignment of better and worse grades, respectively. Results hint in the direction that the effect of anger on grades is stronger than the effect of enjoyment on grades, which means that anger may be a more severe threat to grading reliability than enjoyment. The role of boredom in grading remains to be explored, because the handwriting manipulation failed to induce boredom. The differentiated results warrant further research to identify the specific antecedents and effects of various discrete emotions that naturally emerge in grading situations to find ways to make grading a less aversive task. Thereby, negative consequences for teachers' well-being on the one hand and students' future careers through biased grades on the other, may be prevented. Grading is not only a relevant and under-researched task of teachers', but also of faculty members' profession. Therefore, we aimed to investigate grading emotions in more detail in a faculty sample as a next step.

Research on faculty emotions is scarce, and specifically investigations of faculty members' emotions during specific tasks. Previous research mainly considered emotions while teaching (Stupnisky et al., 2019b), more recent work started to compare emotional experiences between the contexts of teaching and research (e.g., Stupnisky et al., 2016). Grading as a distinct and influential task has not been considered to date, though. Therefore, the second study aimed to investigate six discrete emotions (enjoyment, pride, boredom, anxiety, anger, frustration) faculty may experience during grading. Study 2a compared faculty emotions for grading to emotions for research and teaching (U.S. sample, *n* = 1,226). Mean comparisons showed that grading generally elicited less positive (especially enjoyment,

but also pride) and more negative emotions (especially boredom, but also frustration and to some extent anger) than research and teaching. One exception was anxiety, which was experienced less frequently in grading than in the other two tasks, possibly due to a lack of direct negative consequences when "failing". Study 2b further examined faculty emotions for grading through the lens of control-value theory by identifying emotion-specific appraisal patterns in two countries (U.S., $n = 245$ and Germany, $n = 201$). The cross-country mean level comparisons showed that with respect to grading, U.S. faculty generally experienced higher levels of control, competence, and positive value than German faculty, but perceived similar levels of negative value (cost) with respect to grading. Furthermore, U.S. faculty reported substantially higher frequencies of pride and anxiety and lower frequencies of anger, but similar frequencies of enjoyment, boredom, and frustration, compared to German faculty. The cross-country differences were probably due to the differing circumstances under which grading occurred in the two countries' samples. Touching on possible antecedents of faculty grading emotions, multiple linear regressions revealed that the most important predictor for grading emotions across both samples was cost (negative value), in terms of the extent to which faculty perceived grading as a thankless task that kept them away from more meaningful tasks. Predictive patterns varied across emotions and countries. In the German sample, diagnostic competence was another highly important predictor across all emotions, which consistently predicted all positive and negative emotions in the expected direction. In the U.S. sample, social value was another important value dimension above and beyond negative value that influenced four emotions (pride, anxiety, anger, frustration) in the expected way, but in the German sample it influenced anxiety only. Reasons for the differing patterns between the U.S. and German context were discussed. Overall, results imply that universities should aim to improve the circumstances of grading to optimize the emotional experiences of faculty, perhaps by supporting the development of grading skills or more

substantive acknowledgement of the task of grading. The previous study contributed to understanding faculty grading emotions, but it only tapped at teaching emotions within a context-comparison (as compared to grading and research). As teaching is an integral part of faculty members' duties, the emotions experienced therein warrant research. And although faculty emotions in teaching have been comparably well-researched, there is one form of teaching that received little research attention with respect to emotions thus far: online teaching. The onset of the COVID-19 pandemic, however, brought immediate attention to this constantly-growing form of higher education.

Therefore, we tried to answer the question how moving higher education teaching and learning to online supported environments due to the COVID-19 pandemic affected the teaching experiences of university faculty. To this end, we obtained data from a study by Daumiller et al. (2019) and replicated the study design to compare the experiences of faculty teaching face-to-face classes before the pandemic (Sample 1) and faculty teaching synchronous online classes in a time of pandemic (Sample 2; data collected by the authors during the pandemic). Participants answered (a) a basic questionnaire, and (b) several session-specific diaries (1-10 sessions). Analyses of the basic questionnaire revealed that Sample 1 (n = 101) and Sample 2 (n = 71) were comparable in their perceived autonomy, competence, relatedness, and self-efficacy before the pandemic as well as in their stress at work during the time of data collection. When judged against their own prior experiences, faculty perceived their relatedness (d = 1.34, BF = 1.2e12), teaching satisfaction (d = 0.32, BF = 2.18), and positive valence of teaching (d = 0.36, BF = 5.13) to be reduced during online teaching compared to face-to-face teaching, but not their autonomy and competence. Analyses of the aggregated session-specific diary data revealed that when teaching online in a time of pandemic, the satisfaction of the basic needs for autonomy (d = 0.56, BF = 56.10), competence (d = 0.51, BF = 22.40), and relatedness (d = 0.61, BF = 184.98) was poorer and

emotional experiences less favorable (fewer enjoyment, $d = 0.68$, BF = 823.03; more anger, $d = 0.56$, BF = 59.85; tendency towards more shame, $d = 0.39$, BF = 3.05; similar levels of pride, boredom, and anxiety) compared to face-to-face teaching. The more the digital teaching environment allowed for quasi-live experiences (i.e., seeing more students through their cameras), the more faculty tended to feel related to their students ($\beta = .26$, $p = .03$, BF = 2.04). The results imply that it is important to foster digital teaching skills to enhance the satisfaction of the needs for autonomy and competence and to create digital environments that allow for building relationships with students in order to optimize faculty members' and students' emotional experiences when teaching and learning online. Nevertheless, face-to-face teaching is an important element of university education, which can be complemented, but not fully replaced by online offers after the COVID-crisis.

Overall, we conclude that teacher and faculty emotions and their antecedents and effects in certain contexts are still under-researched and need to be examined in more depth to eventually improve the emotional experiences of teachers and faculty in less pleasant tasks of their professions to subsequently increase well-being and reduce the high burnout rates.

Table of Content

1. Introduction

"Everyone who remembers his own education remembers teachers, not methods, and techniques. The teacher is the heart of the educational system"

– Sidney Hook –

Education is a central pillar in any society. It contributes to shaping the values, beliefs, and attitudes of future generations from elementary school through universities. And although the role of educational institutions and the influence they unfold on educational processes may not be underestimated, students' educational experiences are mainly influenced by their teachers and the learning environments they create. And although it is beneficial for students to attend a good school to excel, students' actual success is probably more strongly determined by their teachers' facilitation of learning and development and their willingness to learn and improve. Teachers managing to initiate classroom interactions that foster relationships, student learning, and spark the innate curiosity of students, are key to any successful learning environment. Because after all, learning should not be reduced to acquiring knowledge that someone tries to convey. Learning should enable individuals to independently acquire and develop their own knowledge, skills, and convictions. In today's world with its richness of information, in which answers are often only one click away, the competence to retrieve, sort, and process new information may not be as important as to reflect critically upon the truthfulness and sources of the obtained information to make an informed decision about ones' own point of view. But to reach this point, "every child deserves a champion: an adult who will never give up on them, who understands the power of connection and insists that they become the best that they can possibly be" (Pierson, 2013, 6:59). In the long run, the best teachers are probably those who make themselves superfluous

because their students learned to learn for and by themselves and to stand in for their convictions.

Conceptions of education that view education as an at least to some extent holistic process already imply that there is much more to the profession of a teacher than "simply" teaching. And indeed, teachers are not only expected to teach content to students according to the curriculum, they are also expected to educate them to become responsible citizens according to the hidden curriculum, to be an important contact person while keeping a professional distance, to assess student performance objectively while considering individual students' performance levels and development, to apply student-centered teaching that activates student learning while often having to test pre-determined outcomes, to incorporate technology-assisted teaching in their classrooms while often not having been trained to do so, to plan and realize all their lessons thoroughly while taking on further responsibilities for extra-curricular school activities (e.g., theatre, sports, science groups), and so on and so forth. This leaves teachers in a tension field between their own ideas, ideals, needs, and capacities on the one hand, and expectations set on them by students, parents, colleagues, institutions, policies, and politics on the other.

What is often missing from this very colorful bouquet of expectations and requirements placed on teachers is that teachers are humans with limited resources, their own needs and desires. Especially political discussions about what teachers should or should not do often lack the component of what is realistic to be asked of teachers. One example was the rapid shift of nearly all teaching in all educational institutions from elementary school through university to distance education during the onset of the COVID-19 pandemic. Most political discussions (at least in Germany) were seemingly concerned with the learning of school students and covered questions, such as "how can student learning be maintained during lockdowns and contact restriction? What do teachers need to do to achieve this goal?

How can be made up for the missed time in class?". But although the COVID-19 pandemic and especially the declared lockdowns and contact restrictions were a burden to everyone, including teachers, hardly any discussions focused on questions such as "how can we support teachers in the rapid transition to distance education to maintain learning? Which technical equipment and methodological skills do they need to implement distance education? How can we help them to manage the extraordinarily increased workload?". Such questions were mainly left to teachers themselves who had to instantly reinvent teaching depending on their own and their students' technical equipment and skills, often in need to find individual solutions to provide students with learning materials; all too often through parents especially in lower grades. Taking into consideration that teachers had comparably high burnout rates even before the pandemic (Hakanen et al., 2006) and acknowledging the incredible challenge they were facing during the pandemic, it seems almost careless to not pay more attention to the people working in this highly important profession. Teaching is not limited to schools, however, and faculty[1] at universities were facing a similar situation. And although universities managed the transition to distance education comparably well by implementing online teaching, students and faculty alike were highly challenged by the rapid shift (Kanning & Ohlms, 2021). Politics, however, seem to largely have overlooked higher education in times of the COVID-19 pandemic. This is mirrored in the fact that bringing school children back into schools and to provide face-to-face teaching is highly prioritized and widely discussed, while higher education will likely be continued in online formats, probably assuming that higher education online teaching and learning are just as effective as face-to-face teaching and that it does not impair university students and faculty severely. The

[1] By faculty or faculty member we refer to all individuals teaching at a higher education institution irrespective of the held degree or exact position, thereby also including for instance doctoral students and external lecturers.

available studies, however, point in the direction that emergency online learning during the COVID-19 pandemic was a less positive experience than face-to-face learning for university students and the highly limited results about faculty point in a similar direction (Besser et al., 2020; Garris & Fleck, 2020; Kanning & Ohlms, 2021). Therefore, the question of returning to face-to-face teaching in higher education and the needs of university students and faculty should not be ignored. The present dissertation tries to contribute to this important topic by directly comparing the emotional experiences and the satisfaction of basic psychological needs of faculty who taught synchronous emergency online classes during the pandemic to faculty who taught face-to-face classes before the pandemic. To our knowledge, this is the first comparative approach to generate knowledge about the teaching situation during the COVID-19 pandemic in direct comparison to face-to-face teaching before the pandemic that is not based on retrospective comparative judgements.

The relative neglect of teachers' and faculty members' situation not only during the COVID-19 pandemic but also in general is worrisome to a certain degree as they are a key factor in maintaining a working educational system. The comparable inattention in political discussions to teachers and faculty as individuals with own needs is to some extent mirrored in educational research: An abundant body of research focusing on students is available, for instance how their learning, interest, and motivation can be fostered and how emotions influence their performance. Within this body of research, teachers (and faculty) were mainly considered from the perspective of how they can contribute to desirable student outcomes. And although it is definitely important to know how they can exert a positive influence on their students, it is not less important to change the perspective and to attend to teachers and faculty as persons whose well-being, job motivation, and satisfaction are important in and of themselves. To this end, investigating the emotions teachers and faculty experience during their workdays is important. And because there is more to being a teacher or faculty member

than teaching, it is necessary to pay attention to the emotional experiences triggered by work tasks outside the classroom. Therefore, the present dissertation tries to narrow the research gap on emotional experiences outside the classroom by taking a look at one time-consuming and unliked, yet for students' further careers highly influential task, namely grading. To better understand the role that emotions play in the context of grading, it is necessary to investigate different discrete emotions (e.g., enjoyment, anger, boredom), their antecedents (e.g., by means of a task-inherent emotion induction or assessing control and value appraisals and investigating their predictive power on different emotions), their effects on grades, and compare the emotions for grading to other tasks, such as teaching and research to obtain a comprehensive view. To our knowledge, these are some of the first quantitative approaches to explicitly investigate teacher and faculty emotions for grading. But before diving into grading emotions, a first step is to pay attention to the manifold activities teachers and faculty perform as part of their professions.

The Professions of Teachers and Faculty

Teachers and faculty have professions that encompass highly diverse tasks and activities that do not end when leaving a classroom. In addition to teaching classes, which is highly varied in and of itself, preparing and post-processing lessons, grading student work, and administrative responsibilities are some further tasks that need to be handled. In addition to these common tasks, teachers additionally need to be in touch with parents and contribute to school life for instance by offering extra-curricular activities, organizing field trips, or student exchange programs, while faculty additionally need to conduct research, another highly complex activity with many specific tasks. And although both professions are highly diverse, most research focused on one task: teaching. Because teaching is one of the most time-consuming and influential tasks that teachers and faculty perform, solid knowledge about it is unarguably important. Within the teaching context, most research focused on how

teachers or faculty can positively influence their students. For instance, research was typically interested in the effects of teaching methods, approaches to teaching, and teacher or faculty characteristics on student engagement and achievement. Fewer studies considered teachers with their emotional experiences as the individuals of interest, although their number increased in the last decades. Earlier studies often dealt with burnout among teachers, which became one of the most well-researched emotional phenomena (Frenzel, 2014) and is fairly high among teachers and faculty (García-Carmona et al., 2019; Watts & Robertson, 2011). And although research on teacher and faculty emotions more broadly gained momentum over the last years, the number of studies is still comparably limited, especially with respect to faculty emotions. Nevertheless, teacher and faculty emotions in and of themselves seem to have been identified as research need, not least because they contribute to the well-being and job motivation of teachers and faculty, which subsequently ensure a working educational system.

Emotions in Education

Emotions have been described as multicomponential episodes that arise as reaction to certain stimuli and which comprise affective, behavioral, cognitive, physiological, and motivational components (K. R. Scherer, 2005; Shuman & K. R. Scherer, 2014). Disagreement prevails whether appraisals of events are to be considered as a component or antecedent of emotions (Shuman & K. R. Scherer, 2014), whereby the present dissertation treats appraisals as antecedents of emotions, as suggested by control-value theory (Pekrun, 2018).

Control-value theory is a renowned theory in the context of education and proposes that specific combinations of subjective control appraisals (range from low to high) and value appraisals (may be positive or negative and range from low to high intensity) determine discrete emotions in achievement contexts. It claims that perceiving a situation as

(un-)desirable constitutes the valence of an emotion (positive or negative), whereas the extent of perceived controllability influences its quality (e.g., anxiety vs. anger). When individuals appraise high value, they typically experience all emotions more strongly, when they appraise high subjective control, they typically experience high levels of positive and low levels of negative emotions (Pekrun, 2006). For instance, if a student believes to be capable of mastering the content for an exam (high control) and considers the exam grade to be very important (high value), then it is assumed that the student experiences high levels of enjoyment when learning. If a student believes, however, to be incapable of sufficiently preparing for an exam (low to medium control) and considers the exam grade to be very important nevertheless (high value), then it is assumed that the student experiences high levels of anxiety with respect to the exam. Although the control-value theory of achievement emotions has been developed and mostly researched in students (Pekrun et al., 2007), it has been applied to faculty emotions in teaching and research (e.g., Stupnisky et al., 2019b). This is lending support for its wider applicability to teaching staff on the one hand and different contexts in higher education, on the other.

Another prominent theory in the context of education is self-determination theory (Deci & Ryan, 2004), which proposes that intrinsic motivation and well-being are influenced by the satisfaction of the basic psychological needs for autonomy, competence, and relatedness. Autonomy refers to having agency over one's actions, which ideally reflect one's self, competence to successfully interacting with and mastering one's environment, and relatedness to experiencing connectedness and mutual caring with others (Deci & Ryan, 2004). Due to the basic nature of the psychological needs, this theory has been widely applied in varied contexts and populations with respect to motivation and well-being. Lately, studies started to explore the presence of and relationships between basic need satisfaction and discrete emotions in teaching (Ebersold et al., 2019; Hagenauer et al., 2015; Hagenauer &

Volet, 2014; Klassen et al., 2012; Löfström & Nevgi, 2013; Russo et al., 2021) and (online) learning settings (Buhr et al., 2019; Butz et al., 2014; Butz & Stupnisky, 2016; Filak & Nicolini, 2018; Marchand & Gutierrez, 2012; Otter et al., 2013). The findings lend support for a wide applicability of self-determination theory in educational settings and for its connection with discrete emotions. Recently, a first attempt has been made to empirically connect the two theories by considering the satisfaction of the basic needs as antecedents of control and value appraisals and subsequent discrete emotions (Buhr et al., 2019), further supporting the idea that basic need satisfaction can be predictive of specific discrete emotions.

Taking into consideration the specific situational and appraisal antecedents of discrete emotions on the one hand and the distinct effects of discrete emotions on the other, it seems important to investigate discrete emotions rather than general positive and negative affect of teachers and faculty when engaging in different tasks of their professions. Acknowledging the important role of emotions in (higher) education (Pekrun, 2019) and their context-specificity (Frenzel et al., 2015, 2016), it seems necessary to extend current emotion research to still under-researched contexts in these varied professions. Such contexts may be important, time-consuming tasks outside the classroom, such as grading student work, or alternatives to face-to-face teaching, such as online teaching.

Teacher and Faculty Emotions in Different Contexts

What do and do we not know about the emotions that teachers and faculty experience in different contexts? With respect to research, a task that only faculty engage in but which is typically at least as time-consuming as teaching, faculty experienced a variety of emotions. They experienced positive emotions (e.g., enjoyment, pride) more strongly than negative emotions (e.g., boredom, anxiety, anger), but perceived research as emotionally less favorable than teaching because they experienced positive emotions less and negative

emotions more strongly than in teaching (Stupnisky et al., 2016, 2019a). Reasons may be that research is a less controllable (especially with respect to publications) and more important, yet competitive (monetary benefits at universities are often dependent on publication rates) endeavor than teaching. Generally, the insights into faculty emotions are scarce, and gaining more knowledge about the emotions experienced by faculty in this central activity is important to assess their contribution to overall well-being and job satisfaction. Although research is an important aspect to pay attention to, the present study focused more on teaching and teaching-related activities.

With respect to classroom teaching, teachers and faculty consistently reported that they experienced predominantly positive (e.g., enjoyment, pride) rather than negative emotions (e.g., boredom, anger, and frustration, anxiety) in teaching (e.g. Frenzel, 2014; Keller et al., 2014; Kordts-Freudinger, 2017; Postareff & Lindblom-Ylänne, 2011; Stupnisky et al., 2016, 2019a; Thies & Kordts-Freudinger, 2019; Trigwell, 2012). Regarding the antecedents of teacher and faculty emotions in teaching, control and value appraisals as well as perceived autonomy, competence, and relatedness contributed to experiencing different positive and negative emotions in teaching (Ebersold et al., 2019; Hagenauer et al., 2015; Hagenauer & Volet, 2014; Klassen et al., 2012; Löfström & Nevgi, 2013; Russo et al., 2021; Stupnisky et al., 2019a, 2019b). Overall, it is good news that teachers and faculty seemingly are fond of one of their main responsibilities. Under these circumstances teaching probably contributes to their motivation and well-being.

But what if the circumstances change drastically and teaching is not the same anymore as it used to be? During the COVID-19 pandemic, most teachers and faculty around the globe were forced to immediately switch from their seemingly appreciated face-to-face to distance teaching. Taking into account that the relationship with students is one of the factors that fostered positive teaching experiences (Klassen et al., 2012), the move to a distance

format lacking direct contact definitely held the potential to impair teaching and learning alike at all levels of higher education. Furthermore, the rapid shift did not allow for a lot of preparation (if any) and only few teachers, faculty, and students had prior experience in online teaching or learning. On top of these personal challenges, most institutions were not equipped to provide such extensive distance education offers. In Germany, for instance, only few schools had learning platforms and online meeting tools in place, not all students had the technical equipment to participate in online distance education, and only few teachers were trained to use such tools. Although universities had used some digital tools even before the pandemic, such as learning platforms to provide learning materials and recordings of select lectures to be watched from home, typically no software was available already that could have covered the massive number of courses that needed to be delivered online. Purchasing the necessary tools to offer online education was not only expensive but complicated even further by having to fulfill high data protection standards in many countries. And across all levels of education, not all students had access to digital devices and the internet to participate in online learning and not all teachers and faculty necessarily had a device they could work with from home to provide online teaching. Therefore, the institutional support for all parties may have left room for improvement in some cases especially during the onset of the pandemic when masses of challenges needed to be tackled. But not only institutions, but also teachers and faculty were facing many highly diverse challenges during these unprecedented times and taking into consideration how aversive the circumstances were, maintaining the provision of education at all has been a great accomplishment. Nevertheless, the pandemic pointed out that digitalization mostly has not reached the educational sector yet. Although it seems to be an ideal to have technology-assisted teaching in classrooms and online or hybrid classes in higher education to complement traditional face-to-face offers, Germany had to realize that it is far from using the potential that the use of digital tools in

education may unfold and opportunities to participate were often limited due to insufficient internet coverage especially in rural areas. This backlog in digitalization was partially mirrored in research on online education more generally: despite the steadily growing number of online offers in higher education over the last years, the insights into online education are still limited in comparison to traditional education, especially with respect to emotional experiences. Similar to the body of research on emotions in face-to-face settings, teacher and faculty emotion research is even scarcer than that about student emotions. But some studies suggested that online teaching is a less pleasant experience than face-to-face teaching because the transition is hard for most as they do not feel ready and online teaching seems to be associated with lower perceptions of competence, impaired relationships with students, and negative emotions (Downing & Dyment, 2013; Kanning & Ohlms, 2021; Meishar-Tal & Levenberg, 2021; Regan et al., 2012; R. Scherer et al., 2021). There is, however, no consistent and compelling body of research on this topic available yet. And given the chance that probably not all (mainly higher) education will move back to the state of affairs before the pandemic, it seems necessary to pay attention to this increasingly important form of education. And although it is unarguably important to learn how to implement effective online teaching, the emotions faculty experience may not be neglected if online teaching becomes an integral part of their work live in the future.

Because tasks outside the classroom and beside research take up another good portion of teachers' and faculty members' workdays, the effects of the emotions experienced while performing them cannot be neglected when being interested in getting a comprehensive impression of the emotions experienced by teachers and faculty at work. Research in this area is few and far between and often based solely on qualitative studies. The possibly most interesting task in teachers' and faculty members' work lives apart from teaching and research is grading, that is assessing student work and assigning a grade to it. Grading has

two sides to it. First, from a student perspective, grading is highly important because it determines their exam grades and thereby GPA, which is the gatekeeper to move forward in education (e.g., GPA-based selective processes when transitioning to different school tracks for secondary school or wanting to enroll in specific subjects at universities) and to good chances on the labor market. Second, from a teacher or faculty perspective, grading typically does not provide any direct benefits to them while at the same time being a typically repetitive and aversive task that is highly time-consuming, especially when grading written student work, such as essays and exams with open (long answer) questions. These circumstances place teachers in an unpleasant and stressful situation because they need to perform a task they do not like, but need to do so diligently and concentrated because objective, reliable, and valid grading is important to them and moreover necessary to assess student performance fairly. Deepening our knowledge on grading is necessary to reach two aims: first, assigning reliable grades that are not biased by factors unrelated to student performance and second, creating an emotionally less aversive grading situation for teachers and faculty to minimize negative effects on well-being. To this end, it seems important to find out which grading biases exist and how they can be mitigated and which factors make grading (especially written work) such an unpleasant task and how these can be positively influenced. Although there are many studies on different grading biases (Malouff & Thorsteinsson, 2016), empirical findings on how to reduce them are scarcer. With respect to grading emotions, the few available studies suggested that grading is an aversive and stressful task, and that especially the prospect of having to assign fail grades or to offer test-feedback sessions triggered anxiety, anger, frustration, sadness, and guilt (Babb & Corbett, 2016; Ilott & Murphy, 1997; Laybourn et al., 2019; Loh & Liew, 2016; Stough & Emmer, 1998). But grading did not only trigger emotions, emotions also influenced grades (Brackett et al., 2013; Townsend et al., 1989). And therefore, research on emotions for grading is at the intersection

of research on grading bias, effects of teacher and faculty emotions on student outcomes, and teacher and faculty emotions as outcome relevant in and of itself, and thereby important for students, teachers, and faculty alike.

But grading is not a uniform task that teachers and faculty need to perform in addition to teaching (and research), it is a task that can take incredibly different forms depending on the circumstances. Imagine the following situations to get an idea of how different grading may be. First example: a mathematics school teacher who holds multiple exams throughout the term in his class and needs to assign points based on the correctness of answers to specific mathematical problems. Apart from possibly accounting for multiple approaches to solving a problem and the question how to treat consequential errors, there is the possibility to differentiate a right from a wrong answer. Summing up the correct answers, grades are determined based on reached points or percentages and the teacher gets an impression of the overall performance level in class. Second example: a language school teacher or faculty member who holds multiple exams throughout the term in their class and needs to assess essays on open topics, such as discussing a provocative statement or pros and cons of a certain topic that has probably been treated in class in some form. In such a setting it is harder to differentiate right from wrong answers and the determination of quality is more subjective and therefore more challenging. The final grade may then be determined by using a grading rubric and summing up the assigned points on predefined criteria or by assigning an overall grade based on a holistic consideration. The essays may inform the teacher or faculty member about the performance level in class and if additional feedback is provided even foster student learning. Third example: a faculty member involved in grading centralized exams, such as the student teachers' state exams in Bavaria (a federal state in southern Germany) that determine the final grade for graduation. In this very special setting, the exam questions are formulated by faculty members from all Bavarian universities and sent to the ministry of education.

Faculty at the different universities often place different weights to specific exam topics and sometimes have different views on the same topics. The ministry then chooses the exam questions and all student teachers in Bavaria work on the same questions. The essays are then sent to faculty from a different university for grading who are not necessarily involved in teaching the preparatory courses for the state exams and not necessarily familiar with the topics they are supposed to grade. As there are typically no grading rubrics available from the authors of the questions, grading the essays turns into a highly subjective endeavor with differences in interpreting questions and therefore varying expectations between a huge number of faculty involved in grading, possibly different views on right and wrong answers to a specific topic, and in the end likely different grades on comparable student performance. Based on these three examples it becomes obvious that although grading is probably always an aversive task, there are situations in which grading may be more aversive than in others.

Drawing on control-value and self-determination theory reasoning it seems likely that grading situations differ with respect to control appraisals and the satisfaction of the need for autonomy (e.g., differences in the extent of determining the exam content and the grading process), control appraisals and the satisfaction of the need for competence (e.g., based on differences in familiarity with the topic, experience in grading), the satisfaction of the need for relatedness (e.g., differences in the closeness of the relationship with students), and value appraisals (e.g., differences in the importance of the closeness to students, perceived utility of the task, perceived interestingness of the task, and perceived costs). Based on these assumptions it seems reasonable that appraisals and need satisfaction in turn influence discrete emotions in grading, although grading is not a typical achievement context in the sense that the outcome is a direct measure of teachers' or faculty members' performance.

Overall, it can be stated that emotions of teachers and faculty are generally under-researched. Most available research focused on one specific task, namely classroom teaching.

Further research tapped at other tasks, such as online teaching, research, and grading, but the body of research on these important aspects of teachers' and faculty members' professions is highly limited and far from offering comprehensive insights. Therefore, the main aim of the present dissertation was to broaden the insights into some of those under-researched contexts of teacher and faculty emotions, namely grading and online teaching. To this end, three studies were conducted and reported herein. The first study tried to shed light on the question whether emotions that are elicited by the grading task itself influence grades, which are an important outcome for students because they influence their further educational trajectories and career opportunities. The second study explored grading emotions in more detail, but in a university faculty sample. The aims were to a) compare the emotional experiences of faculty in different tasks, namely teaching, research, and grading b) identify possible antecedents of grading emotions based on control-value theory reasoning c) compare grading emotions and their antecedents between the U.S. and Germany. The third study investigated discrete teaching emotions and need satisfaction of faculty teaching synchronous online classes during the pandemic compared to faculty teaching face-to-face classes before the pandemic using self-determination theory as theoretical framework.

2. Study 1

Discrete Emotions in Grading Situations: Differentiated Effects of Anger, Enjoyment, and Boredom on Grades

If you were asked to conjure up an image of a teacher, most likely the first picture that would come to your mind is that of a teacher standing in front of a class. However, teachers' jobs do not end once they leave the classroom. Manifold other tasks need to be accomplished in this emotionally demanding profession (Hargreaves, 1998). One of them is grading student work. And as anyone who has ever graded written work knows, laboring over pages and pages of students' essays can make you bite your nails, pull your hair, and in some cases, make you want to throw the essay into the trash. If you read work that is written well and clearly follows the requirements, however, you may enjoy grading. In other words, the act of grading itself elicits emotions – a notion mostly overlooked in teacher emotion research thus far.

Importance of Teacher Emotions

The emotions teachers experience have far-reaching consequences for teachers themselves and for their students. Teachers' emotions have been shown to not only influence teachers' own well-being (Frenzel & Stephens, 2017), but also to transfer to students (Frenzel et al., 2017; Frenzel, Goetz, Lüdtke, et al., 2009), and to have consequences for students' motivation and performance (Rodrigo-Ruiz, 2016). For instance, when math teachers experienced subject-related anxiety, their students showed reduced ability belief and subsequently decrements in math performance (Beilock et al., 2010). Considering such discrete emotions in teacher emotion research is necessary because teacher emotions were better represented by discrete emotions than by combining emotions into positive and negative affect or a single factor (Frenzel et al., 2016). Research on teacher emotions over the last decade has provided insights into teachers' emotional lives while in the classroom (e.g.,

Frenzel, Goetz, Stephens, et al., 2009). Missing from this picture, however, are insights into teachers' emotional lives while engaging in tasks outside the classroom.

Considering that teachers spend roughly half of their weekly working time on teaching (i.e., in the classroom) and the other half on teaching-related activities outside the classroom (Bauer et al., 2007; Philipp & Kunter, 2013), it is obvious that these activities are far from negligible. After preparing lessons (about 20 - 30%), the most time-consuming duty with probably the most far-reaching consequences was grading student work (almost 20%; Bauer et al., 2007; OECD, 2014; Philipp & Kunter, 2013) — an activity that is highly influential for students' futures, because grades were shown to be an important predictor of graduation from high school, performance in and graduation from university (Allensworth & Clark, 2020; Bowers, 2010; Westrick et al., 2015). Because grades are such an important student outcome it is necessary to ensure that grading is fair, reliable, and valid by reducing the influence of non-achievement factors that bias grading (Malouff & Thorsteinsson, 2016). Grading was found to be influenced by expectation-driven factors, such as teachers' impressions of their students' ability, improvement, effort, and classroom participation (e.g., Cizek et al., 1995; Duncan & Noonan, 2007; see also Brookhart et al., 2016), contextual factors, such as order of essay presentation (Hughes et al., 1980; Hughes & Keeling, 1984) and handwriting quality (Briggs, 1970, 1980; Greifeneder et al., 2010, 2012; James, 1927; Markham, 1976), and teacher characteristics, such as experience in grading (Eames & Loewenthal, 1990; Markham, 1976) and mood (Brackett et al., 2013; Townsend et al., 1989). Most relevant for the present study was work which showed that mood influenced grades in emotion-congruent ways (Brackett et al., 2013) and that individuals assigned better grades to well- than to ill-legible work (Greifeneder et al., 2010, 2012), supposedly because the easier and harder processing were affectively positively and negatively connotated, respectively (Reber & Greifeneder, 2017; Winkielman et al., 2003). Although grading is such an

important task that typically has a negative (emotional) connotation (Steinberg, 2008), especially quantitative research on teacher emotions for grading is few and far between.

Grading Emotions

Overall, authors who studied the role of emotions in grading agreed that it is an aversive task that is associated with negative rather than positive emotions. In interviews, teachers described grading written student work as a highly stressful endeavor that came along with many emotional burdens (Loh & Liew, 2016). Similarly, teachers named grading student work as one of the tasks within their profession they procrastinated most often due to its task aversiveness (Laybourn et al., 2019). Furthermore, they reported strong negative emotions, such as anxiety, anger, frustration, sadness, and guilt when asked to think of situations in which they needed to give a fail grade (Babb & Corbett, 2016; Ilott & Murphy, 1997) and when they offered test-feedback sessions (Stough & Emmer, 1998). In line with attribution theory, sympathy and anger mediated the relationship between teachers' controllability appraisals of student's failure and retributive goals when giving feedback in a vignette study: the more participants thought that the reason for a student's failure was controllable, the more they held the student responsible and the more anger and the less sympathy they experienced, which led to the use of more and less retributive goals when providing feedback, respectively (Reyna & Weiner, 2001). This hints at the potential of discrete emotions to not only bias feedback negatively, but grades in general.

Experimental research on emotions in the grading context investigated the effects of mood on grades and revealed mixed results. Townsend and colleagues (1989) asked university students to watch a film as means of mood induction and to subsequently grade multiple short essays of 13-14 year old students on their hopes and aspirations in the next decade. Overall, the results showed that the mood induction worked but wore off quite quickly especially in the negative mood condition and that there were only little effects of

mood on grades. One of the few significant findings indicated that students in the negative mood condition unexpectedly assigned better grades to the first essay than those in the positive mood condition (Townsend et al., 1989). Using a similar design, Brackett and colleagues (2013) induced positive or negative mood via an autobiographical recall procedure and thereafter asked participants to evaluate a middle school student's narrative essay in two identical studies with varied participants (undergraduate students and in-service teachers). Participants judged the same essay using four specific criteria (creativity, spelling/punctuation, composition structure, and vocabulary) and one global criterion (overall performance). In the undergraduate sample, differences between the positive and negative mood induction group emerged for creativity only, whereas in the teacher sample differences emerged for creativity, spelling/punctuation, and overall performance. Results indicated that participants in a positive mood tended to judge the essay more favorably than those in a negative mood and that the effects of mood on grades were more pronounced in the teacher than in the student sample, despite the longer experience in grading student work in the teacher sample. The effects of the external mood induction wore off over time in these studies as well (Brackett et al., 2013).

Although the two studies (Brackett et al., 2013; Townsend et al., 1989) provided first insights into the emotional experiences when grading, eliciting emotions by means of a mood induction held some drawbacks: On the one hand, the mood induction wore off over time, because it was a cue external to the grading situation, unlike a task-inherent cue that would occur naturally and constantly in the grading situation itself. On the other hand, it captured only general mood, which did not allow to disentangle emotion-specific effects (watching a film about Auschwitz probably elicited other negative emotions than writing about a negative moment in life). It is important, however, to consider discrete emotions due to their specific antecedents and effects (Pekrun, 2006). For the present study we chose to focus on three

discrete activity emotions that were frequently experienced by teachers (Frenzel, 2014) and that may occur in grading situations, namely enjoyment, anger, and boredom. Activity emotions are claimed to be triggered by appraisals of the activity itself, not of its outcomes (Pekrun, 2006, 2018). From a control-value perspective, enjoyment typically arises when an individual is capable of dealing with a positively valued situation (i.e., is able to master a pleasant situation), anger when an individual is capable of dealing with a negatively valued situation (i.e., is able — but does not want — to master an aversive situation), and boredom when a situation does not hold any value for the individual, whereby an individual may feel either over- or underchallenged by the task at hand (Pekrun, 2006; Pekrun et al., 2010). In general, positive activating emotions (e.g., enjoyment) focusing on the task to be accomplished have been reasoned to elicit positive outcomes, negative activating emotions (e.g., anger) to elicit negative outcomes unless they trigger failure-avoidance (e.g., anxiety), and negative deactivating emotions (e.g. boredom) to elicit negative outcomes (Pekrun, 2018), which may translate into differentiated effects of emotions on grades.

The Present Study

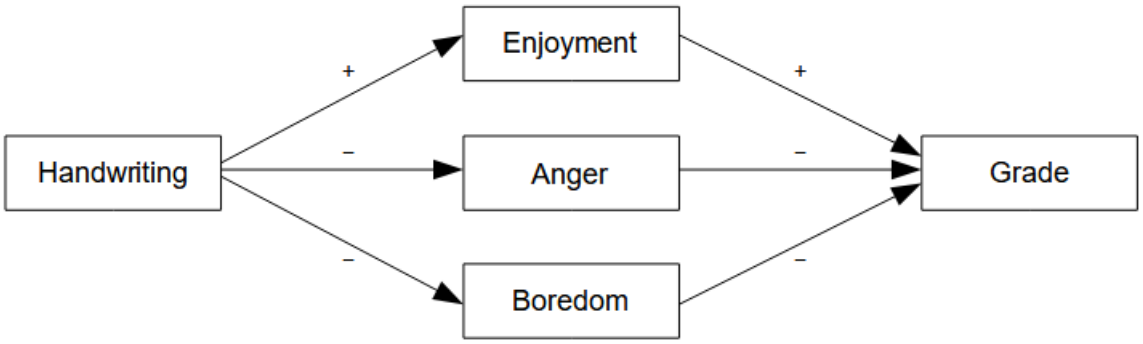
We sought to extend previous quantitative studies (Brackett et al., 2013; Townsend et al., 1989) in mainly two ways. First, by inducing activity emotions via two levels of a naturally occurring, material-inherent cue (good vs. bad handwriting quality) that is continually present during the entire study (as is the case in realistic grading situations), unlike the effects of externally induced mood that wear off over time. Second, by separately assessing discrete activity emotions, that is enjoyment (positive, activating), anger (negative, activating), and boredom (negative, deactivating) in a within-subject design, rather than using a group comparison approach, to disentangle emotion-specific effects.

We hypothesized that within participants the effect of the handwriting quality manipulation (good vs. bad handwriting) on grades was mediated by positive and negative

emotions in emotion-congruent ways (Brackett et al., 2013). More specifically, we assumed that handwriting quality predicted enjoyment positively and anger and boredom negatively, which in turn led to better and worse grades, respectively (see Figure 1.1).

Figure 1.1

Conceptual Model of the Assumed Two-Condition Within-Participant Mediations



Method

Participants and Procedure

Participants were school psychology student teachers at two universities in southern Germany attending the seminar “Diagnostic Case Studies”. The course prepared them for their final examination, which involved writing a summative review and recommendation integrating information provided in a case study (typically test results, as well as student, teacher, and parent interviews). After giving informed consent, participants randomly received an envelope containing the study materials in one of four variants. In the beginning, all participants indicated their current affective state. Then they graded two different essays of similar content quality written in different handwritings one after the other: after having read an essay, they directly assigned a grade to it and self-reported their emotions and

competence when grading the respective essay. At the end of the study, they judged the handwriting quality of both essays and provided their demographic data.

Due to substantial missing data, one dataset was removed from analyses. The final sample comprised $N = 73$ participants ($n = 62$ female), who were on average 25.31 ($SD = 4.30$) years old and had on average been enrolled for 9.54 semesters ($SD = 2.15$). Participants were familiar with the case study ($M = 4.06$, $SD = 1.30$) and rather inexperienced in grading written work ($M = 2.39$, $SD = 1.36$). Before they started grading the two essays, they were in a neutral to slightly positive affective state (positive PANAS: $M = 2.81$, $SD = 0.61$; negative PANAS: $M = 1.41$, $SD = 0.48$) and the experimental groups did not differ regarding their affective state (positive affect: $F(69, 3) = .34$, $p = .80$; negative affect: $F(69, 3) = .36$, $p = .78$).

Measures

Familiarity with the case study ("How familiar have you been with the case study?") and *experience grading written work* ("How much experience do you have with grading essays?") were both measured with single items on a 6-point Likert scale (1 = *not at all/none*, 6 = *very much/very good*). *Affective state* was measured by the German state version of the Positive and Negative Affect Scale (PANAS; Watson et al., 1988; German translation in Schwarzer, 1993), which separately measures positive ("Indicate to what extent you feel this way right now, that is, at the present moment", e.g., "excited", "proud") and negative affect (e.g., "nervous", "scared"). Both subscales (positive and negative affect) were measured on a 5-point Likert scale (1 = *not at all*, 5 = *very much*), consisted of 10 items each, and were internally consistent (Cronbach's α for positive and negative affect was .80 and .86, respectively). *Perceived handwriting quality* of the written material was measured with a single face-valid item on a 6-point bipolar scale ("How did you experience essay x during grading"?; 1 = *hard to read*, 6 = *easy to read*; based on Greifeneder et al., 2010). The discrete

activity emotions *enjoyment*, *anger*, and *boredom* were measured on a 4-point Likert scale (1 = *strongly disagree*, 4 = *strongly agree*; adapted to the grading context from Pekrun et al., 2005) with three items each. The factor structure of the scales was affirmed by a confirmatory factor analysis and the scales for *enjoyment* (e.g., "I had fun grading the text"), *anger* (e.g., "I was angry while grading"), and *boredom* (e.g., "I was so bored during grading, that time seemed to stand still") were internally consistent (Cronbach's α after having read the first/second essay: enjoyment: .90/.87, anger: .74/.87, boredom: .92/.93). Higher values corresponded to experiencing the respective emotion more strongly. *Competence* was measured with two items (e.g., "I felt competent grading the essay") on a 4-point Likert scale (1 = *strongly disagree*, 4 = *strongly agree*), which was internally consistent (Cronbach's α after having read the first/second essay: .74/.87). *Grades* were measured with one item ("Please assign an overall grade to essay x") according to the German grading system on a scale from 1 (*very good*) to 6 (*insufficient*). Straight letter grades could be increased or decreased by values of 0.3. Grades were then reverse-coded, so that higher values represented better grades. The codebook can be found in Appendix A.

For the within-participant analyses all measures were transformed in such a way that each case was assigned one value for each variable in the good and one in the bad handwriting condition – irrespective of the order and the essay in which the good and bad handwriting were presented in the study.

Materials: Essay Construction

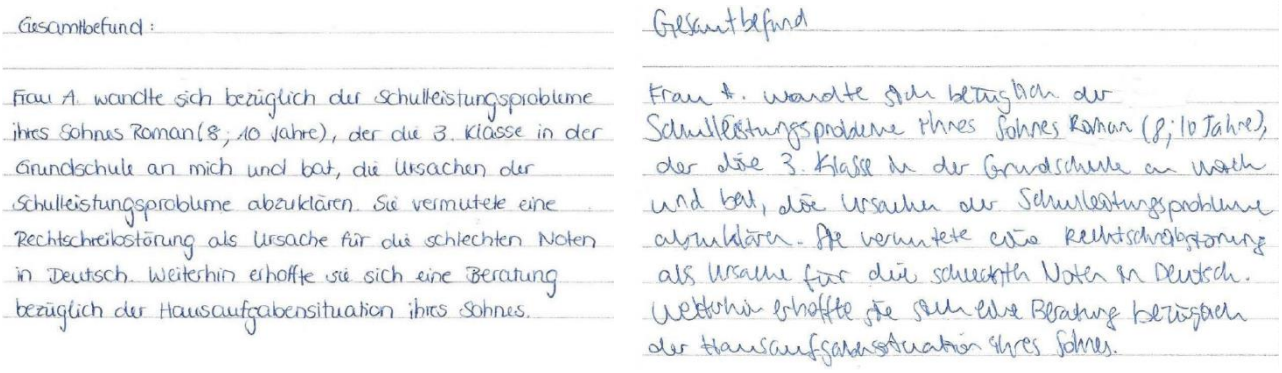
Prior to the pilot study, a course instructor screened real essays from a previous examination cohort and selected two essays (essays A and B) of similar length and content quality on an average performance level (letter grade "C") as study materials. Additionally, multiple handwriting samples were collected to obtain sufficiently realistic, yet maximally divergent handwriting qualities. Thirteen independent raters (n = 11 female) judged their

handwriting quality. The most and least legible female and male handwriting samples were selected and their authors copied both essays verbatim in their typical handwriting as if they were in a time-constrained examination situation (for handwriting samples, see Figure 1.2).

We used the pilot study ($N = 28$; $n = 22$ female) to ensure that the content quality of the two essays was comparable and to select two suitable handwritings for the main study. Results showed that student teachers judged the content quality of both essays similarly (Essay A: $M = 3.95$, $SD = 0.98$; Essay B: $M = 4.25$, $SD = 1.07$; $t\ (27) = -1.23$, $p = .23$) when using a grading rubric to enhance objectivity and accuracy. Both essays would have been considered a "C" and students' average assigned grades converged with the instructor's initial assessment. To control for gender effects, we selected the two female handwritings for the main study. They differed sufficiently in handwriting quality (good handwriting: $M = 5.93$, $SD = 0.26$; bad handwriting: $M = 1.64$, $SD = 0.81$; $t\ (11.50) = 16.99$, $p < .001$) and the bad handwriting was still legible.

Figure 1.2

Handwriting Samples of the Good (Left Side) and Bad (Right Side) Handwriting Condition



Research Design

The present study had a 2 (handwriting quality: good vs. bad) x 2 (essay: essay A vs. essay B) mixed design, with handwriting quality and essay as within-subject factors. To control for possible order effects, materials were counterbalanced, which resulted in four groups (between-subject factor). Participants experienced both levels of each within variable (e.g., one essay in good handwriting was combined with the other essay in bad handwriting; see Table 1.1 for more detailed information about the experimental groups).

Table 1.1

Number of Participants and Combination of Handwriting Quality and Essay in the Experimental Groups

	n	First essay	Second essay
Group 1	17	Essay A, good handwriting	Essay B, bad handwriting
Group 2	17	Essay B, bad handwriting	Essay A, good handwriting
Group 3	21	Essay A, bad handwriting	Essay B, good handwriting
Group 4	18	Essay B, good handwriting	Essay A, bad handwriting

Results

Statistical analyses were performed with SPSS 24 and the macro MEMORE (Montoya & Hayes, 2017). MEMORE simultaneously estimates direct, indirect, and total effects for mediation analyses in two-condition within-participant designs applying a path-analytic framework. We used percentile bootstrap confidence intervals (CI; 95%) with 10,000 bootstrap samples and reported the unstandardized path coefficients. The unstandardized path coefficients regressing the mediator (emotions) and the outcome variable (grade) on the condition (handwriting quality) only can be interpreted as the mean difference within each participant between the good and the bad handwriting condition with regard to the respective

variable. The unstandardized path coefficient from the mediator (emotion) to the outcome (grade) estimates the difference between the good and bad handwriting condition regarding the outcome variable (grade) from the difference between the two conditions in the mediator variable (emotion).

Manipulation Check

To ensure that the handwriting manipulation was successful, participants judged the handwriting quality of the essays in good and bad handwriting at the end of the study. As intended, a paired-samples *t*-test showed that the perceived handwriting quality was significantly better in the good than in the bad handwriting condition (see Table 1.2). Perceived competence in grading the essays did not differ between the handwriting conditions ($t(73) = 0.94$, $p = .35$).

Main Analyses

Means of all variables but boredom showed significant differences between the good and bad handwriting condition (Table 1.2). Table 1.3 shows the correlations between the study variables.

Table 1.2

Means and Standard Deviations (in Parentheses) for All Variables and Comparison Between the Good vs. Bad Handwriting Condition

Handwriting condition	Perceived HW quality***	Anger***	Enjoyment***	Boredom	Grade*
Good HW	5.86 (0.42)	1.55 (0.57)	2.54 (0.76)	1.63 (0.77)	4.36 (0.98)
Bad HW	1.45 (0.73)	2.13 (0.86)	2.04 (0.69)	1.70 (0.79)	4.05 (0.90)

Note: Paired-samples *t*-tests were used to test for differences between the good and bad handwriting (HW) condition.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 1.3

Correlations of All Study Variables in the Good and Bad Handwriting Conditions

	1	2	3	4	5	6	7	8	9	10
1. PHWQ; good HW	—									
2. PHWQ; bad HW	0.07	—								
3. Anger; good HW	0.04	0.05	—							
4. Anger; bad HW	0.09	−0.29 *	0.41 **	—						
5. Enjoyment; good HW	−0.02	−0.02	−0.52 **	0.01	—					
6. Enjoyment; bad HW	−0.08	0.28 *	−0.20	−0.56 **	0.34 **	—				
7. Boredom; good HW	0.04	0.08	0.55 **	0.16	−0.49 **	−0.26 *	—			
8. Boredom; bad HW	0.03	−0.04	0.45 **	0.33 **	−0.33 **	−0.19	0.62 **	—		
9. Grade; good HW	−0.09	−0.05	−0.46 **	−0.03	0.33 **	−0.01	−0.04	0.00	—	
10. Grade; bad HW	−0.21	−0.03	0.09	−0.09	−0.11	0.08	0.28 *	0.17	0.13	—

Note. "Good HW" represents the variable in the good handwriting (HW) condition, „bad HW" represents the variable in the bad handwriting condition. *n* ranged from 71 to 73 participants. PHWQ = Perceived handwriting quality.

* $p < .05$. ** $p < .01$. *** $p < .001$

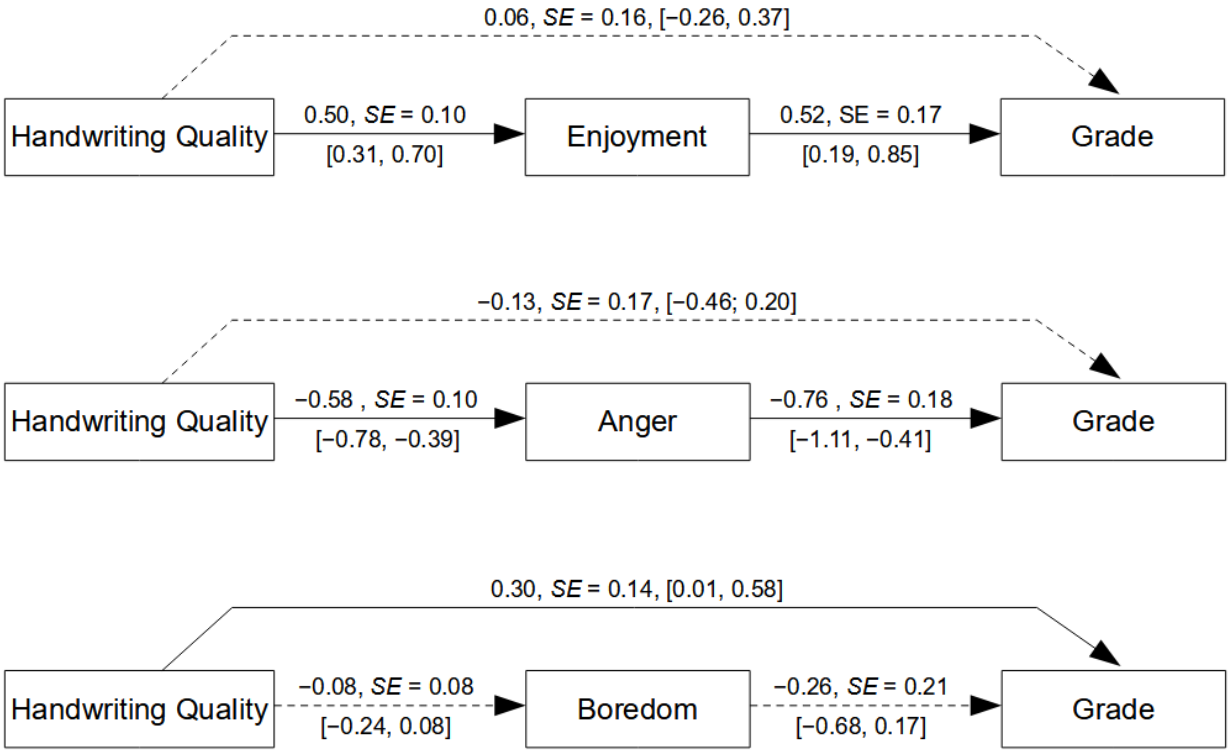
The total effect of handwriting on grades was 0.32 (*SE* = 0.15, CI [0.03, 0.61]), which equaled a grade difference between the good and bad handwriting condition of one third of a letter grade. Separate models (see Figure 1.3) that depict unstandardized path coefficients were estimated for enjoyment, anger, and boredom.

Overall, good handwriting predicted higher level of enjoyment and lower levels of anger, whereas it did not predict boredom. More specifically, participants experienced 0.50 units more enjoyment and 0.58 units fewer anger in the good compared to the bad handwriting condition. Higher levels of enjoyment and anger in turn predicted better and worse grades, respectively, whereas boredom did not predict grades. More specifically, a difference of one unit with regard to enjoyment and anger between the good and bad handwriting condition predicted a difference in grades of 0.52 and −.76 units, that is the grades in the good handwriting condition were half a grade better and three-quarters of a grade worse in the good compared to the bad handwriting condition, respectively. The indirect effect of handwriting on grades through enjoyment was significant (0.26, *SE* = 0.09, CI [0.06, 0.43]) and the model accounted for 15 percent of the variance in grades ($F(2,70) = 6.19$, $p < .001$, $R^2 = .15$). The indirect effect of handwriting on grades through anger was significant (0.44, *SE* = 0.12, CI [0.21, 0.67]) and the model accounted for 23 percent of the variance in grades ($F(2, 70) = 10.70$, $p < .001$, $R^2 = .23$). The indirect effect of handwriting on grades through boredom was not significant (0.02, *SE* = 0.03, CI [−0.04, 0.10]) and the model accounted for six percent of the variance in grades ($F(2, 70) = 2.13$, $p = .13$, $R^2 = .06$). The pattern of findings remained stable when controlling grades for perceived competence. When estimating one single model with enjoyment, anger, and boredom as parallel mediators ($F(6, 66) = 3.71$, $p = .003$, $R^2 = .25$), the total indirect effect (the sum of all three separate indirect effects through each emotion) was significant (0.47, *SE* = 0.12, CI [0.22, 0.73]), indicating that the effect of handwriting on grades was mediated by emotions. The single

indirect effects of the emotions, however, were not significant. Nevertheless, the simultaneous estimation descriptively gave reason to assume that the effect of anger on grades (−0.56, *SE* = 0.27, CI [−1.09, −0.02]) was stronger than the effect of enjoyment on grades (0.27, *SE* = 0.23, CI [−0.20, 0.74]), when controlling for effects of the other two emotions. The effect of boredom was not significant in either of the models. Overall, the results of the parallel model were similar to the results from the separate models.

Figure 1.3

Separate Two-Condition Within-Participant Mediation Models Depicting Unstandardized Path Coefficients for Enjoyment, Anger, and Boredom as Mediators Between Handwriting Quality and Grade



Note. Solid lines depict significant paths. Dashed lines depict nonsignificant paths.

Discussion

The aim of the present study was to investigate whether three discrete activity emotions (enjoyment, anger, and boredom) influenced essay grades and whether these grading emotions could be induced by the grading task itself by means of a naturally-occurring, constantly present task-inherent cue (i.e., handwriting quality). To this end, we tested whether the three discrete activity emotions enjoyment, anger, and boredom mediated the effect of handwriting quality on grades in a two-condition within-participant design.

Handwriting quality, a relatively superficial feature of written student work, holds the potential to bias grading. Previous research showed in between-subject and mixed designs that participants assigned better grades to well-legible compared to less legible student work (Briggs, 1970, 1980; Greifeneder et al., 2010, 2012; James, 1927; Markham, 1976). These findings were replicated in our two-condition within-participant design: participants read two different essays of similar content quality and evaluated the essay in good handwriting around one third of a letter grade better than the essay in bad handwriting. Although the legibility bias may be a serious threat to the reliability and validity of grades, it was not the focus of the present study. Moreover, it has been shown that its effects can be minimized by informing graders about its influence (Greifeneder et al., 2010).

In our study, the handwriting manipulation was primarily used as means to induce discrete activity emotions to explore their presence and effects in grading situations. Overall, the induction was successful because the essay in better handwriting compared to the essay in bad handwriting induced higher levels of enjoyment and lower levels of anger, respectively. This supports the idea that the grading task itself can trigger positive and negative activity emotions. Drawing on control-value theory (Pekrun, 2006), the development of discrete emotions may be due to specific control and value appraisals assigned to the grading task. In the present study participants felt equally capable of grading the two essays irrespective of

handwriting quality (identical control appraisals), but they may have perceived grading the essay in good handwriting as more pleasant (higher positive value) than grading the essay in bad handwriting (higher negative value due to task aversiveness), which may have contributed to experiencing differing levels of enjoyment and anger. This line of reasoning generally aligns with claims that good and bad handwriting quality may result in positive and negative affective reactions, respectively (Reber & Greifeneder, 2017; Winkielman et al., 2003). This general claim, however, did not extend to boredom, although it is another negative activity emotion. In contrast to enjoyment and anger, which are qualified by high positive and negative value respectively, boredom is qualified by a lack of value. And although grading holds the potential to elicit boredom due to its monotonous and repetitive nature (Pekrun et al., 2010), it seems reasonable that handwriting quality is none of the main factors that make grading a tedious task of little or no value.

Mostly in line with research on the effects of mood on grades in between-participant designs (Brackett et al., 2013), our within-participant results showed that one positive and one negative discrete emotion influenced grades in emotion-congruent ways. In line with expectations, higher levels of enjoyment and anger led to the assignment of better (around half a letter grade) and worse grades (around three-quarters of a letter grade), respectively. Unexpectedly, the negative emotion boredom did not influence grades. This may be explained by the very low boredom scores (probably due to the relatively high utility value in our study, caused by the relevance of the essays for the exam preparation) that did not differ between the good and bad handwriting condition. Because the two-condition between-participant mediation uses differences between the two conditions in the predictor variable to predict differences in the outcome variable, the very similar levels of boredom when grading the essays in varied handwriting quality could not result in grade differences.

To summarize, our results showed that the effect of handwriting quality on grades was mediated by enjoyment and anger, but not boredom. Considering estimates from both the separate mediation models and the parallel mediation model, our data furthermore suggested that the effect of enjoyment on grades (separate model: 0.52; parallel model: n.s.) was not as strong as the effect of anger (separate model: −0.76; parallel model: −0.56). This hints in the direction that the legibility bias may be mainly driven by anger (negative emotion) rather than by enjoyment (positive emotion). Moreover, anger may be a more serious threat to fair grading than enjoyment, especially when taking into account that negatively biased grades have more disadvantageous consequences for students than positively biased grades. Due to the unsuccessful boredom induction by means of the handwriting quality manipulation, the role of boredom in grading situations remains open.

Limitations and Future Research

The present study offered first experimental insights into the role of discrete emotions when grading, but was also subject to some limitations. Generally, it is very hard to create an experimental setting that fully resembles an authentic grading situation (i.e., teachers grading many essays of their own students on the same topic and assigning real grades that are relevant for the graders' students but do not have consequences for teachers). In the present study, student teachers rather inexperienced in grading read and assessed only two essays that could be used as exam preparation. These factors may have created a situation that was less repetitive and less aversive than typical grading because the situation held positive value for the student teachers (possibly intrinsic value due to interesting content, utility value due to exam preparation). Therefore, the absolute levels of the experienced emotions may not be representative of real grading situations. While handwriting influenced enjoyment and anger probably because the essay in bad handwriting created more task aversiveness leading to lower positive and higher negative value, boredom was not influenced, possibly because

other factors are responsible for eliciting boredom in grading situations. Based on control-value theory such factors could be very high or low levels of perceived competence triggering boredom due to under- or overchallenge, respectively, in combination with a lack of value (Pekrun et al., 2010), which may be caused by uninteresting essay topics, reading multiple essays on the same topic, and seeing no utility in grading the students' work. Additionally, high costs associated with grading, such as competing and possibly more important and pleasant tasks, or effort and time needed for grading may contribute to experiencing boredom when grading. Therefore, other possibilities to induce boredom in experimental grading settings need to be explored.

To extend and generalize the presented experimental findings on the effects of discrete activity emotions on grading, it is necessary to investigate grading emotions in real grading situations as experienced by in-service teachers to increase external validity. It is likely that grading holds the potential to elicit further discrete emotions, such as pride because students did well on a test or because the teacher managed to handle the grading well or anxiety because a teacher worries about the reliability of assigned grades. It is necessary to understand their occurrence, antecedents, and effects in grading to make the task less aversive and more reliable. The control-value theory framework (Pekrun, 2006) may provide a promising theoretical basis to do so. Meaningful insights into the emotional experiences of in-service teachers when grading may be obtained by conducting interviews with teachers, comparing their emotional experiences in different tasks (e.g., teaching, class preparation, grading) using cross-sectional surveys, or by investigating intra-individual differences when grading for a longer period of time for example by means of event-contingent experience sampling.

Implications

Considering the emotionally challenging nature of teachers' jobs (Hargreaves, 1998) and the high burnout rates among teachers (Hakanen et al., 2006), it is key to foster teachers' well-being to sustain a well-working educational system. Research so far was mainly concerned with investigating teacher emotions when in the classroom. It is important, however, not to neglect emotions experienced outside the classroom, because those activities take up a good portion of teachers' working time (OECD, 2014), are often associated with high task aversiveness (Laybourn et al., 2019), and therefore are likely to unfold negative effects on teachers' well-being. The present study showed that the grading activity itself can elicit emotions. By identifying factors in grading situations that elicit emotions, it may be possible to give teachers some advice on how grading may be made less unpleasant. For example, teachers may encourage students to write essays on their computers to circumvent the negative emotional effects of bad handwriting. Furthermore, it is important to encourage teachers to reflect on their emotional experiences when grading and to find ways to improve them if necessary. This may be possible by means of actively searching for positive value in grading, such as intrinsic value by choosing interesting topics and utility value by finding reasons why grading is useful for the teacher (for instance getting information about students' strengths and weaknesses), or reducing the perceived costs of grading by breaking the work into smaller chunks and to reward themselves after having completed one pile of papers. One important side-effect of such strategies would be that grading bias may be reduced when teachers are aware that their judgements are influenced by their own emotions – irrespective of how these emotions are elicited. In general, attention should be paid to the role of emotions in the teaching profession already in teacher training, to prepare student teachers for their often challenging occupation.

Conclusion

Overall, we conclude that handwriting quality is one naturally-occurring task-inherent cue that may induce specific activity emotions in experimental grading situations. Choosing task-inherent emotional cues for an emotion induction procedure resembles the development of emotions in grading situations probably more closely than externally induced mood. It is important to keep in mind, though, that in real grading situations manifold sources may trigger specific positive and negative activity emotions. To name just a few, these may be features of the student work, such as content quality, structure, spelling, grammar, or creativity, characteristics of the teacher, such as interest in the essay topic, perceived value of grading, or perceived grading competence, and contextual circumstances, such as a large number of essays that need to be graded, competing tasks that need to be completed, high workload, or stress. Due to these various possible origins of emotions in grading situations and their potential to not only bias grades but also influence teachers' well-being (Frenzel & Stephens, 2017), it is necessary to scrutinize the specific antecedents and effects of discrete grading emotions to eventually find ways to make grading less of a "tedious thing to do because you're not reading it for enjoyment" (Baker, 2014, p. 43).

Grading is not only a relevant and under-researched task of teachers', but also of faculty members' profession. Therefore, we aimed to investigate grading emotions in more detail in a faculty sample as a next step. We aimed to contrast emotions for grading against those for research and teaching and to explore antecedents of faculty grading emotions from a control-value perspective in two countries (U.S. and Germany).

3. Studies 2a+b

How Do University Faculty Feel About Grading? Insights From a Control-Value Theory Perspective

Until quite recently, tertiary educational research seems to predominantly have viewed university professors and instructors as logical, rational beings, who are not emotional whatsoever. This image, however, does not seem to reflect the day-to-day realities of faculty very well. In response, the emotional experiences of faculty have gained research attention in the past five to ten years. Qualitative studies found faculty experienced various emotions in different situations throughout their work days (e.g., Hagenauer & Volet, 2014; Lahtinen, 2008; Meanwell & Kleiner, 2014; Postareff & Lindblom-Ylänne, 2015), and a recent survey revealed some of the most common emotions faculty reported to experience were enjoyment, happiness, excitement, pride, anxiety, and frustration (Stupnisky et al., 2019a).

The present research seeks to add to the growing literature on faculty emotions by quantitatively assessing discrete faculty emotions pertaining to an important yet so far neglected task: grading, that is, assessing the quality of student work and assigning a letter grade or pass/fail judgement. In so doing, we focus on grading student papers. We propose that grading such long answer and essay-type (i.e., constructed-response) assignments can be very emotionally arousing, in contrast to grading multiple choice and true or false tests (i.e., selected-response; Reynolds et al., 2009), which is more mechanical and in modern times mostly done automatically by learning management systems.

The present research had two goals: First, we sought to quantify the degree to which faculty experience a range of discrete emotions during grading, as contrasted against two other important faculty task areas, namely teaching and research (Study 2a). Secondly, we aimed to explore the appraisal antecedents of grading emotions in two countries (Study 2b).

Pekrun's control-value theory of emotions (2006) served as a conceptual framework for this research.

Control-Value Theory

In our research, we adopt a definition of emotions which considers them as multi-faceted constructs that comprise affective, behavioral, cognitive, physiological, and motivational components (K. R. Scherer, 2005). In the context of academic emotions, control-value theory (CVT) is a highly prominent theoretical framework (Pekrun, 2018). CVT proposes that subjective control and value appraisals of an activity and its outcome are key cognitive antecedents for an individual's discrete emotions experienced during the activity. Subjective control ranges from low to high and reflects the ability to influence activities or their outcomes. Subjective value ranges from negative to positive and reflects whether activities or their outcomes are appraised as desirable (positive value) or undesirable (negative value). CVT further claims that particular combinations of control and value appraisals in a specific situation shape an individual's emotional experiences. Perceiving a situation as (un-)desirable influences the valence (positive or negative) of an emotion, and feeling in or out of control influences its quality (e.g., anxiety vs. anger). The extent of control and value also influence the intensity of an emotion: high control is typical for strong positive and weak negative emotions, whereas high value intensifies the experience of both positive and negative emotions (Pekrun, 2018). To date, the control-value theory has been predominantly applied to typical achievement contexts, such as students of varying ages studying, attending classes, writing exams, or engaging in online learning environments (e.g., Daniels & Stupnisky, 2012; Peixoto et al., 2017; Pekrun et al., 2011). But it has also been applied and adapted to higher education research, that is, in faculty emotion research (e.g., Stupnisky et al., 2019b).

It is worth noting that CVT makes specific propositions for the elicitation of a range of discrete emotions, yet those existing quantitative studies on faculty emotions often did not assess discrete emotions or combined them into positive and negative affect for further analyses (Stupnisky et al., 2019a, 2019b). This approach captures emotional valence but does not suitably reflect the distinct nature of emotions (Trigwell, 2012). In our studies we chose to investigate six discrete faculty emotions (enjoyment, pride, boredom, anxiety, anger, and frustration) grounded on past empirical findings regarding their frequencies and prominence among faculty in general (Kordts-Freudinger, 2017; Stupnisky et al., 2019a) and relevance judgements for the task of grading in particular.

Contextualizing Faculty Emotions

During their work, faculty engage in diverse tasks in the domains of research and teaching that can be very disparate in nature. While school teachers' emotions have been shown to be largely context-specific (e.g., Frenzel et al., 2015, 2016), comparing emotional experiences between different work contexts has so far been uncommon in faculty emotion research. Most studies focused on a single context, typically teaching (Meanwell & Kleiner, 2014; Postareff & Lindblom-Ylänne, 2015), and further contexts or more specific tasks of faculty work were largely neglected. One exception is the line of research on faculty emotions presented by Stupnisky and colleagues who compared the emotions faculty experienced in two main contexts of faculty work, namely research and teaching. Their results showed that the experienced emotions differed between contexts: research generally elicited lower levels of positive emotions (such as enjoyment, happiness, pride, satisfaction, and relaxation) and higher levels of negative emotions (such as frustration, anxiety, worry, fear, envy, shame, loneliness, and hopelessness) compared to teaching (Stupnisky et al., 2016, 2019a).

Faculty emotions towards research have been shown to be linked with faculty members' research success (Stupnisky et al., 2019a, 2019b), and their emotions towards teaching have been shown to be linked with teaching success (Stupnisky et al., 2019a) and teacher- vs. student-focused approaches to teaching (Kordts-Freudinger, 2017; Trigwell, 2012). This implies that faculty teaching emotions indirectly affect student outcomes. Conversely, we propose that faculty grading emotions are particularly relevant for student outcomes: Course grades provide feedback on student learning and thus shape students' competence beliefs, and even more importantly, they determine GPA and degree attainment, thus subsequently, career opportunities. Prior studies showed that grades may be biased by factors unrelated to student achievement, such as ethnic or racial group, prior performance, or attractiveness (meta-analytical findings; Malouff & Thorsteinsson, 2016). Importantly, also teacher emotions have been shown to influence grades in emotion-congruent ways (Brackett et al., 2013): after a mood induction (positive or negative), participants were asked to evaluate the same middle school narrative essay, and results showed that teachers in a positive mood rated the identical essay more favorably with respect to creativity, spelling/punctuation, and overall performance than those in a negative mood. Because grades have far-reaching consequences for students, and emotions potentially bias grading (see also Study 1), it is important to consider grading as a potentially emotion-arousing faculty task.

Faculty Emotions When Grading

Although grading is an important and time-consuming task for faculty, we know of no study that explicitly addressed the emotional experiences of faculty when grading. We propose that grading can be conceptualized as a task within the scope of teaching, on the one hand, but also as a separate task in addition to teaching, on the other. In certain situations, for example when there are multiple assessments over the semester that inform faculty about student progress, grading may be best described as a specific task within the broader context

of teaching more generally. In contrast, sometimes key final assessments are graded after a longer period of time once class has ended (i.e., final papers), student papers are graded anonymously, or grading happens within the context of centralized exams where faculty were not involved in teaching the examinees at all. In such situations, grading is rather detached from teaching and therefore best described as a separate task. Within the present research, we primarily treat grading as a separate task, while bearing in mind its potential close interconnectedness with teaching. We also consider the different circumstances of grading when comparing across countries that likely have differing cultural norms for teaching and grading.

Within the scope of some qualitative studies on the emotional experiences of faculty, grading was reported to trigger predominantly negative emotions (Lahtinen, 2008) and moreover to elicit negative emotions more frequently than positive emotions (Meanwell & Kleiner, 2014; Postareff & Lindblom-Ylänne, 2015). Specifically, instructors rather frequently reported negative emotions such as worry and shame with respect to assessment. Those negative emotions were reported to occur when there was a lack of clearly right or wrong answers (Hagenauer & Volet, 2014), when instructors more generally doubted that their testing was reliable and valid (Postareff & Lindblom-Ylänne, 2015), also triggered by students questioning their competence to reliably assess student performance (Lahtinen, 2008). In sum, negative emotions often seemed to be linked with low subjective diagnostic competence in terms of the belief that one can assign valid and reliable grading results. The rare reports of positive emotions with respect to assessment pertained to mastering new and innovative forms of assessment (Postareff & Lindblom-Ylänne, 2015).

In sum, grading is a highly relevant task for faculty that is believed to trigger predominantly negative emotions with the potential to detrimentally affect instructors and

students. Nevertheless, studies on this topic are limited with no known quantitative insights available.

Study 2a

Study 2a aimed to replicate Stupnisky et al.'s (e.g., 2016, 2019a) findings on the comparison between faculty emotions for research and teaching, while adding grading as a specific task. In so doing, we expected to replicate Stupnisky et al.'s (2016, 2019a) prior findings that research generally elicited less positive and more negative emotions than teaching. Based on previous qualitative findings (Lahtinen, 2008; Postareff & Lindblom-Ylänne, 2015), we further expected that grading would overall be experienced as largely emotionally aversive, more specifically, that grading would elicit fewer positive and more negative emotions than teaching. We had no a-priori expectations regarding the comparison between grading and research emotions.

Method

This study's final sample consisted of $N = 1{,}226$ faculty from 13 institutions in the U.S. (51% female; aged $M = 49.54$ years, $SD = 11.32$; years of teaching experience $M = 16.09$, $SD = 11.09$) participating in a larger study of the Faculty Survey of Student Engagement (FSSE; Center for Postsecondary Research, 2020), after excluding participants who were missing answers on the central study variables ($n = 88$). Most participants were employed full time (85.28%) as assistant professors (26.23%), professors (24.10%), associate professors (21.98%), or instructors (13.40%) and either tenured (41.24%), on tenure track (21.21%), or not on tenure track (32.19%). They were mainly white (81.53%) and came from more than ten disciplines, the most frequent being arts and humanities (21.80%) and health professions (13.79%). Critical to the current study, faculty reported spending a moderate to large amount of time on grading per semester ($M = 3.42$, $SD = 1.02$; measured on a 6-point rating scale ranging from 0 = *none at all* to 5 = *very large amount*).

Participants first rated the frequency of experiencing six discrete emotions (enjoyment, pride, boredom, anxiety, anger, frustration) with respect to research, teaching, and grading, one after the other, on a 5-point rating scale ("How often do you typically experience the following emotions when conducting research/teaching/grading student papers?"; 0 = *never*, 1 = *rarely*, 2 = *sometimes*, 3 = *often*, 4 = *very often*). Items were based on Stupnisky et al. (2019a) to allow for comparability with previous work. Faculty assessed grading emotions last to ensure that the inclusion of grading as a distinct task did not affect the understanding of "teaching" as a broader context as asked for in the previous question and in previous work. Grading student papers was specifically defined as "assessing the quality of students' work on assigned papers, reports, or other writing tasks (submitted as hardcopy or electronically) … that included open-ended constructed-response essays and long-answer questions that require considerable student writing (NOT selected-response multiple choice and true/false questions)". The codebook is available in Appendix B.

Results and Discussion

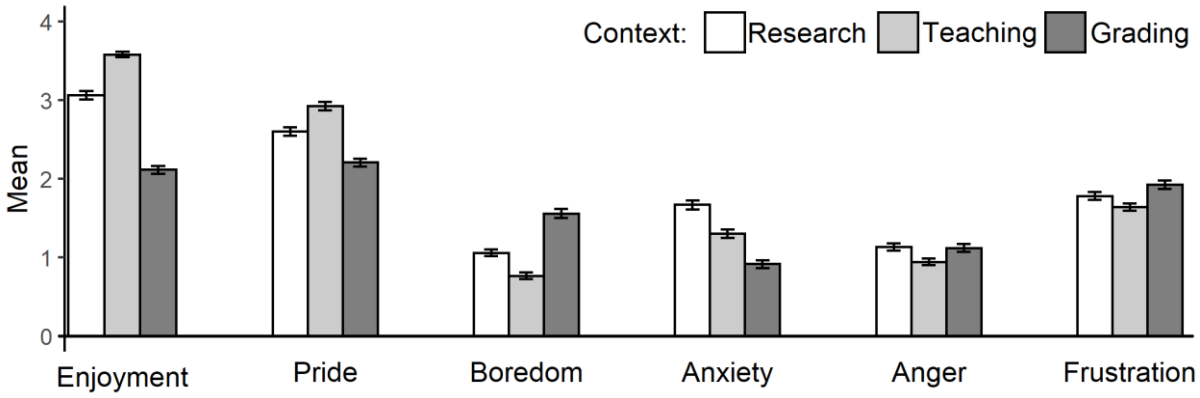
Data was analyzed using R (R Core Team, 2018) and the results of the comparisons across the three contexts for the six emotions are displayed in Figure 2.1 (exact values and effect sizes are displayed in Table 2.1). To account for multiple comparisons, the alpha level was adjusted to $p < .001$.

Overall, results showed that participants reported experiencing positive emotions more frequently than negative emotions across all three contexts. More importantly, and in line with earlier findings, faculty reported experiencing enjoyment and pride significantly less frequently when conducting research than when teaching. Conversely, boredom, anxiety, anger, and frustration were reported significantly more frequently when conducting research than when teaching. This is largely in line with previous findings (Stupnisky et al., 2019a, 2016), with the exception that the frequencies of boredom and anger experienced in research

and teaching had previously been found to be comparable. Regarding anger and frustration though, the effect sizes for the pairwise comparison between research and teaching in our data were small (Cohen's $d < 0.25$) and therefore the significant effects were probably due to the big sample size.

Figure 2.1

Repeated Measures ANOVA Results Comparing the Contexts Research, Teaching, and Grading Including Post-Hoc t*-Tests*



Note. Regarding the repeated measures ANOVAs, the assumption of sphericity was violated, therefore the conservative Greenhouse-Geisser corrected results were calculated (Field et al., 2012). *F*-statistics for all ANOVAs were significant. Post-hoc *t*-tests were Bonferroni corrected. All pairwise comparisons across any two contexts were significantly different from one another ($p < .001$), except for anger during research vs. grading ($p = 1.00$). Error bars represent the 95% CI.

Further, in line with expectations, grading generally triggered positive emotions less frequently, and negative emotions more frequently than both research and teaching. Noteworthy were the large effect sizes for enjoyment (Cohen's $d > 0.80$) which indicated that grading was a task that was notably less enjoyable than both other tasks considered here. Similarly, the frequency of boredom was significantly higher when grading compared to both teaching (Cohen's $d = 0.85$) and research (Cohen's $d = 0.48$). This may originate from grading being a highly repetitive task that is typically not very rewarding, and therefore being a typical context in which boredom due to under-challenge may arise (Acee et al., 2010). One exception to the general pattern of findings was anxiety that was reported less frequently when grading. A reason for this may be that "failures" during grading, different from "failures" with respect to teaching and research, typically do not have any (direct) negative consequences for faculty. Furthermore, if one feels insecure during grading, one can take time to compensate for missing knowledge, while insecurity during teaching or research may not be as readily compensated.

Taken together, findings from this study suggest that grading is a faculty task accompanied by rather negative emotional experiences; particularly when compared against research and teaching more generally. In our next study, we investigated grading emotions in more detail; specifically, we aimed to demonstrate that the predominantly negative emotions during grading are not limited to the U.S. higher education context, and to identify appraisal antecedents that may elicit emotions in faculty when engaging in this important, yet clearly unenjoyable task.

Table 2.1

Study 2a: Descriptive Statistics of Discrete Emotions and Repeated Measures ANOVA Results Including Post-Hoc t*-Tests and Effect Sizes*

	Descriptive Statistics			ANOVA Results		Cohen's *d* for Pairwise Comparisons		
	Research (R)	Teaching (T)	Grading (G)			R vs. T	T vs. G	G vs. R
	M (SD)	*M (SD)*	*M (SD)*	*F*	η^2	*d*	*d*	*d*
Enjoyment	3.06 (0.91)	3.58 (0.62)	2.11 (0.92)	$F(1.87, 2285.87) = 1303.50^{***}$	0.35	0.53	1.58	0.83
Pride	2.68 (0.96)	3.01 (0.96)	2.27 (0.94)	$F(1.94, 2370.47) = 387.62^{***}$	0.09	0.37	0.84	0.40
Boredom	1.09 (0.82)	0.79 (0.76)	1.61 (1.06)	$F(1.88, 2297.84) = 447.34^{***}$	0.12	0.35	0.85	0.48
Anxiety	1.72 (1.05)	1.34 (0.96)	0.95 (0.90)	$F(1.93, 2369.59) = 376.88^{***}$	0.10	0.39	0.43	0.72
Anger	1.17 (0.86)	0.97 (0.78)	1.16 (0.93)	$F(1.89, 2313.97) = 39.92^{***}$	0.01	0.24	0.23	0.02
Frustration	1.84 (0.89)	1.69 (0.85)	1.98 (0.96)	$F(1.90, 2321.50) = 57.55^{***}$	0.02	0.16	0.33	0.14

Note. The assumption of sphericity was violated, therefore the conservative Greenhouse-Geisser corrected results were calculated (Field et al., 2012). Post-hoc t-tests were Bonferroni corrected. Due to multiple comparisons, the alpha level for the pairwise comparisons was adjusted to $p < .001$. All pairwise comparisons across any two contexts were significantly different from one another ($p < .001$), with the exception of anger during grading vs. research ($p = 1.00$).

$^{***}p < .001$.

Study 2b

This study used the CVT of achievement emotions as a theoretical framework. Although originally developed for the context of students' achievement emotions, we propose that CVT also provides a suitable lens through which to view faculty grading emotions because grading is a task that is embedded within a classical achievement context (examinations). A notable difference, however, is that in typical achievement situations an individual's own performance is assessed, whereas when grading, it is another individual's performance that is assessed by the grader whose "grading performance" is not necessarily in the focus of the activity. Therefore, one objective of this study was to identify specific control and value appraisals that seemed relevant for grading.

Control Appraisals in Grading Situations

In the control-value framework (Pekrun, 2006), control appraisals are typically conceptualized in terms of subjective competence in successfully executing a required activity, which seems highly relevant in the context of grading as well. We propose that there are two facets about which instructors may feel more or less competent about when grading: the examined content (content knowledge) and the process of determining grades (diagnostic competence).

At first it seems that all instructors should typically be highly knowledgeable about an exam's content. There are situations, however, in which the graders may be rather unfamiliar with the content, for example, when they just started teaching, substitute for a colleague in an unfamiliar course, or grade state-wide exams with externally designed, randomly selected topics that may not be within the scope of their expertise. It has been reported earlier that a lack of content knowledge can be perceived as threatening in the teaching context (Lahtinen, 2008), and we propose that this could also extend to grading; specifically as grade decisions may be challenged (Stough & Emmer, 1998).

Furthermore, instructors may differ with respect to their perceived ability to diagnose performance levels, that is to determine grades fairly, transparently, and reliably. Although teaching and grading are important aspects in any instructor's work life, there is usually no formal training for either teaching or diagnostics in the higher education context (Murtonen & Vilppu, 2020). One reason may be that the main focus within higher education institutions lies on research success and not so much on teaching success (Cadez et al., 2017). Therefore, instructors' development with respect to grading is highly dependent on their own motivation, need for fairness, and approach to grading, which makes it very likely that individuals vary in their appraisals regarding their diagnostic competence.

In addition to those control-related appraisals pertaining to grading competence, we argue that there are further aspects about the task of grading which may imply varying levels of faculties' action control (Pekrun, 2006). Those pertaining to the grading process (i.e., the design of an exam such as the duration and number and nature of questions, and the grading activity such as use of a rubric) and the exam content (i.e., what is covered in the questions). In some contexts, those aspects may be rather strictly defined by study regulations, in other contexts, instructors may be relatively free in picking the exams' content and formulating questions, defining grading criteria, or organizing the grading process. As a result, we propose that faculty members may vary in their control appraisals pertaining to the grading process and exam content.

Value Appraisals in Grading Situations

The value component within the control-value theory pertains to the importance of an activity or its outcome. Our deliberations about potential facets of the value of grading were inspired by earlier work on students' academic values (Gaspard et al., 2015). As a result, we identified four facets of value that faculty may attach to the activity of grading and its outcomes: diagnostic value, utility value, social value, and cost.

Diagnostic value pertains to the importance of assigning fair, transparent, and reliable grades. Assigning accurate grades may be important to faculty to fulfill their internal grading standards, or to avoid formal appeals against grades by students. *Utility value* may arise, for example, when grading provides faculty with information about their students' current state of knowledge and when this information may be utilized to adapt their teaching. *Social value* may originate from the impact of grading on the relationship with students, as, for example, instructors may feel that assigning poor grades undermines the relationship quality with their students. In contrast to the aforementioned positive value dimensions, *cost* represents the negative value of grading. It may originate from the often non-rewarding nature of the task or the judgement that grading prevents one from doing other, more important tasks, such as course design or research (also discussed in motivation research as so-called opportunity costs; e.g., Gaspard et al., 2015). The importance of each value dimension is most likely dependent on the specific grading situation. For example, teaching a class and having multiple assessments over the semester may provide the possibility for using grades to adapt teaching. Further, seeing students again after grading may boost social values as compared to grading centralized exams implying that one does not even know the examinees.

Country-Specific Differences

Given that we know so little about faculty emotions, there also is a clear lack of comparative research in this field. Yet, there are two studies from the U.S. (Stupnisky et al., 2019a) and Germany (Kordts-Freudinger, 2017) that assessed faculty teaching emotions with single items on comparable scales. A comparison across their findings yielded largely similar results with respect to levels of reported teaching enjoyment, pride, and boredom. Notable differences were that U.S. faculty seemed to experience higher levels of anxiety and frustration and lower levels of satisfaction and anger than German faculty with respect to teaching. As grading has not been specifically considered regarding faculty emotions, there is

so far no evidence on any potential country-specific differences in faculty grading emotions. In the present study, we therefore deliberately recruited samples from the U.S. and Germany, two countries that are similar with regard to socio-cultural aspects, to explore potential differences that may arise across their higher education systems. Based on deliberations about grading in the higher education context in the U.S. and Germany, we suspected that grading occurred under more aversive circumstances in Germany because grading is often completely detached from teaching, or centralized exams on questions potentially outside the expertise of faculty need to be graded. Based on these deliberations, we suspected that German faculty may experience more negative and less positive emotions than U.S. faculty when grading.

Study Aims

Study 2b aimed to investigate six discrete faculty grading emotions (enjoyment, pride, boredom, anxiety, anger, frustration) and their appraisal antecedents from a control-value perspective. To this end we (1) developed multi-item scales to assess grading emotions more reliably than with single-items, (2) identified and measured factors related to control and value appraisals that may be linked with the emotional experiences of faculty when grading, (3) investigated the relative importance of control and value appraisals in predicting the six emotions considered, and (4) made cross-country comparisons with respect to frequency of experienced emotions when grading and the appraisal patterns. Based on the assumptions of control-value theory (Pekrun, 2006), we proposed that to the degree faculty members experience grading as controllable and valuable, they should experience positive emotions during grading. In contrast, the less they feel in control when grading and consider grading as being of negative value, the more negative emotions should arise. Due to a lack of previous research in this context, analyses were deemed largely exploratory.

Method

Procedure

Data was collected through online surveys in the U.S. (Tool: Qualtrics; Qualtrics, 2019) and Germany (Tool: soscisurvey; Leiner, 2019). We asked participants to share their experiences when grading student papers (defined as in Study 2a). After estimating their general grading workload, instructors rated their emotions as well as the control and value appraisal scales regarding one specific student paper they had graded. Overall, participation commitment was high and sample recruitment was effective, thus 84/76% of all participants of the U.S./German sample could be used for analyses. Participants were excluded if either consent for using their data was not given, participants were not involved in grading written student papers regularly or recently, or the survey had been ended early with a completion rate of less than 80 percent of the survey content.

Participants

The final sample consisted of $N = 446$ faculty members in total, of which $n = 245$ were from the U.S. (62.81% female; 87.60% native English speakers) and $n = 201$ were from Germany (52.24% female; 93.03% native German speakers). Most faculty graded student papers regularly (about 80% in both samples), or had done so recently (about 20% in both samples). A notable difference was that U.S. faculty reported to spend around one and a half times more hours on grading per semester ($M = 90.00$, $SD = 119.00$ hrs.) than German faculty ($M = 54.94$, $SD = 53.14$ hrs.). Consistent with this, the U.S. faculty's self-estimated grading workload measured on a scale from 0 (*very little*) to 4 (*very large amount*) was considerably higher than that of the German faculty ($M_{U.S.} = 2.50$, $SD_{U.S.} = 0.84$; $M_{Ger} = 1.85$, $SD_{Ger} = 0.83$).

Most U.S. faculty (aged $M = 47.12$ years, $SD = 10.48$; holding an academic position for $M = 13.32$, $SD = 9.31$ years) were employed full time (86.67%) and either assistant

professors (30.58%), associate professors (26.03%), professors (20.66%), or instructors (20.66%). About half were on tenure track (42.80%) or tenured (41.56%). They were mainly white (92.24%) and came from more than 24 disciplines, the most frequent being medicine and health related (24.07%), education (11.62%), and English (8.30%). U.S. faculty predominantly referred to student papers written in a course they taught themselves (97.95%). Further, most faculty reported grading student papers on undergraduate level (64.20%; Master's level: 23.05%; Doctoral level: 8.23%).

Most German faculty (aged $M = 38.41$, $SD = 10.81$; holding an academic position for $M = 10.50$, $SD = 9.38$ years) were employed full time (67.34%), and either doctoral students (37.19%), professors (30.15%), or postdoctoral researchers (13.57%). About half of them had a temporary employment (51.01%), some were tenured (23.23%), or employed permanently (11.62%). Instructors came from more than 17 disciplines, the most frequent being psychology (29.15%), education (20.60%), and law (16.58%). In this sample, only 74 percent of faculty referred to student papers written in a course they taught themselves. Further, they reported grading student papers on undergraduate (47.26%), Master's (22.93%), or Doctoral level (1.00%). In addition, around one quarter reported about grading experiences in the context of centralized state exams when answering the questionnaire.

Measures

Items were constructed in parallel in English and German through multiple translation-/back-translation by an academic fluent in both English and German, and multiple checks by German and English natives for each of the two language versions, to ensure maximal comparability and linguistic fluency of both language versions. Central measurement properties for all scales are reported in Table 2.2; all scales yielded sufficient to very good internal consistencies in both the U.S. and German sample.

Identical to Study 2a, in a first step emotions were measured on a 5-point rating scale (0 = *not at all*, 1 = *rarely*, 2 = *sometimes*, 3 = *often*, 4 = *very often*) indicating the frequency with which the same six discrete emotions (enjoyment, pride, anxiety, anger, boredom, frustration) were typically experienced when grading. Aside of such a slim, emotion-noun-based assessment of experienced frequency, in this study, we sought to add full-sentence items per emotion, to better capture the richness and complexity of emotional experiences by incorporating different facets of emotions including affective, behavioral/motivational, and physiological emotion indicators (see also, Pekrun et al., 2005). These items were formulated using the Academic Emotions Questionnaire (Pekrun et al., 2005) as a starting point, while adapting them to match the grading task from an instructor's perspective. As a result, three to five items were used to measure the six target emotions.

Control over the grading process and *paper content* were measured with dichotomous items (0 = *no*, 1 = *yes*) indicating the freedom to determine specific aspects when designing the student paper. Because the aspects can be independent of each other, we did not expect them to form uniform scales. Nevertheless, higher sum scores indicated more freedom with respect to the determination of very different aspects of the grading process and the content of the paper and therefore represent higher control appraisals.

Content knowledge, diagnostic competence and all value dimensions (*diagnostic value, utility value, social value, cost*) were measured on a 5-point rating scale (0 = *strongly disagree*, 1 = *somewhat disagree*, 2 = *neutral*, 3 = *somewhat agree*, 4 = s*trongly agree*). Higher values indicated a greater extent of content knowledge, diagnostic competence, and greater importance of the value dimensions. The codebook is available in Appendix B.

Table 2.2

Measurement Properties of Central Study Variables

			Reliability (Cronbach's α)	
Emotion	Items	Sample item [a]	U.S.	Germany
Enjoyment	4	For this student paper I gladly do my grading.	.85	.82
Pride	3	For this student paper I am proud of how well I handle the grading.	.61	.70
Boredom	4	For this student paper I get so bored I have problems concentrating.	.91	.89
Anxiety	5	For this student paper I worry whether I'm able to cope with grading.	.79	.78
Anger	4	For this student paper I get so angry I feel like throwing the papers in the trash.	.81	.82
Frustration	3	For this student paper grading frustrates me because it takes so much time.	.75	.72
Control - grading process [b]	5	For this student paper, do you get to decide the type of questions, e.g., essay, long answer, multiple choice, true/false, etc.?	.72	.66
Control - paper content [b]	3	For this student paper, do you get to formulate the questions yourself?	.47	.75
Content knowledge	2	For this student paper I have good knowledge of the content area.	.90	.82
Diagnostic competence	4	For this student paper I can accurately judge the quality of each paper.	.80	.77
Diagnostic value	4	For this student paper it is important to me that I accurately judge the quality of each paper.	.79	.80
Utility value	2	Grading this student paper is important to me because the results tell me how to adjust my teaching.	.65	.79
Social value	1	Grading this student paper is important to me because the judgements impact my relationships with the students.	–	–
Cost	2	Grading this student paper is a thankless task.	.74	.62

Note. [a] For the emotion scales in addition to the noun-based assessment identical to Study 2a.

[b] These variables were not expected to form uniform scales.

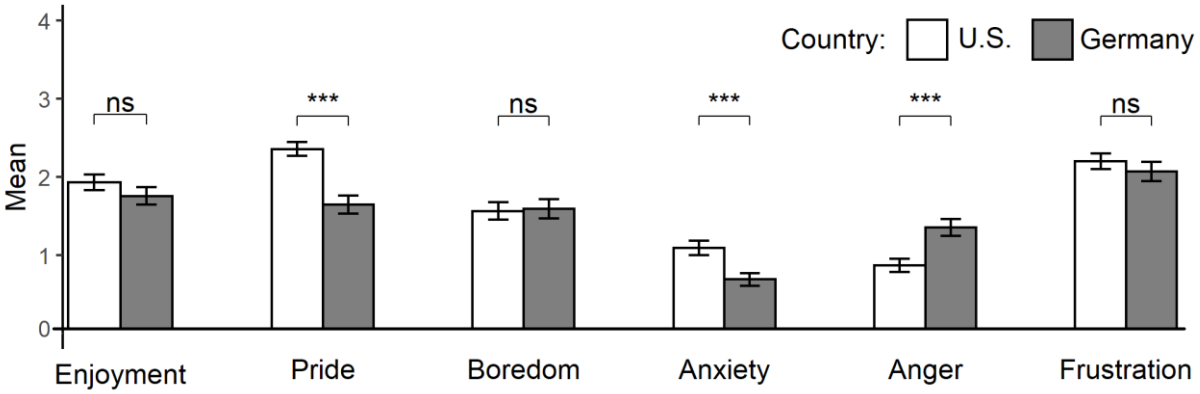
Results

Country Comparisons for Mean Levels of Grading Emotions and Control and Value Appraisals

Data was analyzed using R (R Core Team, 2018). Mean level comparisons across the U.S. and German ratings were analyzed using Welch's independent samples *t*-test[2] (due to multiple comparisons, the alpha level for the pairwise comparisons was adjusted to $p < .001$). Results for grading emotions are depicted in Figure 2.2, exact values and effect sizes in Table 2.3, correlations in Table 2.4. With respect to emotions experienced when grading, U.S. faculty reported substantially higher frequencies of pride and anxiety and lower frequencies of anger. Reported frequencies of enjoyment, boredom, and frustration, were similar across the U.S. and German faculty.

Figure 2.2

Cross-Country Comparison Results for Grading Emotions



Note. Due to multiple comparisons, the alpha level was adjusted to $p < .001$. Error bars represent the 95% CI.

$^{***}p < .001$

[2] Welch's *t*-test is suitable for use with independent samples while not requiring homogeneity of variance ; variances are estimated separately for both samples and the Welch modification to the degrees of freedom is used in case that variances are unequal (Rasch et al., 2011; Ruxton, 2006)

Mean level comparisons of the control and value appraisal ratings across the U.S. and German are depicted in Table 2.2. Overall, findings suggest that U.S. faculty reported to be substantially more in control of their paper's content and grading process, felt more competent, and valued grading more than German faculty, while experiencing the same extent of cost.

Table 2.3

Descriptive Statistics, Independent Samples t*-Tests, Effect Sizes of Cross-Country Comparisons*

	U.S. sample	German sample				Cohen's
	M (*SD*)	*M* (*SD*)	*t*	*df*	*p*	*d* [a]
Enjoyment	1.88 (0.80)	1.71 (0.79)	2.31	428.34	.021	0.22
Pride	2.37 (0.71)	1.64 (0.86)	9.63	386.64	<.001	0.93
Boredom	1.55 (0.91)	1.58 (0.90)	−0.34	428.84	.734	−0.03
Anxiety	1.07 (0.73)	0.66 (0.61)	6.50	442.97	<.001	0.61
Anger	0.84 (0.70)	1.34 (0.80)	−6.92	398.42	<.001	−0.66
Frustration	2.21 (0.82)	2.07 (0.88)	1.68	412.56	.094	0.16
Control - grading process	4.59 (0.95)	2.20 (1.45)	20.01	331.23	<.001	1.94
Control - paper content	2.66 (0.65)	1.89 (1.17)	8.26	298.98	<.001	0.81
Content knowledge	3.77 (0.56)	3.42 (0.60)	6.18	413.22	<.001	0.59
Diagnostic competence	3.39 (0.55)	3.02 (0.58)	6.76	418.25	<.001	0.65
Diagnostic value	3.75 (0.39)	3.45 (0.50)	6.93	375.80	<.001	0.67
Utility value	3.15 (0.73)	2.36 (1.04)	9.03	347.12	<.001	0.87
Social value	2.23 (1.12)	1.24 (1.11)	9.26	427.50	<.001	0.88
Cost	1.98 (1.09)	2.14 (0.93)	−1.67	442.45	.097	−0.16

Note. Due to multiple comparisons, the alpha level was adjusted to $p < .001$. All variables were measured on a scale from 0 to 4, except for the control dimensions which were measured with yes/no ratings (possible max. grading process: 5, possible max. paper content: 3).

[a] Positive numbers indicate higher values for the U.S.

Table 2.4

Study 2b: Correlations Between Study Variables in the U.S. and German Sample

	pride		boredom		anxiety		anger		frustration		ctrl-gp		ctrl-pc		cont. know.		diagn. com.		diagn. val.		utility val.		social val.		cost		age		experience	
	US	GER	US	GER	US	GER	US	GER	US	GER	US	GER	US	GER	US	GER	US	GER	US	GER	US	GER	US	GER	US	GER	US	GER	US	GER
Enjoyment	**.61**	**.53**	**-.52**	**-.48**	**-.24**	-.06	**-.31**	**-.26**	**-.57**	**-.51**	.04	*.15*	.00	.08	.02	.03	*.15*	**.25**	.05	*.14*	.20	**.30**	.06	.22	**-.54**	**-.58**	.17	*-.17*	.06	*-.18*
Pride	—	—	**-.26**	*-.15*	-.08	.02	-.11	-.11	**-.26**	-.09	.01	.09	.06	.03	.18	.09	**.26**	*.15*	.17	.02	**.25**	.20	.20	.11	**-.33**	-.21	.08	-.12	.04	-.21
Boredom			—	—	**.22**	.08	**.42**	**.36**	**.52**	**.44**	.08	-.03	.03	.02	.06	-.10	-.10	-.20	-.05	-.12	-.11	-.23	.11	-.03	**.48**	**.42**	-.18	.01	-.09	.07
Anxiety					—	—	**.42**	.19	**.51**	**.33**	-.06	.05	.07	.02	-.04	**-.39**	**-.27**	**-.43**	-.01	-.18	.04	-.04	**.31**	**.26**	**.36**	.05	-.11	-.20	-.08	-.20
Anger							—	—	**.53**	**.55**	.01	-.04	.09	-.09	.07	-.11	-.08	-.23	.03	-.10	-.02	-.01	*.16*	.11	**.40**	.22	-.11	-.02	-.02	-.06
Frustration									—	—	.03	*-.15*	.05	-.10	.03	-.12	*-.14*	**-.27**	.06	-.06	-.02	-.08	.18	.00	**.59**	**.42**	-.11	.04	-.05	.01
Control grading process											—	—	**.51**	**.47**	.11	.13	*.15*	.01	.04	*-.14*	*.13*	**.24**	*-.13*	*.15*	.06	-.08	.05	.04	.01	.05
Control paper content													—	—	**.31**	.20	*.16*	.22	.04	.01	.18	**.40**	-.08	**.24**	.04	-.14	.03	.19	.01	.12
Content knowledge															—	—	**.46**	**.40**	**.26**	.21	*.15*	.22	.03	-.09	.05	-.09	.04	.20	.09	*.18*
Diagnostic competence																	—	—	**.37**	**.54**	.12	**.26**	-.03	.01	-.16	**-.23**	.06	.23	.10	.21
Diagnostic value																			—	—	**.24**	.22	.05	.01	-.02	-.10	.02	.08	.08	.06
Utility value																					—	—	.19	**.30**	.18	**-.26**	.03	*.14*	-.01	.09
Social value																							—	—	.05	*-.15*	.06	-.02	.06	-.04
Cost																									—	—	*-.16*	-.04	-.08	.01
Age																											—	—	**.75**	**.87**

Note. For easy comparison, U.S. and German coefficients are placed adjacent to each other.

Coefficients in bold are significant at $p < .001$, underlined coefficients at $p < .01$, and coefficients in italics at $p < .05$.

Appraisal Patterns of Emotions for Grading

We performed separate multiple regression analyses for each emotion as predicted by the control and value appraisal dimensions. Years of experience in academia were included as control variable. It proved largely unrelated to the reported grading emotions, with the exception of years in academia being negatively linked with enjoyment and pride in the German sample. Table 2.5 depicts the detailed results of those multiple regressions. A key overarching finding across all discrete grading emotions considered in the present study was that perceiving grading as an aversive task was the most prominent factor determining the emotional experiences of faculty when grading: Cost in the sense of negative value (i.e., perceiving the task of grading as thankless and experiencing that it keeps one from engaging in other, more meaningful or important tasks) was associated with lower frequencies of all positive emotions and higher frequencies of all negative emotions. Furthermore, across the U.S. and German samples, social value was positively and diagnostic competence negatively linked with experiences of anxiety, whereas diagnostic competence was positively linked with experiences of pride during grading.

Otherwise, the predictive patterns differed across samples. Social value emerged as second most important predictor for the U.S. sample, also being linked positively with pride and frustration, which is in line with predictions of the value component of CVT. Diagnostic competence emerged as second most important predictor in the German sample in that higher diagnostic competence was positively linked with all positive and negatively linked with all negative emotions considered in this study. This pattern of findings is in line with the control component of CVT.

Table 2.5

Results of Separate Multiple Regressions for Each Emotion in Both Samples

	Enjoyment		Pride		Boredom		Anxiety		Anger		Frustration	
	U.S.	GER	U.S.	GER	U.S.	GER	U.S.	GER	U.S.	GER	U.S.	GER
Control – grading process	.08	.15*	.01	.07	.08	.00	−.10	−.01	−.05	−.01	−.04	−.18*
Control – paper content	−.04	−.20**	−.02	−.17	−.05	.22**	.16*	.07	.11	−.10	.08	.03
Content knowledge	.07	−.10	.16*	.04	.08	−.02	−.06	−.23**	.01	.00	−.07	.04
Diagnostic competence	.06	.29***	.15*	.22*	−.05	−.20*	−.24***	−.39***	−.06	−.24*	−.07	−.30**
Diagnostic value	−.04	−.04	.03	−.16	−.01	.07	.07	.07	.08	.03	.09	.11
Utility value	.06	.16*	.08	.16	−.04	−.20*	.07	−.01	−.03	.12	.05	.07
Social value	.07	.06	.19*	.00	.09	.04	.28***	.20**	.13*	.16	.15**	.09
Cost	−.51***	−.47***	−.29***	−.13	.45***	.31***	.32***	−.04	.39***	.16*	.59***	.35***
Years in academia	.01	−.22***	−.01	−.24**	−.06	.11	−.06	−.09	.00	.01	−.02	.09
R^2	.30	.46	.22	.15	.24	.23	.29	.31	.19	.12	.39	.24

Note. Regression analyses in the U.S. sample were run with $n = 220$ and in the German sample (GER) with $n = 171$ participants. Reported are standardized βs.

*$p < .05$. **$p < .01$. ***$p < .001$.

Discussion

The key findings of Study 2b were that U.S. faculty, compared to German faculty, reported more pride, anxiety, control, competence, positive value, and less anger with respect to grading. Levels of reported enjoyment, boredom, frustration, and cost were similar across the two countries' samples. Furthermore, across both samples, grading emotions were mainly influenced by the extent to which faculty felt it came at a cost; that is, the more aversive they perceived grading to be, the more negative and the less positive emotions they reported to experience. In addition, there were some noteworthy variations of appraisal patterns across countries. For example, social value was an important predictor of grading emotions in the U.S. sample, but not in the German sample; alternatively, diagnostic competence was another important predictor of grading emotions in the German but not in the U.S. sample.

Country Comparisons for Mean Levels of Emotions, Control and Value Appraisals

Results of the cross-country comparisons on grading emotions partially deviated from those of teaching emotions, which were based on the comparison between Stupnisky et al. (2019a; U.S.) and Kordts-Freudinger (2017; Germany). In line with the findings for teaching emotions, U.S. faculty experienced more anxiety, less anger, and similar levels of enjoyment and boredom when grading compared to German faculty. While U.S. faculty experienced more frustration and similar levels of pride compared to German faculty when teaching, they experienced similar levels of frustration and more pride when grading.

Looking at the appraisal antecedents of grading emotions, U.S. faculty generally experienced higher levels of control, competence, and positive value than German faculty. We propose that this pattern of findings can be explained by the differing circumstances under which grading occurred in the two countries' samples. Almost all U.S. faculty (97.95%) graded papers of students enrolled in their classes and may even have graded multiple assignments over the course of a semester, therefore being familiar with the content

and the students. Moreover, they could determine most aspects of the grading situation. In contrast, almost 25 percent of the German faculty reported their emotional experiences with regard to grading state exams, which has various implications: 1) the corresponding exam tasks are provided by the ministry and some faculty might be unfamiliar with some of the questions, 2) time for grading is typically very limited and may therefore not allow faculty to fulfill their own grading standards, and 3) grading involves scoring anonymized exams of students from other universities unknown to the faculty. These aversive circumstances under which one quarter of the German sample responded may explain the lower mean scores regarding control, competence, and value with respect to grading in the German sample.

Relevance of Specific Appraisals for Grading Emotions

Results from the regression analyses were mostly in line with predictions from CVT (Pekrun, 2006), although not all facets of control and value were as influential as expected. Subjective control in terms of competence is one key predictor for emotions according to CVT (Pekrun, 2006), which influences the quality and intensity of an emotion. This was reflected in the importance of diagnostic competence in the German sample: higher diagnostic competence predicted higher levels of all positive and lower levels of all negative emotions. This finding is also in line with qualitative insights which revealed that negative grading emotions were associated with a lack of competence (Hagenauer & Volet, 2014; Lahtinen, 2008; Postareff & Lindblom-Ylänne, 2015). It is surprising, however, that diagnostic competence influenced only two out of six emotions (pride and anxiety) in the U.S. sample and that content knowledge influenced only one emotion in each sample (pride in the U.S., anxiety in Germany). One explanation may be that the means of diagnostic competence and content knowledge were rather high especially in the U.S. sample and therefore ceiling effects may have occurred. Alternatively, a lack of content knowledge may not be of such great importance when grading because it can be compensated through effort

when there is enough time (e.g., reading up on the exam content). Also, there is no direct interaction with students or colleagues in the context of grading, so corresponding failures may go unnoticed. The extent of control over the paper content and the grading process may not have been as important predictors for grading emotions because they may not have reflected the subjective judgement of whether the external regulations impair faculty's subjective feelings of control very well.

According to CVT (Pekrun, 2006), value appraisals are another important predictor for emotions, influencing the valence and intensity of an emotion. In the context of emotions for grading, negative value (i.e., cost) seemed to be more influential than positive value because it predicted almost all emotions in both samples comparably strongly. In line with predictions from CVT, higher cost was associated with lower frequencies of positive and higher frequencies of negative emotions. This implied that the grading activity was generally appraised as undesirable and therefore elicited rather negative emotional experiences. This strong effect may have undermined possibly smaller effects of positive value appraisals. Nevertheless, social value additionally emerged as predictor of four out of six emotions (pride, anxiety, anger, frustration) in the U.S. sample and of anxiety in the German sample, which implied that the more instructors felt that their grading had an effect on the relationship with students, the more frequently they experienced emotions during grading, above and beyond the costs of grading. This showed that the relationship with students shaped the emotional experiences of faculty even outside the classroom. Further in line with CVT, the higher the utility value of grading, the more joy and less boredom was reported in the German sample, but these effects were small and did not replicate across other emotions and the U.S. sample. However, this finding gives a hint that faculty perceived grading as useful, with positive consequences for the emotional experiences during grading. It should be noted that the scale used in this study has not yet fully covered all possible facets of utility that

could be attributed to grading. Surprisingly, diagnostic value did not predict any of the emotions. This may have been due to the very high means and potential ceiling effects in both samples, which implied that although grading is an aversive task, all faculty deemed it important to do well on it, in terms of assigning reliable, fair, and transparent grades.

Taken together, we observed considerable country differences in mean levels of control and value appraisals which were clearly more unfavorable in the German sample as judged from the perspective of CVT (Pekrun, 2006). Intriguingly, those appraisal differences did not fully reflect in corresponding country differences in all grading emotions: while the levels of enjoyment, boredom, and frustration were rather similar across both countries, U.S. faculty reported considerably more pride and anxiety, and less anger than their German colleagues. One explanation for the similarity of reported enjoyment, boredom, and frustration despite differing contexts may be that emotions were most strongly influenced by costs – the only appraisal dimension in which U.S. and German faculty did not differ. Furthermore, emotion ratings predominantly result from internal comparisons (e.g., how much do I enjoy grading vs. writing a paper?), or external comparisons within one's known reference group and cultural context (e.g., do I enjoy grading more than my colleagues?). As a result, even though the grading circumstances were comparably "benign" in the U.S. context, faculty's emotional ratings of this task were still rather negative.

Study Limitations

This research is the first to systematically quantitatively examine faculty grading emotions from a control-value perspective in two different countries with high socio-cultural similarities, yet noticeable differences in the realities of their academic contexts: the U.S. and Germany. Limitations among the presented studies included that some of the newly developed scales showed some deficiencies with respect to psychometric measurement quality. Some scales showed relatively poor reliabilities (utility value and pride in the U.S.,

cost in the German sample) and in the U.S. sample some scales showed high means (i.e., potential ceiling effects), which may imply that the self-report measures used were not able to distinguish between individuals located high on these dimensions, and that the restricted variance undermined finding empirical support for any covariances with the emotion scales. Despite these shortcomings, we believe the scales' psychometric properties were sufficient to provide a first look into antecedents of faculty grading emotions in a structured way. Nevertheless, future research that focuses on refined scales to measure both grading emotions and their control and value antecedents more accurately, and replication of the findings presented herein, will be necessary to substantiate the conclusions drawn from this research.

Implications

It has been argued previously that it is meaningful to consider faculty emotions as they emerge in different contexts of the diversified profession of university faculty (specifically, research versus teaching; e.g., Stupnisky et al., 2019a). That line of research has provided compelling, and to some degree concerning, evidence that the context of research is experienced as emotionally more aversive than the context of teaching. In the present research, we replicated those findings and additionally proposed to yet consider another very specific task of faculty, namely grading. Our key finding here was that faculty experience the task of grading as even more emotionally aversive than research. Our data showed that these predominantly negative emotional experiences seemed to be driven by the fact that faculty members generally attached high costs to the task of grading and that some faculty additionally seemed to doubt their diagnostic competencies, and thus experienced enhanced anxiety, boredom, anger, and frustration during grading, and reduced joy and pride. Acknowledging that grades likely are biased by emotions (Brackett et al., 2013), this may imply that some students receive unduly poor grades, with resulting undue disadvantages on the labor market. From this perspective, it seems imperative that faculty are supported in

optimizing their emotional experiences during grading, but also from the perspective of university administrators or faculty recruiting staff who aim at optimizing faculty well-being. Based on our findings, we see two potential pathways for optimizing faculty experiences. First, universities could make an effort to – more than is already the case – acknowledge large grading workload, and reward efforts in designing psychometrically sound as well as learner- and learning-oriented exam formats, for example in terms of "grading awards" in addition to the traditional "teaching awards." Secondly, universities may want to make an effort to optimize the contexts for grading by increasing faculties' competence beliefs with respect to the task of grading, for example by offering opportunities for training on how to design exams (see e.g., https://writingproject.fas.harvard.edu/files/hwp/files/hwp_brief_guide_assignments.pdf). Ideally, such policies could help to reduce faculty appraisals of the task of grading as aversive and thankless, and thus eventually lead to more productive emotions and more reliable and fair grades.

Studies 2a+b contributed to understanding faculty grading emotions and Study 2a tapped at teaching emotions but only within a context-comparison (as compared to grading and research). As teaching is an integral part of faculty members' duties, the emotions experienced therein warrant research. And although faculty emotions in teaching have been comparably well-researched, there is one form of teaching that received little research attention with respect to emotions thus far: online teaching. The onset of the COVID-19 pandemic, however, brought immediate attention to this constantly-growing form of higher education.

4. Study 3

“I’m Tired of Black Boxes!”: Faculty Teaching Emotions in Emergency Online and Face-to-Face Teaching

Although online teaching and learning opportunities at universities have increased over the last decade, the majority of lectures and seminars were still offered in classical face-to-face teaching settings (Filak & Nicolini, 2018; Salikhova et al., 2020). The onset of the COVID-19 pandemic in early 2020, however, forced universities worldwide to ad hoc shift their teaching to purely digital online environments (Marinoni et al., 2020). While this forced shift boosted digitalization in the higher education context, it also imposed great challenges for students who needed to demonstrate very high self-managerial and self-directed learning skills to master digital learning (Sun & Rueda, 2012) and for faculty[3] who needed to quickly re-think and adapt their established ways of teaching to transfer their classes to online environments. Given the acknowledged importance of emotions in higher education (Pekrun, 2019), as of today, surprisingly little is known about the emotional experiences of students learning online, and even less so about faculty teaching online (Naylor & Nyanjom, 2020; Valverde-Berrocoso et al., 2020). The radical change to online education during the onset of a pandemic was no pleasant experience for students (Aristovnik et al., 2020; Besser et al., 2020; Garris & Fleck, 2020; Padrón et al., 2021), but again insights into faculty teaching experiences during the onset of the pandemic are highly limited. To narrow this knowledge gap, the present study aimed to compare emotional experiences of faculty teaching synchronous online classes using an online meeting tool during a time of pandemic against those teaching face-to-face classes before the pandemic adopting self-determination theory (SDT) as theoretical framework.

[3] By faculty or faculty member we refer to all individuals teaching at a higher education institution irrespective of the held degree or exact position, thereby also including for instance doctoral students and external lecturers.

SDT and Faculty Emotions in Face-to-Face Teaching

Self-determination theory has been used frequently as a framework to inform research in educational settings. SDT proposes that an individual’s motivation and well-being depend on the satisfaction of three basic psychological needs: the needs for autonomy, competence, and relatedness. Individuals experience autonomy when they feel like they themselves cause their actions, which should align with their self, competence when they successfully use their skills to interact with their environment, and relatedness when they sense a connection and mutual caring with others (Deci & Ryan, 2004). Ample research in diverse settings and populations showed that the satisfaction of the basic needs is not only beneficial for autonomous forms of motivation (e.g., Stupnisky et al., 2018) and well-being (e.g., Milyavskaya & Koestner, 2011), but also for further outcomes, such as achievement and engagement (e.g., Jang et al., 2009), satisfaction (e.g., Crick et al., 2020), affect (e.g., Ebersold et al., 2019; Holzer, Lüftenegger, Käser, et al., 2021; Sheldon & Filak, 2008), and discrete emotions (e.g., Klassen et al., 2012).

Faculty have a profession that comprises highly diverse tasks, such as teaching, research, and administration, which may trigger different emotional reactions. Grounded in the context-specificity of emotions (e.g., of school teachers; Frenzel et al., 2015, 2016), it has been shown that faculty experienced varying levels of discrete emotions while engaging in different tasks. Generally, compared to research, university teaching in face-to-face settings triggered more positive emotions, such as enjoyment and pride, and fewer negative emotions, such as anxiety and frustration (Stupnisky et al., 2016). Quantitative and qualitative findings on faculty emotions when teaching face-to-face classes revealed that faculty reported positive emotions, such as enjoyment and pride more often than negative emotions, such as boredom, anger, and frustration (Kordts-Freudinger, 2017; Postareff & Lindblom-Ylänne, 2011; Thies & Kordts-Freudinger, 2019; Trigwell, 2012).

Bringing the two lines of research together, SDT offered insights into potential antecedents of faculty emotions in face-to-face teaching settings. Although the results were not fully consistent, especially when it came to the relative importance of autonomy, competence, and relatedness, the bigger picture showed that the satisfaction of one or multiple psychological needs had positive relationships with or effects on faculty members’ intrinsic motivation (Esdar et al., 2016; Stupnisky et al., 2018), positive emotions (Hagenauer & Volet, 2014; Löfström & Nevgi, 2013), as well as general satisfaction and teaching satisfaction (Crick et al., 2020; Larson et al., 2019; Seipel & Larson, 2018). In the same line, research on school teachers showed that autonomy, competence, and relatedness were positively associated with positive affect and enjoyment, and negatively with negative affect and negative emotions, such as anxiety and anger during face-to-face teaching (Ebersold et al., 2019; Hagenauer et al., 2015; Klassen et al., 2012; Russo et al., 2021). Taken together, in face-to-face teaching stronger positive than negative emotions were experienced and higher levels of need satisfaction were typically associated with more positive and less negative emotions.

SDT and Emotions in Online Education

Although online teaching offers steadily increased over the past years (Filak & Nicolini, 2018), SDT has been used frequently as a framework in educational settings, and the importance of emotions in higher education has been acknowledged (Pekrun, 2019), the insights into both need satisfaction and emotional experiences in online education are limited, especially with regard to faculty (Naylor & Nyanjom, 2020; Valverde-Berrocoso et al., 2020). To the best of our knowledge, to date no study explicitly addressed faculty need satisfaction in online teaching. One study, however, reported reduced satisfaction with the interaction between faculty and students during emergency online education during the COVID-19 pandemic (Kanning & Ohlms, 2021). Moreover, technical problems, low levels

of student involvement in class, and missing face-to-face contact with students were named frequently as negative aspects of online teaching (Wasilik & Bolliger, 2009), which may be considered as factors influencing need satisfaction. A few studies touched on faculty teaching emotions in online education before the pandemic and the radical transition of almost all face-to-face teaching to online teaching during the onset of the COVID-19 pandemic elicited some more research into the emotional experiences of faculty teaching during these unprecedented times. Many of the results obtained shortly after the onset of the pandemic need to be interpreted with caution, though, because they often reported single case studies, reflections of few individuals, or were lacking a control group. This being said, faculty reported experiencing a variety of different positive and negative emotions during the transition to online teaching depending on their attitudes, abilities, and institutional support (Naylor & Nyanjom, 2020). Even during emergency online teaching, faculty seemingly experienced positive emotions more strongly than negative emotions (Meishar-Tal & Levenberg, 2021), which is in line with findings from face-to-face teaching (e.g., Thies & Kordts-Freudinger, 2019). Nevertheless, the level of enjoyment was clearly reduced in online compared to face-to-face teaching (Kanning & Ohlms, 2021). Negative emotions such as stress and frustration were reported to stem from the use of technology or problems with it, a lack of pedagogical skills, and an inability to adjust teaching based on student reactions like facial expressions and postures, due to a lack of visual feedback in online teaching (Downing & Dyment, 2013; Regan et al., 2012). First correlational findings align with these statements because perceived competence in online teaching was positively related to positive emotions and negatively related to negative emotions (Meishar-Tal & Levenberg, 2021). Furthermore, relatedness with students was named as a factor contributing to teaching satisfaction (Downing & Dyment, 2013) and teaching satisfaction was lower in online teaching during compared to face-to-face teaching before the pandemic (Kanning & Ohlms, 2021). To

conclude, the few available findings hint at reduced levels of competence (e.g., technical problems), relatedness (e.g., missing face-to-face contact), and teaching satisfaction, and a relation between faculty need satisfaction and emotions as well as teaching satisfaction in online teaching.

To complement the limited findings on faculty experiences in online settings, we drew on university students’ experiences in online learning to further inform our reasoning. Undergraduate students enrolled in online classes reported lower overall need satisfaction and higher need dissatisfaction than students enrolled in face-to-face classes (Wang et al., 2019). More specifically, irrespective of whether students judged their online experiences retrospectively against their own face-to-face experiences or students taking the same hybrid class online were compared to their on-campus counterparts, students consistently reported similar levels of perceived autonomy but reduced levels of relatedness in online learning (Butz et al., 2014; Butz & Stupnisky, 2016; Filak & Nicolini, 2018; Otter et al., 2013). In retrospective judgements, perceived competence was reduced in online classes compared to face-to-face classes (Filak & Nicolini, 2018), whereas the group comparisons did not show differences with respect to perceived competence (Butz et al., 2014; Butz & Stupnisky, 2016). Studies on the relation between need satisfaction and discrete emotions showed that self-efficacy (closely linked to perceived competence) predicted hope positively and anxiety and frustration negatively in both online and face-to-face learning (Marchand & Gutierrez, 2012). Autonomy and relatedness predicted enjoyment positively and boredom negatively in students attending massive open online courses (MOOCs; Buhr et al., 2019) and competence most strongly influenced positive emotions in online learning during the COVID-19 pandemic, generally lending support for the applicability of SDT in online learning.

Taken together, research on online teaching and learning showed that faculty members’ need satisfaction especially of competence and relatedness may be thwarted in

online compared to face-to-face settings, similar to students’ experiences. Overall, previous findings support the applicability of SDT to emotion research in online settings.

Faculty Need Satisfaction in Emergency Online Teaching During a Time of Pandemic

Before trying to outline factors that may affect need satisfaction in online teaching, which in turn were assumed to influence emotional experiences and teaching satisfaction, it must be noted that online teaching before and during a time of pandemic may differ. While faculty teaching online before the pandemic had typically chosen the format voluntarily, faculty teaching online during the COVID-19 pandemic were forced to abruptly switch to the online format — a shift for which most were not or only partially ready (R. Scherer et al., 2021). And although the transition to online teaching has been associated with challenges no matter whether it occurred before or during the pandemic (Downing & Dyment, 2013; Kuladinithi et al., 2020; Regan et al., 2012; Rusly et al., 2021; Wasilik & Bolliger, 2009; Yarmand et al., 2021) and results may therefore likely be generalized, the present study focused on emergency online teaching that had to be implemented ad hoc during the onset of a pandemic. Therefore, we try to outline the exceptional circumstances that may have affected faculty members’ need satisfaction in emergency online teaching during the onset of a pandemic.

Regarding the satisfaction of the need for autonomy, a lack of freedom in determining the content, activities, or policies in class has been identified as thwarting perceived autonomy of graduate teaching assistants in face-to-face teaching (Kajfez & Matusovich, 2017). The rapid shift from familiar face-to-face to unfamiliar online teaching may well have impaired faculty members’ perceived autonomy because the new format may not have aligned very well with their ideas about teaching. Furthermore, they may have lacked the pedagogical skills to implement effective online teaching according to their beliefs (Downing & Dyment, 2013). German faculty, howver, generally have a lot of freedom with respect to

teaching and may adapt the content and design of a class to fit their expertise and personal demands. Even during the transition to an online format faculty were rather autonomous because they could decide for each of their classes how to implement it online. To name just a few options, faculty may have chosen to deliver a lecture synchronously through an online meeting tool, to record all sessions and make them available as asynchronous offers, or to provide relevant content in text-based self-learning units. Further, they may have chosen to turn seminars, tutorials, or similarly interactive and student-centered classes into a lecture-style class to focus on content delivery or to create an interactive class by using for instance different features of synchronous online meeting tools (e.g., polls, breakout rooms, etc.), collaborative online tools (e.g., etherpads, online mind maps, etc.), and learning platforms (e.g., forums, wikis, etc.). Therefore, given these manifold options, faculty presumably chose the teaching approach that fit their own preferences, teaching conceptions, and competencies best and therefore their perceived autonomy should not have been impaired too severely because the transition to online teaching allowed for control and agency.

Regarding the satisfaction of the need for competence, training and previous experiences have been identified as factors positively influencing teaching competence in graduate teaching assistants teaching face-to-face (Kajfez & Matusovich, 2017). During the onset of the pandemic, most faculty lacked experience and training in online teaching, whereby their perceived competence was likely impaired (Downing & Dyment, 2013). With respect to online teaching, the concept of competence may be conceived broader than in face-to-face teaching because faculty not only needed to master the manifold teaching task itself and acquire new didactical concepts to teach content, but also master the new digital tools and acquire technical methods to implement online teaching in the first place. And indeed, faculty reported technical problems and issues with engaging students in class discussions as challenges in online teaching (Rusly et al., 2021; Wasilik & Bolliger, 2009). Therefore, the

unprepared shift from familiar face-to-face to new online teaching probably impaired faculty members’ perceived competence (Downing & Dyment, 2013). As both pedagogical and technical skills have been reported to influence perceived competence to teach online (Downing & Dyment, 2013), it is likely that faculty who experienced more technical problems perceived lower levels of competence because they could not deliver classes as planned.

Regarding the satisfaction of the need for relatedness, a positive relationship with students contributed to the satisfaction of the need for relatedness during teaching more strongly than a positive relationship with colleagues (Kajfez & Matusovich, 2017; Klassen et al., 2012). This hints at the importance of relatedness within the teaching context. The physical distance between faculty and their students, as well as among students, is the probably most outstanding difference between online and face-to-face teaching and was stated as a negative aspect of online teaching (Wasilik & Bolliger, 2009). Whereas content may have been covered in similarly effective ways especially in online lectures (Euzent et al., 2011), for which lecture recordings had been rather common even before the pandemic, many of the interactions in which faculty and students engaged in before, during, and after class in face-to-face teaching may not have been resembled sufficiently in online environments to build relationships and experience a sense of belonging and mutual caring (Yarmand et al., 2021). Therefore, faculty very likely perceived their relatedness with students being severely impaired in online compared to face-to-face teaching, just as students perceived their relatedness with faculty being impaired (Filak & Nicolini, 2018). Some implementations of online teaching, such as synchronous online meetings, bear the potential to foster relatedness more than others, such as mainly asynchronous self-learning units. This assumption aligns with ideas of social presence theory (Lowenthal, 2010) and related findings which showed that asynchronous videos helped university students to perceive their instructor and peers as

real persons and to feel closer to them (Borup et al., 2012), that is to foster relatedness. But even within synchronous classes via online meeting tools that allow for quasi-live experiences in real time, the impression of interacting with real students within such online classes may depend on whether faculty could see (i.e., cameras are activated) and hear (i.e., audio is turned on for communication) their students rather than teaching black boxes and answering questions from a chat window. This idea was corroborated by faculty reporting that a lack of (visual) feedback from their students due to mainly turned off cameras and passive and non-responsive students made it hard for them to get a feeling for their class, to know whether students could follow or not (Regan et al., 2012; Rusly et al., 2021; Wasilik & Bolliger, 2009; Yarmand et al., 2021), that is to connect with their students and feel related. Therefore, it is likely that the more students were visible to faculty during an online class, the more related they felt to their students.

To sum it all up, we conclude that SDT is a suitable framework to investigate faculty emotions in face-to-face teaching settings and very likely in online teaching settings, due to the needs’ basic and universal nature (Deci & Ryan, 2004) and first hints at SDT’s applicability in online teaching (Meishar-Tal & Levenberg, 2021). In face-to-face teaching settings, faculty experienced various mainly positive emotions, whereby the satisfaction of the three basic psychological needs for autonomy, competence, and relatedness increased the experience of positive affect and discrete emotions, such as enjoyment, and decreased the experience of negative affect and discrete emotions, such as anxiety and anger (Ebersold et al., 2019; Hagenauer & Volet, 2014; Klassen et al., 2012; Löfström & Nevgi, 2013; Russo et al., 2021). Some first studies on online teaching before and during the COVID-19 pandemic suggested that the transition to online teaching was challenging due to factors that could be associated with autonomy, competence, and relatedness (Downing & Dyment, 2013; Kuladinithi et al., 2020; Regan et al., 2012; Rusly et al., 2021; Wasilik & Bolliger, 2009;

Yarmand et al., 2021), which may hint at decreased need satisfaction and therefore less positive and more negative emotions in online compared to face-to-face teaching. None of the available studies systematically compared online teaching experiences during the pandemic to experiences in face-to-face teaching before the pandemic. Therefore, it remains to be explored whether faculty members’ teaching experiences differ between emergency online and face-to-face teaching.

The Present Study

Grounded in SDT, this study aimed to compare the satisfaction of the three basic psychological needs for autonomy, competence, and relatedness, as well as discrete teaching emotions and teaching satisfaction of faculty teaching synchronous online classes during the COVID-19 pandemic to faculty teaching face-to-face before the pandemic. To this end, we obtained faculty data from before the pandemic from Daumiller and colleagues (2019) and replicated their diary-design for the data collection during the onset of the COVID-19 pandemic to compare faculty experiences between the two samples.

Drawing on research of faculty members’ and university students’ experiences in online teaching and learning and deliberations about the exceptional circumstances when face-to-face teaching was ad hoc shifted to online teaching in a time of pandemic, we expected faculty indicating comparable levels of satisfaction of the need for autonomy, a reduced satisfaction of the need for competence, and a clearly reduced satisfaction of the need for relatedness in emergency online teaching compared to face-to-face teaching. Based on the positive and negative relationships between need satisfaction and positive and negative emotional experiences in teaching, respectively
(Ebersold et al., 2019; Hagenauer & Volet, 2014; Klassen et al., 2012; Löfström & Nevgi, 2013; Meishar-Tal & Levenberg, 2021; Russo et al., 2021), and the expectedly reduced levels of need satisfaction, we furthermore expected less favorable emotional experiences (less joy

and pride, more boredom, anger, anxiety, and shame) and lower levels of teaching satisfaction in online compared to face-to-face teaching.

Based on the deliberations and findings about different aspects of an online teaching environment that may influence need satisfaction, we expected that more technical problems predicted fewer satisfaction of the need for competence. Furthermore, we expected that online environments allowing more for quasi-live video-based interaction (quantified by the average number of participants sharing their videos during a session) predicted higher satisfaction of the need for relatedness.

Method

This study including the reported hypotheses has been preregistered (https://aspredicted.org/blind.php?x=fa6ur8) before the start of data collection of the sample teaching online during the COVID-19 pandemic. The data and analysis script will be made accessible through OSF upon publication.

Procedure and Measures

We obtained data collected in the context of a different study (Daumiller et al., 2019; PsyArXiv: p4nhu) before the pandemic (Sample 1, teaching face-to-face) and replicated the survey design to collect corresponding data during the pandemic (Sample 2, teaching online): faculty of both samples were asked to complete a basic questionnaire pertaining to their working conditions before the pandemic with respect to basic need satisfaction (German adaptation of Sheldon & Hilpert, 2012) and self-efficacy (German adaptation of Nie et al., 2012; faculty of Sample 2 did so retrospectively while the pandemic had already set in), and to indicate their current stress at work (Schulz & Schlotz, 1999). Sample 2 was additionally asked to report basic need fulfillment, teaching satisfaction, and the positive valence of teaching in the current time of pandemic as judged against their own experiences before the pandemic (adapted from the session-specific questionnaire items). Subsequently, all faculty

were asked to choose one of their classes and to fill in a session-specific questionnaire ideally three to six times directly after having taught the indicated class (online classes had to be taught synchronously using an online meeting tool). The session-specific questionnaire tapped at basic need fulfillment (adaptation of Janke & Dickhäuser, 2018), discrete emotions (based on Goetz et al., 2016), and teaching satisfaction (self-developed by Daumiller et al., 2019). Sample 2 additionally indicated technical aspects of their online environment, such as the number and approximate time fraction of activated student cameras during the session. This information was used to calculate the number of on average visible students across a session. Measurement properties and example items of all central study variables are depicted in Table 3.1. The codebook is available in Appendix C.

Sample Description

For the purpose of this study, participants who had answered the basic questionnaire only were excluded from further analyses, which resulted in $N = 172$ faculty from different German universities in total. Sample 1 comprised $n = 101$ faculty (52.81% female; aged $M = 40.01$, $SD = 10.42$; work experience of $M = 9.33$, $SD = 7.84$ years; obtained from Daumiller et al., 2019) who taught face-to-face in classrooms before the pandemic. Sample 2 comprised $n = 71$ faculty (63.33% female; aged $M = 39.57$, $SD = 10.95$; work experience of $M = 9.01$, $SD = 8.28$ years) who taught online during one of the first academic terms within the first year of the COVID-19 pandemic (between March 2020 and March 2021) by offering synchronous classes using an online meeting tool (e.g., AdobeConnect, BigBlueButton, WebEx, Zoom), which is the form of online teaching that resembles face-to-face teaching most closely. We explicitly did not focus on flipped classroom settings or asynchronous offers implemented through online learning platforms, because they are too different from face-to-face teaching to make direct comparisons.

Table 3.1

Number of Items, Sample Items, and Cronbach’s Alpha for Both Samples of All Study Variables

	No.	Sample item	Cronbach’s α in Sample 1	2
Basic questionnaire				
Experiences before the pandemic (rated retrospectively by Sample 2)				
Basic needs		Typically, in my teaching …		
Autonomy [a]	6	I am free to do things my way.	.67	.76
Competence [a]	6	I also master difficult things well.	.73	.77
Relatedness [a]	6	I feel close and connected to colleagues who are important to me.	.78	.76
Self-efficacy [a]	9	Typically, in your teaching, how well do you accomplish to … use varied teaching methods?	.83	.82
Experiences during time of data collection				
Stress at work [b]	8	How often did you experience times when you had too many commitments to fulfill?	.94	.94
Technical problems [a]	4	There are technical problems all the time.	–	.84
Teaching experiences during the pandemic judged against prior experiences (Sample 2 only)				
Basic needs		Compared to my typical experiences in non-online teaching so far, I feel like …		
Autonomy [c]	2	I can determine how I design my teaching.	–	.83
Competence [c]	2	I can handle my teaching well and competently.	–	.86
Relatedness [c]	2	I feel like I'm socially connected.	–	.60
Teaching satisfaction [c]	1	I'm satisfied with my teaching.	–	–
Positive valence [c]	1	the teaching is pleasant.	–	–
Session-specific questionnaire				
Emotions		In today's session, I experienced ...		
Enjoyment [a]	1	enjoyment	–	–
Pride [a]	1	pride	–	–
Boredom [a]	1	boredom	–	–
Anger [a]	1	anger	–	–
Anxiety [a]	1	anxiety	–	–
Shame [a]	1	shame	–	–
Teaching satisfaction [a]	1	Overall, I am satisfied with today's session.	–	–
Basic needs		In today's session, I felt …		
Autonomy [a]	2	able to act autonomously.	.96	.88
Competence [a]	2	like I was competent.	.94	.68
Relatedness [a]	2	close and connected to my students.	.91	.89

Note. No. = Number of items. [a] 8-point agreement scale (1 = *no agreement*, 8 = *full agreement*). [b] 5-point rating scale (1 = *never*, 5 = *very often*). [c] 9-point semantic differential (−4 = *less* (i.e., worse during the time of pandemic), 0 = *equal*, 4 = *more* (i.e., worse before the pandemic).

Outlier analyses revealed that according to Cook’s distance measure, some of the cases could be considered as outliers with regard to at least one of the outcome variables. Because the response behavior of those participants did not show any unusual patterns, however, and it would have been a random decision to determine an outlier with respect to self-reported experiences, we decided to retain those a little more extreme cases in the final sample. The data structure showing which questionnaire had been answered how often is depicted in Table 3.2.

Statistical Analyses

For the analyses in the present study, the answers from the session-specific questionnaires were aggregated across all available sessions. Data was analyzed with R (R Core Team, 2020), using Welch’s independent and one-sample *t*-tests and simple linear regressions. To complement the frequentist approach, Bayes factors (BF) were additionally determined. A Bayes factor indicates the likelihood of the alternative hypothesis compared to the null hypothesis given the observed data, that is, a BF of 5 would indicate that the alternative hypothesis is five times more likely than the null hypothesis given the data. To interpret the results, the following rules were applied: a BF of 1–3 was considered as anecdotal or weak evidence, a BF of 3–30 as positive to strong evidence, a BF of 30–150 as strong to very strong evidence, and a BF of > 150 as decisive evidence (Jarosz & Wiley, 2014).

Results

Results of all mean level comparisons are depicted in Table 3.3. The samples did not differ with respect to gender, age, work experience, and working conditions before the pandemic, such as weekly teaching hours, basic need satisfaction (autonomy, competence, relatedness), and self-efficacy. Faculty before and during the pandemic experienced comparable levels of stress at work and spent a similar amount of time on teaching, while

faculty in Sample 2 spent slightly less time on research (small effect size, anecdotal evidence as judged by the BF). When contrasting their experiences during the time of pandemic against their own experiences before the pandemic, faculty of Sample 2 reported to experience comparable levels of autonomy and competence, but clearly reduced levels of relatedness (large effect size, decisive evidence as judged by the BF), and slightly reduced levels of teaching satisfaction (small to medium effect size, anecdotal evidence as judged by the BF) and positive valence with regard to teaching (small to medium effect size, weak to positive evidence as judged by the BF).

Table 3.2

Data Structure of the Answered Basic and Session-Specific Questionnaires

	Sample 1		Sample 2	
	n	%	*n*	%
Participants completed …				
Basic questionnaire only (excluded from analyses)	6	–	53	–
Basic questionnaire and session-specific questionnaire	89	88.12	60	84.51
Session-specific questionnaire only	12	11.88	11	15.49
Final sample size	101		71	
Frequency of answered session-specific questionnaires				
1	3	2.97	19	26.76
2	6	5.94	8	11.27
3	7	6.93	6	8.45
4	17	16.83	11	15.49
5	68	67.33	14	19.72
6	–	–	12	16.09
10	–	–	1	1.41

Note. Data of Sample 1 (face-to-face teaching) was obtained before and data of Sample 2 (online teaching) during the COVID-19 pandemic. Participants were instructed to answer the session-specific questionnaire 3 to 6 times.

Table 3.3

Mean Level Comparisons of All Study Variables

	Sample 1		Sample 2					
	M	*SD*	*M*	*SD*	*t*	*p*	*d*	BF
Sample characteristics								
Age	40.01	10.42	39.57	10.95	−0.25	.806	0.04	0.19
Work experience	9.33	7.84	9.01	8.28	−0.23	.815	0.04	0.18
Stress at work	3.18	0.94	3.24	0.95	0.36	.719	0.06	0.19
Weekly teaching hours	6.75	4.21	6.76	4.64	0.01	.991	0.00	0.18
Time spent on teaching	17.15	10.29	19.42	13.93	1.08	.284	0.19	0.33
Time spent on research	19.31	12.10	14.32	12.04	−2.47	**.015**	0.41	2.81
Working conditions before the pandemic[a]								
Autonomy	5.97	1.02	6.01	0.97	0.24	.809	0.04	0.18
Competence	6.20	0.97	6.27	0.87	0.49	.624	0.08	0.20
Relatedness	6.09	1.32	6.41	1.15	1.54	.126	0.25	0.50
Self-efficacy	5.94	0.89	5.99	0.80	0.38	.708	0.06	0.19
Teaching experiences during the pandemic judged against teaching experiences before the pandemic								
Autonomy	–	–	0.33	1.62	1.60	.116	0.21	0.47
Competence	–	–	0.05	1.38	0.28	.780	0.04	0.15
Relatedness	–	–	−1.97	1.46	−10.41	**<.001**	1.34	$1.2e^{12}$
Teaching satisfaction	–	–	−0.55	1.74	−2.45	**.017**	0.32	2.18
Positive valence	–	–	−0.77	2.10	−2.82	**.006**	0.36	5.13
Session-specific teaching experiences								
Enjoyment	6.51	0.99	5.70	1.45	−4.10	**<.001**	0.68	823.03
Pride	4.01	1.54	4.07	1.79	0.25	.803	0.04	0.17
Boredom	2.06	1.02	2.37	1.36	1.61	.109	0.26	0.63
Anxiety	1.55	0.82	1.83	1.37	1.50	.137	0.25	0.57
Anger	1.65	0.78	2.32	1.62	3.23	**.002**	0.56	59.85
Shame	1.37	0.63	1.75	1.29	2.26	**.026**	0.39	3.05
Teaching satisfaction	6.55	0.90	6.01	1.29	−3.06	**.003**	0.50	20.21
Autonomy	7.09	0.96	6.51	1.15	−3.48	**<.001**	0.56	56.10
Competence	6.81	0.93	6.32	1.02	−3.23	**.002**	0.51	22.40
Relatedness	5.44	1.30	4.55	1.67	−3.78	**<.001**	0.61	184.98

Note. Negative *t*-values indicate lower values of the respective variables in online teaching during the pandemic.

[a] Rated retrospectively by Sample 2.

The comparison of the experiences when teaching their chosen (synchronous) class showed that faculty teaching online during a time of pandemic reported to experience fewer autonomy (medium effect size, strong evidence as judged by the BF), slightly fewer competence and teaching satisfaction (medium effect sizes, positive evidence as judged by the BF), and clearly fewer relatedness (medium effect size, decisive evidence as judged by the BF) compared to those teaching before the pandemic. Furthermore, faculty teaching online reported to experience clearly fewer enjoyment (medium to large effect size, decisive evidence as judged by the BF), more anger (medium effect size, strong evidence as judged by the BF), slightly more shame (small to medium effect size, anecdotal evidence as judged by the BF), and comparable levels of pride, boredom, and anxiety (small effect sizes, evidence in favor of null hypothesis as judged by the BFs), compared to faculty teaching face-to-face before the pandemic. Regression analyses indicated that technical problems did not predict perceived competence ($\beta = -.15$, $R^2 = .02$, $p = .24$, BF = 0.47, $n = 60$) and that the number of students that were on average visible during an synchronous online class tended to influence faculty members’ perceived relatedness with students ($\beta = .26$, $R^2 = .07$, $p = .03$, BF = 2.04, $n = 71$).

Overall, correlations among the basic needs, discrete teaching emotions, and teaching satisfaction as measured directly after teaching a class were in line with SDT, that is need satisfaction positively correlated with positive emotions and teaching satisfaction and negatively with negative emotions (see Table 3.4).

The strength of a diary approach is that the aggregated scores across multiple sessions are a more reliable measure of in-situ experiences than a single measurement. Therefore, the relatively high percentage of participants in Sample 2 that answered the session-specific questionnaire only once may have skewed the results. To rule out this opportunity, all analyses were re-run excluding participants that had answered the sessions-specific

questionnaire once (see Table 3.5 for detailed results). Supporting the findings reported above, the result patterns remained stable, but expectedly the evidence as judged by the Bayes factors was stronger in the bigger sample (including single measurements), because with increasing sample size Bayes factors tend to rather consistently develop in favor of the alternative hypothesis, in case there is an effect.

Discussion

The present study compared the satisfaction of the three basic psychological needs for autonomy, competence, and relatedness as well as emotional experiences and teaching satisfaction of faculty teaching online during to faculty teaching face-to-face before the COVID-19 pandemic. Overall, both samples were well comparable, the only small difference regarded the hours spent on research during the time of data collection, which was less during the time of pandemic. As the hours spent on teaching were comparable, this may imply that during the pandemic either research efforts were impaired directly, for instance because testing in laboratories was impossible due to contact restrictions and hygiene regulations, or that less time could be spent on research, because the time necessary for administrative and organizational tasks increased, teaching could not be neglected, and therefore research activities had to be reduced.

To get a general impression of faculty members’ teaching experiences during the pandemic, we asked faculty teaching during the pandemic to judge their current overall online teaching experiences against their own typical face-to-face teaching experiences before the pandemic (within-person comparison). Faculty members’ overall impression was that their own autonomy and competence did not differ before and during the pandemic, but that in online teaching during the pandemic their relatedness with students was severely reduced and that teaching was less satisfying and pleasant than in face-to-face teaching before the pandemic. These findings align with comparisons between students attending the

Table 3.4

Correlations Among All Study Variables for the Complete Sample and Faculty That Answered the Session-Specific Questionaire at Least Twice

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1. AUT−B	–	.22	.34*	.23	−.11	.06	−.34	−.11	−.22	−.18	.21	.03	−.13	.11	−.04	−.13	−.21	.06	−.19	.33*	.25	.14	.07
2. COM−B	.22**	–	.31	.49***	−.09	.15	.02	.10	−.20	−.08	−.10	−.23	−.02	−.03	−.01	−.09	−.19	−.01	−.15	.06	.32*	.04	.21
3. REL−B	.32***	.32***	–	.11	.04	−.10	−.19	−.32	−.27	−.30	.01	.13	.14	.05	.07	.04	−.22	−.02	−.20	.14	.11	.07	.06
4. SE−B	.23**	.45***	.09	–	−.11	.09	.04	.11	−.10	−.05	.08	.10	−.03	.04	.07	−.05	−.23	−.10	−.19	.14	.38**	.19	.22
5. AUT−C	−.07	−.02	.08	−.11	–	.63***	.41	.48	.66***	−.14	−.35	.04	−.05	.16	.31	.12	−.01	.04	.17	.30	.19	.03	.30
6. COM−C	.10	.23	−.04	.00	.61***	–	.58**	.79***	.60**	−.06	−.23	−.19	−.32	.03	.39	.11	.11	.09	.13	.10	.16	−.01	.16
7. REL−C	−.29*	.04	−.15	.04	.45***	.49***	–	.61**	.56*	−.18	−.12	−.30	−.04	.28	.52*	.07	.38	.02	.35	.06	.06	.32	.26
8. TS−C	−.10	.15	−.23	.00	.54***	.79***	.55***	–	.51	−.06	−.26	−.16	−.20	.02	.35	.07	.14	−.02	.18	−.10	.10	.05	.28
9. PV−C	−.20	−.08	−.11	−.17	.68***	.60***	.55***	.57***	–	.02	−.02	−.10	−.10	.13	.33	−.02	.19	.05	.23	.13	.09	.17	.23
10. STR−G	−.21**	−.10	−.32***	−.03	−.13	−.09	−.12	−.03	.01	–	.29	−.06	−.18	−.12	.00	−.11	.12	.08	.18	−.20	−.13	−.03	−.19
11. TP−G	.14	−.16	−.10	.10	−.38**	−.31*	−.17	−.31*	−.14	.31*	–	−.31	−.17	−.05	−.16	−.16	−.05	.03	−.03	−.01	−.03	−.09	−.19
12. AS−S	.01	−.02	.06	.06	−.10	−.17	−.27*	−.19	−.13	−.07	.07	–	.04	−.06	−.06	.15	−.23	.05	−.22	.01	.03	−.09	.07
13. VS−S	−.08	−.02	.08	.03	.02	−.23	.08	−.14	−.06	−.16	−.20	−.10	–	.24	−.08	−.18	−.08	−.07	−.05	.10	.00	.15	.20
14. ENJ−S	.10	−.04	.03	.05	.11	−.06	.33*	−.08	.07	−.12	−.09	−.37**	.28*	–	.38***	−.27	−.16	−.42***	−.23	.57***	.57***	.56***	.70***
15. PRI−S	.00	−.03	.04	.07	.20	.25	.51***	.20	.27*	.00	−.12	−.22	.09	.41***	–	.05	.18	.09	.13	.22	.25	.41***	.35**
16. BOR−S	−.09	−.14	.03	−.03	.12	−.03	−.05	.02	−.02	−.11	−.12	−.07	−.14	−.26***	.00	–	.25	.26	.19	−.16	−.18	−.10	−.19
17. ANX−S	−.15	−.20*	−.22**	−.12	−.08	−.12	.16	−.09	−.10	.16	.20	−.09	−.08	−.15	.10	.24**	–	.43***	.77***	−.31*	−.41***	.00	−.28
18. ANG−S	.02	−.07	−.03	−.09	−.03	.00	−.13	−.10	.02	.05	.04	.00	−.11	−.38***	.05	.39***	.31***	–	.51***	−.36**	−.34**	−.25	−.47***
19. SHA−S	−.15	−.14	−.16	−.11	.03	−.13	.14	−.07	−.02	.17*	.15	−.07	−.07	−.20**	.04	.23**	.79***	.37***	–	−.32*	−.49***	−.10	−.33**
20. AUT−S	.32***	.05	.15	.15	.25	−.02	.09	−.15	.13	−.21*	−.04	−.02	.17	.54***	.19*	−.11	−.29***	−.32***	−.26***	–	.67***	.42***	.60***
21. COM−S	.20*	.30***	.10	.32***	.25	.20	.21	.20	.25*	−.13	−.15	−.09	.16	.53***	.29***	−.16*	−.46***	−.28***	−.51***	.66***	–	.46***	.75***
22. REL−S	.11	.05	.05	.16*	.13	.12	.43***	.16	.19	−.01	−.17	−.28*	.26*	.58***	.45***	−.21**	−.08	−.34***	−.19*	.34***	.47***	–	.50***
23. TS−S	.04	.19*	.01	.18*	.16	.12	.31*	.18	.18	−.17*	−.14	.00	.26*	.68***	.40***	−.26***	−.34***	−.42***	−.39***	.53***	.74***	.53***	–

Note. Correlations of the whole sample are depicted below the diagonal. Correlations of those faculty who answered the session-specific questionnaire at least twice are depicted above the diagonal. AUT = autonomy, COM = competence, REL = relatedness, SE = self-efficacy, TS = teaching satisfaction, PV= positive valence, STR = stress, TP = technical problems, AS = number of attending students, VS = Visible students (on average), ENJ = enjoyment, PRI = pride, BOR = boredom, ANX = anxiety, ANG = anger, SHA = shame. 144 are reported as experienced before (B) the pandemic. 5-9 are experiences during the pandemic compared (C) against own experiences before the pandemic (Sampe 2 only). 10-11 General (G) experiences during time of data collection. 12-23 Session-specific (S) information and experiences.

Table 3.5

Mean Level Comparisons of All Study Variables (Sample Reduced to Faculty Who Answered the Session-Specific Questionnaire at Least Twice)

	Sample 1		Sample 2					
	M	*SD*	*M*	*SD*	*t*	*p*	*d*	BF
Sample characteristics								
Age	40.05	10.48	38.81	10.67	−0.64	.522	0.12	0.23
Work experience	9.42	7.83	8.93	7.89	−0.35	.729	0.06	0.20
Stress at work	3.19	0.94	3.18	0.98	−0.05	.957	0.01	0.19
Weekly teaching hours	6.66	4.14	6.76	4.29	0.14	.892	0.02	0.19
Time spent on teaching	16.89	10.05	18.40	13.61	0.67	.503	0.13	0.25
Time spent on research	19.51	12.03	14.70	11.84	−2.23	**.028**	0.40	1.74
Working conditions before the pandemic [a]								
Autonomy	5.96	1.03	6.02	0.95	0.35	.724	0.06	0.20
Competence	6.19	0.97	6.33	0.91	0.86	.389	0.15	0.27
Relatedness	6.09	1.32	6.53	1.11	2.07	**.041**	0.35	1.09
Self-efficacy	5.94	0.89	5.95	0.86	0.04	.965	0.01	0.19
Teaching experiences during the pandemic judged against teaching experiences before the pandemic								
Autonomy	–	–	0.36	1.61	1.54	.131	0.22	0.47
Competence	–	–	0.16	1.28	0.86	.396	0.12	0.22
Relatedness	–	–	−1.89	1.51	−8.59	**<.001**	1.25	$2.4e^{8}$
Teaching satisfaction	–	–	−0.60	1.57	−2.60	**.012**	0.38	3.19
Positive valence	–	–	−0.68	2.07	−2.26	**.029**	0.33	5.13
Session-specific teaching experiences								
Enjoyment	6.48	0.98	5.88	1.29	−2.97	**.004**	0.55	19.22
Pride	4.02	1.53	4.20	1.60	0.66	.512	0.11	0.23
Boredom	2.09	1.02	2.24	1.12	0.81	.419	0.14	0.25
Anxiety	1.57	0.83	1.74	1.15	0.96	.339	0.18	0.31
Anger	1.65	0.77	2.27	1.37	2.99	**.004**	0.60	44.27
Shame	1.39	0.64	1.69	1.10	1.87	.066	0.37	1.56
Teaching satisfaction	6.54	0.90	6.11	1.12	−2.38	**.020**	0.44	3.42
Autonomy	7.05	0.94	6.52	1.17	−2.81	**.006**	0.51	10.30
Competence	6.78	0.92	6.38	0.89	−2.55	**.012**	0.43	3.23
Relatedness	5.47	1.26	4.79	1.46	−2.85	**.005**	0.51	9.74

Note. Negative *t*-values indicate lower values of the respective variables in online teaching during the pandemic.

[a] Rated retrospectively by Sample 2.

same hybrid classes online or face-to-face (Butz et al., 2014; Butz & Stupnisky, 2016). In this general comparison faculty may not have perceived impairments of autonomy and competence possibly because they focused on their autonomy in choosing content, their preferred implementation of online teaching (e.g., synchronous sessions or asynchronous self-learning units), or their chosen teaching approach (e.g., teacher or student-centered), as well as on stable aspects of teaching competence that apply to online and face-to-face teaching, such as content knowledge. Technical aspects of online teaching and the need to try and learn new pedagogical methods were possibly neglected. The difference that was most salient was the reduced relatedness with students, which probably impaired the general teaching situation in online teaching. These retrospective comparisons need to be interpreted with caution, though, because they compared general teaching experiences and may have been affected by memory biases.

Differences in the Session-Specific Experiences Regarding the Satisfaction of the Basic Needs, Teaching Emotions, and Teaching Satisfaction

In our main analyses we compared the session-specific experiences of faculty that taught before the pandemic to faculty that taught during the pandemic in two independent samples (between-person comparison) with respect to basic need satisfaction, teaching emotions, and teaching satisfaction.

Basic Need Satisfaction. Faculty teaching online during the pandemic reported lower levels of perceived autonomy than faculty teaching face-to-face before the pandemic. Although unexpected, it is reasonable that faculty members’ autonomy was impaired when involuntarily teaching online during the pandemic. Meta-analytic findings on antecedents of autonomy in the workplace suggested that workload and job demands were negatively, and perceived organizational support positively related to autonomy (Van den Broeck et al., 2016). Especially in the beginning of the pandemic, the workload to design classes and the

job demands in general were very high, and institutions were partially unable to react to the new circumstances quickly enough to provide high quality support. In addition to such general circumstances that may have impaired autonomy, faculty may have felt limited in designing individual sessions (Regan et al., 2012), for instance because not all well-known and well-working teaching methods from face-to-face teaching could be transferred to the online environment, the available online meeting tool did not offer desired features (e.g., breakout rooms, polling, interactive whiteboard, etc.), the technical equipment to implement the desired form of online teaching was not available, or faculty could not adapt their teaching to the current needs of students by reading visual cues because of the absence of such cues due to deactivated videos. Such impairments in realizing teaching as intended may have reduced perceived autonomy within classes.

Faculty teaching online during the pandemic reported lower levels of perceived competence than faculty teaching face-to-face before the pandemic, which is in line with expectations. On top of generally increased workload and job demands and a possible lack of organizational support (Van den Broeck et al., 2016), a lack of experience and training for online teaching may have reduced perceived competence (Downing & Dyment, 2013). Some faculty may not have been aware of best practices to design online learning, such as avoiding controlling language, providing personalized feedback to help students master challenges, and fostering groupwork (Wang et al., 2019). But even when faculty were aware of best practices, when teaching was shifted to an online format this rapidly, they may not have had enough capacities or the technical and pedagogical skills to implement them. Within an online teaching session, faculty may have felt unprepared to master the teaching task, unable to teach as effectively and in as high quality as in face-to-face settings with well-tried methods, or incapable of handling the technology very well, thereby experiencing reduced levels of perceived competence.

Faculty teaching online during the pandemic reported clearly lower levels of perceived relatedness than faculty teaching face-to-face before the pandemic. These perceptions mirror the physical distance between faculty and their students and is in line with our expectations and with the experiences of faculty and university students in online teaching and learning contexts (Butz et al., 2014; Butz & Stupnisky, 2016; Filak & Nicolini, 2018; Regan et al., 2012). Before the COVID-19 pandemic, the importance of relatedness in the educational context may have been underestimated, because it develops very easily in face-to-face settings by being in the same room, thereby interacting with and getting to know each other. Only when this all too natural mechanism did not kick in due to contact restrictions, faculty and students may have recognized the important role relatedness plays in teaching and learning. Although there are possibilities to form relationships with students in online contexts, for instance by self-disclosure, that is to reveal personal information, responding in a timely manner, and using humor (Song et al., 2016; Sung & Mayer, 2012), such offers probably cannot adequately replace classroom interactions that naturally transform many individuals into a group over the course of a semester. Due to the highly salient physical distance, the satisfaction of the need for relatedness was most severely impaired in online teaching.

Overall, it can be concluded that faculty teaching online in a time of pandemic experienced a reduced satisfaction of their basic needs for autonomy, competence, and relatedness compared to their colleagues teaching face-to-face before the pandemic, whereby relatedness was most severely thwarted. Our correlative findings were overall in line with SDT, that is the satisfaction of the needs for autonomy, competence, and relatedness were positively related to positive emotions and teaching satisfaction and negatively related to negative emotions. Therefore, the reduced levels of need satisfaction bear the potential to be

accompanied by less favorable emotional experiences in online compared to face-to-face teaching.

Teaching Emotions and Teaching Satisfaction. Results regarding teaching emotions were diverse and not consistently in line with expectations, which assumed lower levels of positive and higher levels of negative emotions due to the reduced levels of need satisfaction in online compared to face-to-face teaching. While faculty reported clearly less enjoyment, more anger, and a tendency towards more shame, they reported similar levels of pride, boredom, and anxiety in online teaching during compared to face-to-face teaching before the pandemic.

Of the positive emotions enjoyment but not pride was reduced in online compared to face-to-face teaching. Enjoyment has been shown to be influenced strongly by autonomy and relatedness (Klassen et al., 2012) as experienced in a teaching situation and to be reduced in online teaching compared to face-to-face teaching (Kanning & Ohlms, 2021). Our correlational results support strong relationships between all three needs and enjoyment. Taken together, results showed that teaching was clearly less enjoyable in online compared to face-to-face teaching probably due to reduced need satisfaction, which is in line with expectations and previous research. We extended previous research on SDT and positive emotions in general by assessing a further discrete positive emotion in our study, namely pride, which has not been considered from an SDT perspective before. Pride was experienced to a similar extent in both teaching settings. In contrast to enjoyment that is claimed to arise when performing an activity, pride is claimed to be elicited when successfully mastering an activity (Pekrun, 2006). Possibly, taking pride in mastering teaching could take different forms: faculty may have been proud when they managed to deliver high quality teaching in face-to-face settings, but they may have been just as proud when they accomplished to teach online despite the challenging circumstances. These different subjective definitions of

success may explain why despite reduced levels of need satisfaction and positive correlations between need satisfaction and pride, faculty members’ reported pride in online compared to face-to-face teaching was not reduced.

Of the negative emotions anger and with a tendency shame were enhanced in online compared to face-to-face teaching, but not boredom and anxiety. Anger has been shown to be influenced strongly by all three needs (Klassen et al., 2012) as experienced within a teaching situation, which is corroborated by our correlative findings. The claim that anger arises when an individual needs to perform activities that hold negative valence (Pekrun, 2006) seems to hold in online teaching during the pandemic, because teaching was a task that needed to be accomplished but that was seemingly perceived as less pleasant in online than in face-to-face teaching based on faculty members’ retrospective judgements. Taken together, the clearly increased anger in online compared to face-to-face teaching is in line with expectations and previous research. Shame is claimed to be experienced when the outcome of an activity is considered as failure and when oneself is held responsible (Pekrun, 2006). Within an SDT framework this idea mainly translates into insufficient satisfaction of the need for competence. Indeed, our correlative findings showed a strong negative correlation between competence and shame. Taken together, a tendency to experience more shame probably mainly due to the reduced levels of perceived competence is in line with expectations and previous research. Boredom has been shown to be negatively related to need satisfaction in the workplace and in (MOOC) students (Buhr et al., 2019; Sulea et al., 2015; van Hooff & van Hooft, 2017). Therefore, it seemed plausible that the reduced need satisfaction in online teaching increased the experience of boredom. Nevertheless, there were no differences observed with respect to boredom between online and face-to-face teaching. Our correlative findings only lend support for weak negative relationships between boredom and competence and relatedness, which may indicate that the reasons for boredom when teaching online may

not be sufficiently covered by basic need satisfaction. It may be the case that because teaching is a highly diverse task that holds a lot of variety and typically requires a highly active role of faculty may prevent increased levels of boredom in any form of teaching in contrast to work in general or learning, which may encompass more phases of inactivity and possibilities to experience boredom. Anxiety has been shown to be mainly influenced by competence (Klassen et al., 2012), which aligns with our correlative findings. Therefore, the reduced competence in online teaching could have been accompanied by higher levels of anxiety, but it was not. Despite their reduced perceived competence, faculty seem to have felt competent enough – as indicated by the still high levels of reported competence – to master their teaching and therefore to not feel more anxious when teaching online than face-to-face, possibly due to their prior teaching experience in face-to-face teaching. The reduced levels of competence may have been mainly due to technical insecurities, which possibly may not have affected anxiety during the onset of a pandemic as severely because they could be attributed to the adverse circumstances everyone was facing and trying to overcome together rather than to skills faculty should have. Taken together, based on compelling information we expected enhanced anxiety in online teaching, but based on our data we could only speculate why anxiety did not differ between online and face-to-face teaching.

Teaching satisfaction was reduced in online compared to face-to-face teaching, which is in line with previous findings on the same topic based on retrospective judgements (Kanning & Ohlms, 2021). Because teaching satisfaction has been shown to be influenced by the satisfaction of the basic needs (Crick et al., 2020; Larson et al., 2019; Seipel & Larson, 2018), which was corroborated by our correlative findings, and need satisfaction was reduced, the reduced teaching satisfaction is in line with expectations and previous research.

Overall, faculty teaching online during the pandemic reported less favorable emotional experiences, because they experienced one positive emotion (enjoyment) clearly

less strongly and some of the negative emotions (anger, shame) more strongly than faculty teaching face-to-face. Taken together, online teaching in a time of pandemic was a less pleasurable activity due to reduced need satisfaction, especially of the need for relatedness, less favorable emotional experiences, and reduced teaching satisfaction.

Factors in an Online Teaching Environment Influencing Need Satisfaction

Competence is experienced when individuals successfully interact with their environment (Deci & Ryan, 2004). Based on this definition and previous findings (Downing & Dyment, 2013) it seemed reasonable that experiencing more technical problems in online teaching may decrease the perceived competence of faculty because it would be a sign of hindered interaction with their environment. Based on our data we cannot support or reject this claim, though. Although the assumption is well-grounded, unfortunately the measurement with respect to this research question was insufficient. While the technical problems were assessed in the basic questionnaire on a general level, the satisfaction of the need for competence when teaching was assessed after individual sessions. Measuring at two different levels most likely prevented establishing a meaningful relationship. Having said this, the negative correlations between technical problems and autonomy and competence as judged during the pandemic against before the pandemic point in the direction, that there may be effects in the expected direction if variables were measured on the same level. Nevertheless, at this point it remains unclear whether technical problems that occur when teaching a synchronous online class may have effects on the perceived competence experienced during that class.

Relatedness is experienced when individuals feel connected with others (Deci & Ryan, 2004). We suspected that faculty who saw a larger number of their students would experience a stronger satisfaction of the need for relatedness. Our findings lend some support for this claim, but the results need to be interpreted with caution because the evidence is not

strong. Nevertheless, the finding is remarkable when taking into consideration that in typical online classes the maximum number of visible students is highly limited, because many students prefer to not activate their videos due to (data) privacy issues or internet bandwidth limitations (Yarmand et al., 2021) and because faculty share their screen most of the time which allows to simultaneously see only a limited number of students (typically around 5 videos). Therefore, finding weak evidence despite a highly limited range in the number of on average visible students (Median = 2.4 students, 90% of all values below 10 students) and a relatively small sample size (n = 71) is promising. The currently explained variance, however, leaves room for improvement. It is likely that other factors, such as students responding to questions quickly by speaking up or using the chat, providing feedback by using emoticons, and generally active class participation may additionally contribute to the satisfaction of the need for relatedness.

Limitations and Directions for Future Research

This study shed first light on the experiences of faculty teaching online, in particular teaching online during a pandemic. The onset of the COVID-19 pandemic challenged faculty not only at the workplace by having to develop new approaches to teaching, re-organizing ongoing studies, and establishing new administrative procedures, but also in private life because it was a time of personal and mental challenges. It cannot be excluded that these omnipresent stressors outside the job context to some extent carried over to the experiences when teaching. Therefore, the presented findings may not fully generalize to online teaching in general and need replication in online teaching that is not affected by a pandemic to establish general claims.

Such replications need to be conducted in different countries, because the degrees of freedom in teaching vary considerably between higher education systems and therefore the effects may be more pronounced in countries with a lesser extent of self-determination in

teaching. To obtain stronger evidence for the smaller effects it is furthermore necessary to increase the sample size in such replication studies. Replicating a study on the teaching experiences in online teaching during the onset of a pandemic seems almost impossible, though, as soon as a longer period of time has passed and the experiences can only be reported retrospectively. Nevertheless, despite the relatively small sample size, the presented findings are convincing, not least because even when excluding participants who had answered the session-specific questionnaire only once and thereby reducing the sample size even more, the pattern of findings remained stable.

Within the scope of this study, it was not possible to focus on many different specific factors that may hinder or foster the satisfaction of the basic needs and discrete emotions in online teaching. It is necessary, however, to identify the origins of need satisfaction and emotions in online teaching to develop recommendations on how to make online teaching more attractive and pleasant for faculty. To this end, future research could assess further factors in synchronous online environments that may influence need satisfaction, such as perceived limitations in useable teaching methods, technical problems within the session (rather than technical problems in general), and perceived active participation of students. It could further be explored whether a lack of visual feedback from students due to missing videos could be compensated by introducing new features into online meeting software that could use face or behavior detection software to provide faculty with live class-aggregated information about student attention, confusion, and engagement to support adaptive teaching without the need for students to share sensitive video data (Whitehill et al., 2014; Yarmand et al., 2021). A further direction to deepen our understanding of online teaching more generally is to identify factors in asynchronous teaching that contribute to need satisfaction and positive emotional experiences within faculty and students to develop online classes that may blend

the best of both synchronous and asynchronous approaches, namely the opportunity for direct interaction combined with self-directed and therefore flexible learning phases.

A strength of the diary-design that was not fully taken advantage of within the presented study is the possibility to investigate within-person variability over the course of multiple sessions. In combination with the assessment of further possible sources of need satisfaction, this approach would contribute to understanding how stable or variable basic need satisfaction and emotional experiences are over time and which aspects of an online teaching environment trigger them. Furthermore, it remains to be explored why despite the consistently lower need satisfaction not all discrete emotions were affected as expected (similar levels of pride, boredom, and anxiety). It is possible that certain emotions are affected by the satisfaction of different needs in specific ways and that the general assumption of higher need satisfaction being associated with more positive and fewer negative emotions is too simplistic. The clearly differing strengths of correlations between single need satisfaction dimensions and the discrete emotions lend some support for this idea. In this regard it may be a fruitful endeavor to further explore the relations between SDT and control-value theory (Pekrun, 2006) because it is well possible that the satisfaction of the basic needs may result in specific control and value appraisals, which in certain combinations predict specific discrete emotions. A first step into this direction has been made in a student sample (Buhr et al., 2019), but it is necessary to consider all three needs and to extend the findings to different discrete emotions also in faculty to determine whether and if so how basic needs can be considered as antecedents of control and value appraisals and thereby discrete emotions in online teaching settings.

Implications

Although online teaching in a time of pandemic was overall reported to be a less pleasant experience than face-to-face teaching, online and hybrid teaching provide

opportunities for higher education after the pandemic. To name only a few, faculty could structure their work more flexible when offering some of their classes in an online format, students could gain flexibility in their studies, and institutions’ room shortage could be mitigated to some degree. Hybrid offers may even need to be a necessary next step towards after-crisis teaching to fulfill students’ and faculty members’ desire for face-to-face teaching and to allow individuals with heightened risk of an COVID-19 infection and a severe course of disease to at the same time avoid direct contact with others. To encourage faculty and students to continue with online education after the pandemic, it seems necessary to foster positive online teaching and learning experiences and to advance teaching quality in these new formats. Faculty members’ perceived autonomy and competence likely increased with growing experience in online teaching, that is having discovered teaching methods that are effective in online teaching by try and error and having developed confidence in using methods and technology more flexibly. Nevertheless, institutional and mainly individualized support in designing new online or hybrid classes (Downing & Dyment, 2013) as well as regular exchange about teaching experiences among faculty seem imperative to further develop online teaching offers. To face the challenge of severely reduced relatedness not only between faculty and students but also among students, it is necessary to actively foster interaction and timely communication to create a feeling of belonging (Sung & Mayer, 2012). One possibility to do so may be to stop screen-sharing when discussing questions and to encourage students to share their videos in such situations to create a feeling of being in class together and to be connected. Furthermore, to support relationships among students it may help to trigger self-disclosure by for instance prompting students to talk about one private question in a small group before starting content-focused groupwork or discussions in breakout sessions (Akcaoglu & Lee, 2016; Shackelford & Maxwell, 2012; Sung & Mayer, 2012). It seems likely that with increased skills to satisfy the needs for autonomy,

competence, and relatedness in online teaching faculty and students can savor the experience to a greater extent. This may be especially true when online offers complement rather than forcedly replace face-to-face teaching. In post-pandemic face-to-face teaching, faculty may still want to make use of technology they got to know and value during online teaching, such as live polling, the use of etherpads or online mind maps, because such technologies offer opportunities for live interaction not only in seminar-size but also in larger lecture-size groups that cannot be realized through teaching that is not technology-supported.

Although online classes can substitute rather well for some aspects of face-to-face teaching and learning, it became very obvious that higher education institutions are not only a place of knowledge generation, transmission, and advancement, but also a place that enables people to connect with each other, build relationships, and interact as social beings. Therefore, even though online education has advantages, the crucial role of a successful social integration into the university community, which probably happens more easily in face-to-face settings, should not be underestimated.

Conclusion

To conclude, SDT was a suitable framework to investigate faculty teaching emotions in online teaching during a time of pandemic. Overall, online teaching in a time of pandemic was a less pleasurable experience than face-to-face teaching before the pandemic as indicated by lower levels of need satisfaction, positive emotions, and teaching satisfaction as well as higher levels of negative emotions. For the future it is important to take away the positive aspects of this forced step towards digitalization, such as more flexibility for faculty and students when adding comprehensive, well-designed hybrid or online solutions to the classical face-to-face offers and to integrate collaboration-supportive technology in classroom teaching, to make higher education more attractive and accessible for a more diverse student body.

5. General Discussion

This dissertation aimed to shed light on the emotional experiences of teachers and faculty in two so far under-researched contexts, namely grading written student work and (emergency) online teaching. To this end, we conducted three studies and on the one hand explored the occurrence of grading emotions in teachers and faculty, their effects on grades, and their antecedents from a control-value perspective, and on the other hand compared the emotional experiences, teaching satisfaction, and the role of the satisfaction of the needs for autonomy, competence, and relatedness between faculty offering emergency online teaching during and faculty offering face-to-face teaching before the COVID-19 pandemic.

With respect to grading, the findings support the idea that the grading task itself is an activity that elicits a variety of discrete emotions in teachers and faculty. The first study showed that a relatively superficial feature of written student work, namely handwriting quality, can induce enjoyment and anger, but not boredom, within participants experiencing two different handwriting conditions. Enjoyment and anger in turn influenced grades, that is an individual who experienced higher levels of enjoyment and lower levels of anger between the good and bad handwriting condition, assigned better grades to an essay. These within-participant findings were in line with previous research showing that externally induced mood biased grades in emotion-congruent ways in an between-participant design (Brackett et al., 2013). Our findings extended previous research by showing that not only external sources but also specific features of an essay itself have the potential to elicit different emotions, that grading emotions can vary within individuals when grading multiple essays, and that there are specific antecedents and effects of different positive and negative emotions when grading (e.g., handwriting quality induced anger but not boredom, which both are negative activity emotions). Unfortunately, the effects that boredom unfolds on grades could not be explored because the handwriting manipulation did not induce boredom. As grading is a highly

repetitive task, it nevertheless holds the potential to elicit boredom in other ways and the effects of boredom warrant further investigation. Study 2a showed that research, teaching, and grading are activities in which faculty experience different levels of specific discrete emotions, lending further support for the context-specificity of emotions (Frenzel et al., 2015, 2016) and the assumption that although grading and teaching may in some instances be rather intertwined, grading may be viewed as a separate task from teaching due to the distinct emotional pattern. On average, teaching was the context with the most favorable emotional experiences, qualified by high levels of enjoyment and pride and generally lower levels of negative emotions, with boredom being experienced least strongly. In comparison to teaching, research was emotionally slightly less favorable (small to medium effect sizes) because faculty reported to experience consistently lower levels of positive and higher levels of negative emotions. In comparison to both teaching and research, grading was emotionally clearly less favorable (mostly medium to large effect sizes) because especially the levels of enjoyment and pride were clearly lower and the levels of boredom were clearly higher than in both other contexts. The relatively high boredom levels lend support for the assumption that grading triggers boredom but due to other reasons than handwriting quality. Taken together, the findings supported qualitative claims that grading is an unpleasant task (e.g., Postareff & Lindblom-Ylänne, 2015)within faculty members' profession (and very likely also teachers' profession) that itself triggers a variety of discrete emotions.

We were, however, not only interested in showing that grading is an unpleasant task but even more interested in identifying factors that make grading unpleasant. Study 2b explored possible appraisal antecedents of grading emotions from a control-value theory perspective in two countries (U.S. and Germany) and yielded the main finding that the perception of cost seemed to be the most important predictor of positive and negative emotions in grading. That is, fewer cost was consistently associated with higher levels of

positive and lower levels of negative emotions across both samples. Cost measured the negative value of grading, which represented the thanklessness of the grading task and the extent to which it kept faculty away from more meaningful tasks. So overall, the more aversive grading was perceived, the more unfavorable emotions were experienced. The further results were rather inconsistent across the two samples, though. In addition to cost, diagnostic competence (i.e., the ability to assign fair grades) and social value (i.e., the importance of the relationship with students) were among the frequent predictors of grading emotions in the German and U.S. sample, respectively. The lack of predictive power of the other appraisals may be attributed to ceiling effects especially in the U.S. sample, and therefore although the present findings did not support all assumptions based on control-value theory, we cannot rule out that the used dimension may in fact contribute to experiencing grading emotions (especially those factors that were significant in the German sample with more variance in predictor variables but not in the U.S. sample, that is content knowledge, diagnostic competence, and utility value). The differences in predictive patterns may be explained by the different circumstances under which grading occurs in higher education in the U.S. and Germany. Thinking back to the description of possible grading scenarios in the introduction and after having exchanged information among the authors of Study 2 about how grading of written student work is implemented at higher education institutions in the U.S. and in Germany, it seems that grading in the U.S. is often more integrated in the teaching process (typically multiple assessments over the term) than it is in Germany (typically one assessment at the end of the term). Moreover, grading situations seem to differ to a greater extent within German (situations range from grading multiple assessments in one's own classes to grading centralized state exams without being involved in the preparatory classes) than in U.S. higher education (more precise institutional guidelines and no centralized exams). These first anecdotal impressions were largely mirrored in our data:

U.S. faculty consistently reported to have clearly more control and positive value with respect to grading, that is they indicated having more control over the grading process and the paper's content (i.e., could determine in their classes what to test, how, and when), being more knowledgeable about the content and being better able to assign fair grades, and perceiving grading as being of more positive value in terms of the importance of assigning fair grades (very high in both samples, though), the utility of grading (i.e., the extent to which grading could be used to inform their teaching), and the importance of the relationship with students. These differences may be mainly attributed to the around 25 percent of German faculty grading state examinations who get to know the examination topics only once they have to grade them, are highly restricted in their time for grading often not allowing for thorough procedures, and do not know the students. The larger variability of grading situations within Germany was mirrored in larger standard deviations in the German sample with respect to control over the grading process and the paper's content and the utility value of grading (depends on whether faculty teach the class). Considering the seemingly more favorable circumstances of grading in the U.S. compared to Germany, it is surprising that U.S. faculty did not consistently experience more positive and less negative emotions than German faculty. In fact, they did not differ with respect to enjoyment, boredom, and frustration. U.S. faculty expectedly experienced more pride and less anger, but unexpectedly also more anxiety. Because anxiety is an outcome emotion that is competence-dependent and experienced more intensely when value is high, the higher social value reported by U.S. faculty seems to have intensified the experiences of anxiety compared to German faculty who may not be confronted with the consequences of their assigned grades as often when they only assign one grade at the end of the academic term. The combination that cost was the consistently strongest predictor of all emotions and that its extent did not differ between the two samples may explain the comparably similar emotional experiences of faculty in both

samples despite the differences in control and positive value appraisals. The rather similar emotional experiences also lend support for the assumption that when faculty judge their emotional experiences with respect to grading, they use their own emotional experiences in other activities and maybe reports of colleagues' experiences as points of reference to evaluate their emotions for grading, but they do not compare their situation with the situation of other faculty outside their own frame of reference (e.g., in other countries with more or less favorable circumstances).

Taken together, the results on grading showed that grading is one of the tasks faculty and very likely also teachers do not favor. Assigning fair grades to student work was very important to faculty as indicated by very high diagnostic value ratings and cost seemed to be the main predictor of experiencing grading emotions. Grading was accompanied by different positive and negative emotions, but the overall emotional experience when grading was clearly less favorable than when conducting research or teaching. These comparably negative emotional experiences may be triggered by the circumstances of grading (e.g., other tasks that need attention, number of essays that need to be graded) or the grading task itself (e.g., handwriting quality, content quality, writing style, topic) — but whatever the sources of grading emotions are, grading emotions have the potential to bias grades to the disadvantage of students and therefore warrant further investigations to mitigate their effects on grades.

With respect to teaching, findings showed that "classical" face-to-face teaching was a task that faculty enjoyed and which was overall not only experienced more favorable than research and grading, but also more favorable than emergency online teaching. Although it is normal that job demands change constantly, they typically do so rather slowly and it is possible for employees to adjust without greater problems. When circumstances change drastically within a very short time, however, as it was the case during the onset of the COVID-19 pandemic, a generally pleasant task as teaching may become stressful and

aversive. Study 3 showed exactly that: faculty who were forced to teach online instantly during the COVID-19 pandemic without time for preparation (and did so in synchronous online meetings), experienced less autonomy and competence with respect to teaching, clearly less relatedness with their students, clearly less joy, more anger and slightly more shame when teaching, and fewer teaching satisfaction than faculty who taught face-to-face before the pandemic. The levels of pride, boredom, and anxiety were comparable. Based on previous research we had expected consistently higher levels of positive and lower levels of negative emotions due to the reduced need satisfaction, but pride, boredom, and anxiety may be emotions that need an additional appraisal process that was not mirrored in basic need satisfaction to be elicited and explained. Pride is outcome-dependent and faculty may have defined "successful teaching" differently in online and face-to-face teaching and therefore felt similarly proud in both formats. Boredom has been shown to be experienced relatively rarely in faculty teaching in general (Study 2a; see also e.g., Stupnisky et al., 2016, 2019a, 2019b) and maybe the teaching task is so diverse and the faculty members' role mostly so active that it is no context to typically trigger boredom irrespective of the format (online vs. face-to-face) of synchronous teaching. Because anxiety is competence-dependent and perceived competence levels were still high in online teaching, anxiety did not seem to be severely affected by the format of synchronous teaching. It remained unclear whether technical problems within a session may have influenced the perceived competence of faculty, but we found first hints that specific features of an online environment may contribute to the satisfaction of specific needs (e.g., average number of students with activated videos tended to influence faculty members' perceived relatedness with students) and thereby to more favorable emotional experiences. It is an interesting future endeavor to examine the predictive role of the distinct need satisfaction dimensions with respect to discrete teaching emotions especially in online settings. To which extent the results can be

generalized to online teaching that is not affected by a pandemic is still an open question, but probably many aspects do transfer to online teaching more generally.

Overall, the three presented studies contribute to the body of research on teacher and faculty emotions in still under-researched tasks, namely grading and online teaching. Although the findings provided some first insights, more research is needed to understand the mechanisms at play and to enable individuals to influence them in such a way that aversive tasks can be made less unpleasant or maybe even more pleasant to not impair teacher and faculty well-being and satisfaction in the long run.

Perspectives on Grading

Deliberations About Theories

Unarguably, grading in general and especially of written work is no pleasant task per se, but still teachers and faculty differ in how aversive they perceive grading to be and in the emotions that they experience when grading. And likely the emotions experienced during grading even vary within one person over the course of grading multiple essays. Although the presented studies provided some first insights into this important task, the findings not only answered but also raised questions. We argued that although grading is no typical achievement situation, control-value reasoning may likely be applicable because grading is a task that is embedded in an achievement context, namely assessing written student work. For students, the outcome of an achievement situation is typically rather easily determined and in most cases a grade (or pass/fail result). But is the grade itself also the outcome of the grading activity for teachers and faculty? Or is it rather assigning an appropriate and fair grade? And if so, who would decide whether teachers or faculty were successful in their endeavor or not? And if their performance is not assessed, is a theory based on the assumption that emotions are elicited in achievement situations really applicable to grading? Looking at grading from a control-value perspective under the assumption that deciding on a grade can be performed

more or less well, it was comparably obvious that multiple factors may contribute to perceived control in grading situations. Although the list may not be exhaustive, general control over how grading can be organized and how the exam is designed probably contribute to the extent one feels in control of grading and assigning a (fair) grade. More in line with the conception of control according to control-value theory, competence with respect to the content of the exam (content knowledge) and the assignment of fair grades (diagnostic competence) were also deemed important to perform well in grading. Control-value reasoning further suggests that an activity itself or its outcome may be appraised to have a certain extent of positive and negative value. But just as it seems hard to precisely identify the outcome of grading from the perspective of teachers and faculty, it seems similarly challenging to identify value in grading that is relevant to teachers and faculty directly because the "outcome" of grading typically does not have direct consequences for teachers and faculty themselves, but mainly for their students to whom the given grade is typically very important. Thinking the situation through drawing on value dimensions proposed by control-value and expectancy-value theory (Gaspard et al., 2015; Pekrun, 2006), graders may find intrinsic value in grading if essays deal with interesting content, which the grader may enjoy reading. Attainment value may be appraised if it is important to graders to do well in grading, which is closely linked to the typical idea of value in control-value theory, and may refer to the importance of assigning fair grades (which is highly important to faculty; see Study 2b). Utility value may include the usefulness of grading for teachers and faculty themselves (questionable), for their career (questionable), for their life more generally (highly questionable), or for their job (possible if grading feeds back to inform their teaching). Negative value may be appraised if a situation requires a lot of effort to be performed or hinders one from doing other things, which is reflected in the perceived costs of grading. And although the factors derived from theory that were included in Study 2b predicted grading

emotions in control-value theory-congruent ways if there were relations, the list does not seem to be exhaustive of all factors contributing to the emergence of grading emotions (comparably low level of explained variance in Study 2b). That other factors, for instance a superficial feature of student work such as handwriting quality may affect grading emotions was already shown in Study 1. Due to the high overlap between teaching quality, student performance, and grading, it seems reasonable that graders may not only experience emotions for grading based on their own control and value appraisals and success or failure in grading, but also based on their students' performance and development. For instance, graders may not only feel proud when grading because they are doing well on the task, but because their students did well in the exam. They may not only enjoy grading because they like the task, but because students improved a lot from one assessment to the next. They may be angry not only because they are stopped from doing more important things, but because they know that a talented student just did not put in enough effort to receive a good grade. They may not only be frustrated because grading is a thankless task, but also because they realize that a student may still not have understood the main point of what they had tried to teach. Taking such interconnections into account, there are surely triggers for grading emotions that were not captured with the chosen control-value approach to grading emotions. Therefore, it would be a fruitful endeavor to conduct in-depth interviews with teachers and faculty alike to identify further sources of grading emotions and to validate their prevalence in grading quantitatively as a subsequent step. Once more is known about the sources of grading emotions, it may be possible to develop empirically-based recommendations or even interventions to make grading less aversive.

Deliberations About Practice

From a theoretical perspective and based on some (anecdotal) findings from our data, increasing positive and especially decreasing negative value may contribute to more

favorable emotional experiences when grading. Apart from choosing essay topics that a grader finds interesting whenever possible, there is probably not that much that can be done to increase the intrinsic value of grading. It is typically also not possible to reduce the required workload for other tasks. Therefore, to reduce the perceived cost of grading, it would be necessary to make it less aversive. And to this end it may be necessary to find a few ways to make grading less unpleasant and maybe more useful. In situations in which teachers and faculty have influence on grading, one idea may be that they try to integrate grading more into the teaching and learning process rather than seeing it as a tool to only assess current student performance. For instance, in higher education students often have some weeks (rather than hours in an examination) to produce essays that are in the end graded. Grades on such essays are most of the time accompanied by written feedback to justify the grade. In such cases, grading consists of two steps: pointing out strengths and weaknesses of the work and deriving a grade based on the judgements. In such situations it may be more conducive to student learning (and not significantly more work) to provide written formative feedback on an essay before the submission deadline and to give students the opportunity to improve their essays before final submission. With such a procedure the utility value may increase because the formative feedback may foster student learning (if this is important to faculty) and also inform the grader about which aspects of the topic may not have been covered sufficiently in class yet. Furthermore, it may make grading less aversive because the probability of having to assign a “fail”, which was mentioned as causing negative emotions (Babb & Corbett, 2016), may be reduced when students get a chance to rework their assignments. Upon final submission, it is sufficient to only assign the grade because the work intensive feedback has been provided before. If teachers or faculty want to foster learning even more it would probably be beneficial to provide information on what improved since the last version. If it is not possible for the grader to provide individual feedback on each essay

before the submission, it may also be a good option to organize peer feedback, that is students read and comment on strengths and weaknesses of their colleagues' essays. This approach provides students with the opportunity to revise their work based on feedback before the final submission and thereby hopefully improving the overall quality without the need to read every essay. If such procedures are to be used, it is necessary, however, to teach and practice them so that the feedback is useful for students.

In the school context, the more typical situation is that an examination is written within a few hours and that students need to answer questions or discuss topics without being allowed to draw on external resources to produce their piece of work. Such situations are unpleasant for most students and even elicit test anxiety in some, which impairs student performance (von der Embse et al., 2018). Especially in subjects in which there are no clearly right or wrong answers (e.g., writing discussions in language subjects) students may perceive a relatively low level of control that may contribute to experiencing negative emotions. And if negative student emotions decrease their performance, grading the student works will be even less pleasant for graders because the performance is worse and thereby also the grades that need to be assigned. To make writing and grading such open examinations less unpleasant, one approach may be to increase control appraisals in students and teachers. One possible approach may be to develop a grading rubric that states all aspects that need to be covered in a certain assignment together in class and discuss which characteristics of the written assignment define the quality for each criterion. Then students could write a practice exam on a comparable topic and subsequently grade the practice exam of one of their classmates by applying the grading rubric. This would allow students to see how others approach such an assignment, to critically reflect on how strengths and weaknesses with respect to certain criteria may look like, and to internalize the grading criteria. In the examination they would know what they will be graded on and this may give them some

more control over the situation and thereby reduce the likelihood of experiencing negative emotions. For teachers, using a grading rubric often makes grading easier, more time-efficient, and less biased and sharing the grading rubric with their students up front gives them the security of having been transparent about how grades are going to be determined. These aspects may contribute to perceiving more control over grading. Additionally, good exam preparation enhances the chances of reading more good essays which is typically more pleasant. Overall, grading better essays in a shorter amount of time with more confidence in the fair assignment of grades probably contributes to experiencing grading as less aversive.

Even more uncontrollable is the situation of student teachers writing their state exams in Bavaria and faculty grading them as outlined in the introduction. In this specific context, the process of writing and grading the exams seems sometimes more like a lottery than a valid assessment of student performance. At this point one may even ask whether assessing content knowledge about pedagogical topics is a good measure of assessing pedagogical skills needed for teaching in the first place, but the current form of assessment is probably not going to change any time soon. For student teachers the situation is highly uncontrollable because the topics that may be asked are too numerous to learn all of them in-depth especially with the requirement to cite papers to backup claims without being allowed to use external sources during the examination. Therefore, most students are forced to concentrate on the most likely topics and theories that may explain processes in different contexts, leaving a certain level of insecurities. But let us assume for a moment that it was possible to sufficiently prepare for the state exams. One would assume that such centralized exams provide the opportunity to assess student performance rather reliable across all students taking the exam because they assess the quality of answers to the same questions. Unfortunately, not only writing the state exams is highly uncontrollable, but also grading them. Although faculty do know the weeks in which they are supposed to grade the state

exams, they often receive the essays later than indicated and those weeks are typically rather busy with other tasks as well, leaving only little time for grading, which may prevent some faculty from doing so diligently. It is natural that expectations and interpretations vary between individuals and also faculty, but when all of them are supposed to grade answers to the same questions without being provided with information on what the author of a question meant and what can be expected from students, the assigned grades will likely vary between graders given similar content quality. Moreover, faculty are in some cases not familiar with the topics they are supposed to grade because they are not working in the field and would in an ideal world familiarize themselves with the topic before grading. Due to time constraints, this is sometimes not possible. For both problems a detailed grading rubric including expected answers that needs to be submitted alongside the questions to the ministry, may help to reduce at least some of the variation between graders, provide faculty that are unfamiliar with the topic quickly with the necessary knowledge to assess performance, and make grading more time-efficient when faculty only need to tick off correct answers or assign a number on a specific criterion rather than having to work out what their expectations are, how they want to weigh different aspects, and how to communicate the results. It may further be conducive to more reliable grading to have some common criteria for good writing within the rubrics for all tasks, such as clarity and structure, grammar, and quality of citations, rather than leaving weighing these aspects to each grader who may place different importance on them. In addition to providing a grading rubric it may also be important to find a common ground between universities and their faculty about the question which topics are deemed to be important for all student teachers and which views on the topics may be agreed upon by everyone involved in grading. This is necessary to bypass the problem that a student may be taught for example how bullying in class should be handled based on the insights one faculty member may have gained based on a heterogeneous body of research, but is graded by a

faculty member of another university who reached a different view on the same topic and may therefore judge the answer as incorrect. These suggestions may be little first steps into the direction of making the taking and grading of state exams in Bavaria a little less unpleasant for both student teachers and faculty.

But what can be done if the grading task itself cannot or only to a very small extent be influenced by a grader? In such cases, it seems highly unlikely that the grading activity itself can be made less unpleasant. But maybe it is possible even in such highly aversive situations to organize the grading process in such a way that it is a little less unpleasant. For instance, it may be helpful to break grading down into smaller chunks and to reward oneself after having reached certain interim goals to create at least some form of extrinsic motivation. Grading smaller portions may also prevent the development of boredom to some extent, because the repetitiveness when grading a lot of essays in a row is reduced. It may also help to grade in a nice environment, for instance by meeting with colleagues that need to grade as well and share the same problems (increasing the satisfaction of the need for relatedness), creating a room that one feels comfortable in, or turning on some background music to get through the stack of paper (some form of reducing task aversiveness). Especially if grading cannot be made more pleasant, it may be a sign of appreciation of the mental and emotional effort and the time that teachers and faculty invest in grading written work, if institutions made this cumbersome work more visible. This may be realized by conferring awards for exemplary grading practices or even providing financial compensation for extensive grading duties.

Perspectives on Teaching

Synchronous (Emergency) Online Teaching

With respect to online teaching, a differentiation between school and university settings is necessary. While online teaching in higher education has constantly grown over the last years, this was not common in schools, which focused more on integrating digital

devices in teaching practices. Nevertheless, the onset of the COVID-19 pandemic boosted online teaching in both schools and universities immediately. Based on our findings and generally in line with previous research, online teaching at higher education institutions during the COVID-19 pandemic was a less pleasant experience for students (Kanning & Ohlms, 2021) and faculty (more specifically qualified by fewer satisfaction of the needs for autonomy, competence, and relatedness, and fewer positive and more negative emotions; Study 3) than face-to-face teaching before the pandemic. Leaving aside the age difference of students and the more compulsory nature of schooling, teaching is the same task for teachers and faculty, which makes it likely that teachers perceived their situation similarly. And although our study provided some first hints at the emotional experiences and the satisfaction of the basic needs of faculty in synchronous emergency online teaching, the circumstances that foster positive experiences in online teaching largely remain to be further explored. We could show that aspects of a synchronous online learning environment hold the potential to foster basic need satisfaction in faculty (e.g., more activated student cameras tended to increase faculty members' perceived relatedness with students). Based on this first hint, it is important to identify further aspects that may contribute to positive experiences in online teaching. From a self-determination perspective this may be possibly by improving the perceived autonomy, competence, and relatedness of teachers and faculty. For instance, institutions could provide more varied or more powerful digital tools and support offers that help faculty to improve their teaching skills to be able to use more digital tools on the one hand, but also to use those available more flexibly and in line with their teaching conceptions. It may also be helpful to simplify the access to tools, for instance by identical login information for different tools or even a platform that aggregates the offers and enables access to all tools with one login. Additionally, trying to build connections with students seems to be imperative to a good online classroom climate (this may probably be even more

important in smaller classes than in lecture settings). Overall, it seems therefore important that faculty include interactive elements in their teaching (e.g., live polling, whiteboards, etherpads, breakout sessions) and encourage students to activate their cameras, provide visual feedback using the often-available emoticons, speak up actively or use the chat to create a feeling of working on a topic together. Facilitating the exchange among colleagues about their experiences with certain methods in online teaching may not only create an opportunity to learn new methods but also foster the relatedness with colleagues and thereby contribute to more positive experiences.

And although SDT provides some valuable insights on how to foster positive emotional experiences in synchronous online teaching, it remains to be explored whether the three need satisfaction dimensions can accurately explain the occurrence of specific discrete emotions. It may be possible that control-value theory reasoning (especially effects of value appraisals) may be needed to explain the emergence of teaching emotions in online teaching settings in more detail. For example, it is likely that the satisfaction of the needs for autonomy and competence contribute to control appraisals (e.g., choosing one's preferred teaching approach and applying methods one feels able to implement successfully) and that the satisfaction of the need for relatedness may contribute to positive value to some extent (e.g., feeling connected creates a pleasant situation). But what need satisfaction may not influence is how important and valuable (apart from pleasant) teachers and faculty perceive a task, in this case teaching, to be. Nevertheless, exploring such connections seems to be helpful in generating more comprehensive knowledge about the emergence of discrete teaching emotions in different settings.

Having seen that emergency online teaching was less pleasant than face-to-face teaching but also having outlined that there may be various ways to make online teaching more pleasant, the question is whether we want to move back to the state of affairs before the

COVID-19 pandemic or if we want to build on the gathered experiences and use them to improve teaching and learning after the pandemic? But before we can tell how education may look like after the pandemic, it is necessary to clarify how the way to get there could look like.

Transition to Post-Pandemic Teaching

Political debates about school teaching clearly prioritize pure face-to-face teaching in all school types even if the pandemic situation worsens again and post-pandemic online teaching offers are not considered as an option. Taking into account that teaching and learning in general and in primary schools in particular are often dependent on direct interactions with teachers, showing and directly discussing things, and performing activities together, the urge to ensure classroom teaching in schools seems justified. It would additionally be challenging for some families to ensure the supervision of minors when parents need to work, to provide an adequate learning environment (i.e., desk in an own and silent room), and to support self-directed learning also in post-pandemic distance learning. Overall, the return to regular classroom teaching and learning seems imperative to ensure learning at school and providing students with as equal chances of participating in education as possible. Because face-to-face teaching is the largest part of teachers' jobs that they typically enjoy, the return will also contribute to more positive (emotional) experiences for school teachers.

The situation at universities is a little different, though. Although political debates about how or when face-to-face teaching should be resumed in universities are practically not existent in Germany, universities do discuss the next steps. Although the majority of faculty and students seem to prefer switching back to face-to-face teaching once it is possible, this leaves 15 to 20 percent of faculty and students in higher education who would still prefer to teach and learn online after the COVID-19 pandemic, respectively (Kanning & Ohlms,

2021). During the transition back to mainly face-to-face teaching it is important to consider two sides. On the one hand, it seems important to make face-to-face offers possible for faculty and students who want to savor the experience of being together in one room and engage in learning activities together, and on the other hand to respect the need of some faculty and students to offer and attend classes online because they are at risk of having a severe course of disease and do not want to risk an infection when attending face-to-face classes or because students had to move back home due to financial difficulties caused by the pandemic and cannot commute to university. Therefore, at the present time it seems unwise to switch back to face-to-face teaching as rapidly as switching from face-to-face to emergency online teaching during the onset of the pandemic. Therefore, in the current situation online (if faculty need to avoid face-to-face teaching) or hybrid offers (if students need to avoid face-to-face learning yet attend classes) seem to be the necessary next steps towards post-pandemic higher education. Especially with respect to hybrid offers (i.e., some students being in class and some attending the same class synchronously from home) empirical research is scarce, but likely similar approaches to improve teaching and learning as in synchronous settings can be applied. Nevertheless, it seems important that faculty prepare for this new challenge and that institutions support them in this endeavor and provide at least the technical equipment to realize this teaching format.

Post-Pandemic (Online) Teaching

Thinking ahead to a post-pandemic future, complementary online and hybrid offers could probably open the door to more flexible work and study models. For instance, faculty could work from home even on days when they need to teach or more flexibly distribute their class preparation according to their needs if they offered asynchronous online classes. Moreover, scientific exchange may become easier over further distances because external lecturers could offer workshops, classes, and especially short guest lectures without the

necessity to travel to the university, which is always associated with a great investment of time an money. Faculty may even be given the freedom to move elsewhere while being able to continue their work position. Especially those students who have children or need to work a lot to finance their studies may profit from flexible online offers with mainly asynchronous elements because it would be easier to combine their personal life with their studies. Online offers may additionally pave the way to more part-time study programs without a lot of extra effort, because students could choose to attend face-to-face (or hybrid) classes if the timeslots suit their schedule or move to asynchronous online offers if they do not. And despite these more flexible offers, universities would not have to offer extra classes in the evenings or on weekends to cater for such special needs.

To realize such ideas, it is necessary to learn more about how asynchronous online offers can be an effective and enjoyable endeavor for students and faculty alike. Assuming that faculty may choose to deliver classes face-to-face, hybrid, or online after the pandemic, depending on their own preferences and institutional demands, it seems likely that only those students and faculty who want to engage in online classes participate in these offers. Therefore, online learning and teaching should not be as aversive as forced emergency online learning and teaching and motivation levels should typically be higher. These are good preconditions to create good online learning and teaching environments. From a theoretical perspective, satisfying the basic needs of students and faculty seems promising to foster positive asynchronous online experiences. For instance, to foster competence it may be beneficial to have clear structures and a course outline right from the beginning so that everyone knows what is coming next and feels able to prepare and follow the content and activities. Additionally, offering some flexibility in submitting assignments or the opportunity to choose between assignments may foster perceived autonomy in students. Experiencing autonomy in designing the assignments and the different activities may foster perceived

autonomy in faculty. Building connections may be especially challenging in asynchronous settings. Nevertheless, faculty may foster students' relatedness among each other by encouraging the use of a profile picture and indicating some personal information (e.g., hobbies, spoken languages) in the learning platform profiles, providing opportunities for exchange (e.g., chat, forum), and initiating group activities in which students need to collaborate and exchange ideas (e.g., group works, peer feedback). They may enhance the relatedness of students with them by replying to questions quickly, share instructions through a video recording rather than in writing, and offering office hours to meet synchronously. Some of these ideas probably foster the relatedness between students and faculty in both directions. To increase their satisfaction of the need for relatedness, faculty may ask students to provide short videos, either introducing themselves or presenting a topic they were working on during class, and also try to connect with colleagues, for instance by co-teaching classes, sharing activities or input that may be used in similar courses, and generally exchanging their experiences and best practices in online teaching. These are but a few examples that may have the potential to foster good learning and teaching in asynchronous online classes, but they need an empirical foundation to find out which factors contribute to successful and enjoyable learning and teaching in asynchronous online settings.

Nevertheless, maintaining complementary hybrid or online offers after the COVID-19 pandemic may not be the only positive aspect of the unexpectedly long excursion into a world of digital higher education. It may even be possible to transfer some of the elements that faculty started to value in online environments to face-to-face settings. Especially in large lecture settings with hundreds of students in a lecture hall it has typically been challenging to interact with students, for instance because a microphone was needed to hear them, students did not want to speak up in such large groups, and overviews of opinions were limited to yes or no questions that could be answered by raising one's hand. To enhance the

interactivity especially in large or hybrid group settings, it may be possible to use some online tools known from the online meeting tools in face-to-face teaching (possibly faculty would have to use tools that are available outside the meeting tool, but plenty are available). Some good options may be live polling (e.g., for getting an overview of opinions on comparably open questions, finding out whether certain concepts have been understood, practicing multiple choice exam questions, etc.), working on shared whiteboards (e.g., brainstorming ideas, sharing short answers, commenting on others' ideas, etc.), or online mind maps (e.g., for visualizing relationships between different constructs that have been covered over the course of the academic term together).

Overall, the pandemic fostered progress in online teaching and learning and it is up to institutions, students, teachers, and faculty to shape the future. They need to decide which valuable elements will be integrated in post-pandemic face-to-face teaching across all educational institutions and which complementary online-based offers will be added to post-pandemic higher education.

Perspectives on Education

Institutional education may be seen as an attempt to provide all students with equal learning opportunities and equal chances for further education and career opportunities. Throughout the course of education, educational decisions set the direction for further educational opportunities. And unfortunately, just like single grades are biased by manifold factors (e.g., handwriting quality and discrete emotions, see Study 1; student test anxiety, von der Embse et al., 2018; socio-economic background, Becker & Lauterbach, 2016), educational decisions are also biased, for instance by socio-economic status (Becker & Lauterbach, 2016). Socio-economic status not only affects student education directly (i.e., competencies acquired at home contribute to success expectations and performance) but also indirectly (i.e., through decisions for or against higher school tracks or higher education given

similar student performance; Becker & Lauterbach, 2016). One important educational decision in Germany regards the choice of school tracks for secondary school after grade 4. This decision is highly dependent on students' grades in grade 4 and the primary school teachers' recommendation for a specific school track and is in the end mainly made by the parents.

Taking into consideration that all three factors that contribute to this decision process are likely factors that contribute to a biased decision in one direction or the other, such educational decisions may need a critical analysis, especially when they occur very early in the educational process. The possibly inequity-supporting effects of biased grades and socioeconomic background may be countered by postponing the decision about school tracks to a later point in time or by giving up on the idea that dividing students into different schools is necessary to provide appropriate learning opportunities. Maybe a more integrative system that promotes and challenges students not based on their overall GPA and places them in a generally more or less demanding school, but on strengths and weaknesses in different subjects and places them in classes with an optimal challenge level within the same school, may be able to cater better for the individual needs of students.

But no matter how one wants to optimally foster students based on their performance, the question how performance should be defined and assessed remains. The current view on performance seems to place more emphasis on separating students based on their grades rather than to encourage them to develop, improve, and achieve together on different performance levels. Although it will be almost impossible to implement such approaches in school systems, it is worth to think about alternative ways of assessing performance, especially when taking into account that test situations are highly unpleasant and sometimes anxiety-inducing for students, grading situations are aversive, stressful, and emotionally unfavorable for teachers, and that in the end the grades may not even be a reliable and valid

estimation of true performance (cf. test anxiety and grading bias). Maybe it would be important to take into consideration for instance students' motivation, effort, and social interaction with classmates to derive at a more holistic estimation of classroom performance or competence than is possible with a grade from an examination trying to capture a momentary state of knowledge. Maybe it would be necessary to define achievement as knowledge gains, the ability to actively engage with content, and incorporate feedback rather than as an absolute level of performance. I do not have a solution at hand how to improve a whole educational system, but I believe that there is more to learning and teaching than knowledge transmission and more to performance and competence than knowledge testing.

Conclusion

Teachers and faculty are at the heart of education and their well-being and satisfaction are important in and of themselves and moreover contribute to maintaining a working educational system. To foster positive emotional experiences at work it seems important that teachers and faculty perceive themselves to be autonomous, competent, and connected to others and that they see positive value in the tasks they are performing. As in every other occupation, teachers and faculty need to accomplish tasks they do like and tasks they do not like. Grading is one of the tasks they typically do not like, on which they have to spend a comparably large amount of time, though. Therefore, this dissertation attended to the occurrence, effects, and especially antecedents of grading emotions. It can be summarized that emotions that are either triggered by the grading task itself or by other sources likely have effects on grading and grades. Although the use of grades as performance measure may be questioned rather generally, it seems important to make grading as pleasant as possible to mitigate its possibly detrimental effects for teachers' and faculty members' emotional lives. To this end, teachers and faculty may design grading in such a way that it is useful for teaching and contributes to student learning, reappraise the grading situation and try to find

value in grading, or create their own reward structures to facilitate extrinsic motivation. Institutions may contribute to make grading less thankless by valuing the task more, for instance by awarding grading awards for exemplary grading practices or providing monetary compensations for extraordinary high workload with respect to grading.

Although the working conditions of teachers and faculty alike may be stressful (and even more so during a time of pandemic), at least the workplace of faculty is typically better equipped. Universities typically provide the technical equipment they need to work (e.g., laptops, software, printers, etc.) and also an office at university where they can work on tasks outside the classroom (e.g., design studies, write manuscripts, prepare lessons, grade exams, offers office hours for students, etc.). This is usually not the case for teachers who often do not have a room to work in within the school building to accomplish job duties that take place outside the classroom. By having to take work home and using private devices to perform their job duties, the lines between work and private life may blur more easily and may make it even harder to maintain a healthy work-life balance. In addition to improving the work environment, it would be beneficial to improve teaching conditions for teachers and faculty by having smaller classes or a teaching assistant to be better able to cater for the individual needs of students and to be able to facilitate student learning more effectively.

Because teachers lay the foundation of education and their working conditions are partially detrimental to their educational mission (possibly even more so than for faculty), it seems imperative to push forward measures that increase well-being and decrease high stress levels and burnout rates. And we should exhaust all possibilities to support teachers within the current system, so that they can concentrate on pursuing their jobs with enthusiasm and heart blood to become a “champion” who teaches anyway (Pierson, 2013), one of the exceptionally good teachers that you remember when you think back to your own education.

6. References

Acee, T., Kim, H., Kim, J.-I., Chu, H.-N., Kim, M., Cho, Y., & Wicker, F. (2010). Academic boredom in under- and over-challenging situations. *Contemporary Educational Psychology*, *35*(1), 17–27. https://doi.org/10.1016/j.cedpsych.2009.08.002

Akcaoglu, M., & Lee, E. (2016). Increasing social presence in online learning through small group discussions. *The International Review of Research in Open and Distributed Learning*, *17*(3), 1–17. https://doi.org/10.19173/irrodl.v17i3.2293

Allensworth, E. M., & Clark, K. (2020). High school GPAs and ACT scores as predictors of college completion: Examining assumptions about consistency across high schools. *Educational Researcher*, *49*(3), 198–211. https://doi.org/10.3102/0013189X20902110

Aristovnik, A., Keržič, D., Ravšelj, D., Tomaževič, N., & Umek, L. (2020). Impacts of the COVID-19 pandemic on life of higher education students: A global perspective. *Sustainability*, *12*(20), Article 8438. https://doi.org/10.3390/su12208438

Babb, J., & Corbett, S. J. (2016). From zero to sixty: A survey of college writing teachers' grading practices and the affect of failed performance. *Composition Forum*, *34*. https://compositionforum.com/issue/34/zero-sixty.php

Baker, N. L. (2014). "Get it off my stack": Teachers' tools for grading papers. *Assessing Writing*, *19*, 36–50. https://doi.org/10.1016/j.asw.2013.11.005

Bauer, J., Unterbrink, T., Hack, A., Pfeifer, R., Buhl-Grießhaber, V., Müller, U., Wesche, H., Frommhold, M., Seibt, R., Scheuch, K., & Wirsching, M. (2007). Working conditions, adverse events and mental health problems in a sample of 949 German teachers. *International Archives of Occupational and Environmental Health*, *80*(5), 442–449. https://doi.org/10.1007/s00420-007-0170-7

Becker, R., & Lauterbach, W. (Eds.). (2016). *Bildung als Privileg*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-11952-2

Beilock, S. L., Gunderson, E. A., Ramirez, G., & Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proceedings of the National Academy of Sciences*, *107*(5), 1860–1863. https://doi.org/10.1073/pnas.0910967107

Besser, A., Flett, G. L., & Zeigler-Hill, V. (2020). Adaptability to a sudden transition to online learning during the COVID-19 pandemic: Understanding the challenges for students. *Scholarship of Teaching and Learning in Psychology*. Advance online publication. https://doi.org/10.1037/stl0000198

Borup, J., West, R. E., & Graham, C. R. (2012). Improving online social presence through asynchronous video. *The Internet and Higher Education*, *15*(3), 195–203. https://doi.org/10.1016/j.iheduc.2011.11.001

Bowers, A. J. (2010). Grades and graduation: A longitudinal risk perspective to identify student dropouts. *The Journal of Educational Research*, *103*(3), 191–207. https://doi.org/10.1080/00220670903382970

Brackett, M. A., Floman, J. L., Ashton-James, C., Cherkasskiy, L., & Salovey, P. (2013). The influence of teacher emotion on grading practices: A preliminary look at the evaluation of student writing. *Teachers and Teaching*, *19*(6), 634–646. https://doi.org/10.1080/13540602.2013.827453

Briggs, D. (1970). The influence of handwriting on assessment. *Educational Research*, *13*(1), 50–55. https://doi.org/10.1080/0013188700130107

Briggs, D. (1980). A study of the influence of handwriting upon grades using examination scripts. *Educational Review*, *32*(2), 186–193. https://doi.org/10.1080/0013191800320207

Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., Stevens, M. T., & Welsh, M. E. (2016). A century of grading research: Meaning and value in the most common educational measure. *Review of Educational Research*, *86*(4), 803–848. https://doi.org/10.3102/0034654316672069

Buhr, E. E., Daniels, L. M., & Goegan, L. D. (2019). Cognitive appraisals mediate relationships between two basic psychological needs and emotions in a massive open online course. *Computers in Human Behavior*, *96*, 85–94. https://doi.org/10.1016/j.chb.2019.02.009

Butz, N. T., & Stupnisky, R. H. (2016). A mixed methods study of graduate students' self-determined motivation in synchronous hybrid learning environments. *The Internet and Higher Education*, *28*, 85–95. https://doi.org/10.1016/j.iheduc.2015.10.003

Butz, N. T., Stupnisky, R. H., Peterson, E. S., & Majerus, M. M. (2014). Motivation in synchronous hybrid graduate business programs: A self-determination approach to contrasting online and on-campus students. *Journal of Online Learning & Teaching*, *10*(2), 211–227. https://jolt.merlot.org/Vol10_No2.htm

Cadez, S., Dimovski, V., & Zaman Groff, M. (2017). Research, teaching and performance evaluation in academia: The salience of quality. *Studies in Higher Education*, *42*(8), 1455–1473. https://doi.org/10.1080/03075079.2015.1104659

Center for Postsecondary Research. (2020). *Faculty Survey of Student Engagement [Survey Instrument]*. https://fsse.indiana.edu

Cizek, G. J., Fitzgerald, S. M., & Rachor, R. A. (1995). Teachers' assessment practices: Preparation, isolation, and the kitchen sink. *Educational Assessment*, *3*(2), 159–179. https://doi.org/10.1207/s15326977ea0302_3

Crick, K. A., Larson, L. M., & Seipel, M. T. (2020). Non-tenure track faculty satisfaction: A self-determination model. *Journal of Career Assessment*, *28*(3), 425–445. https://doi.org/10.1177/1069072719870681

Daniels, L. M., & Stupnisky, R. H. (2012). Not that different in theory: Discussing the control-value theory of emotions in online learning environments. *The Internet and Higher Education*, *15*(3), 222–226. https://doi.org/10.1016/j.iheduc.2012.04.002

Daumiller, M., Janke, S., Rinas, R., Hein, J., Dickhäuser, O., & Dresel, M. (2019). *Temporal variability and domain specificity of university instructors' achievement goals and associations with affective experiences*. PsyArXiv. https://doi.org/10.31234/osf.io/p4nhu

Deci, E. L., & Ryan, R. M. (2004). *Handbook of self-determination research*. University Rochester Press.

Downing, J. J., & Dyment, J. E. (2013). Teacher educators' readiness, preparation, and perceptions of preparing preservice teachers in a fully online environment: An exploratory study. *The Teacher Educator*, *48*(2), 96–109. https://doi.org/10.1080/08878730.2012.760023

Duncan, C. R., & Noonan, B. (2007). Factors affecting teachers' grading and assessment practices. *Alberta Journal of Educational Research*, *53*(1), 1–21. https://cjc-rcc.ucalgary.ca/index.php/ajer/article/view/55195

Eames, K., & Loewenthal, K. (1990). Effects of handwriting and examiner's expertise on assessment of essays. *The Journal of Social Psychology*, *130*(6), 831–833. https://doi.org/10.1080/00224545.1990.9924637

Ebersold, S., Rahm, T., & Heise, E. (2019). Autonomy support and well-being in teachers: Differential mediations through basic psychological need satisfaction and frustration. *Social Psychology of Education*, *22*(4), 921–942. https://doi.org/10.1007/s11218-019-09499-1

Esdar, W., Gorges, J., & Wild, E. (2016). The role of basic need satisfaction for junior academics' goal conflicts and teaching motivation. *Higher Education*, *72*(2), 175–190. https://doi.org/10.1007/s10734-015-9944-0

Euzent, P., Martin, T., Moskal, P., & D. Moskal, P. (2011). Assessing student performance and perceptions in lecture capture vs. face-to-face course delivery. *Journal of Information Technology Education: Research*, *10*(1), 295–307. https://doi.org/10.28945/1515

Filak, V. F., & Nicolini, K. M. (2018). Differentiations in motivation and need satisfaction based on course modality: A self-determination theory perspective. *Educational Psychology*, *38*(6), 772–784. https://doi.org/10.1080/01443410.2018.1457776

Frenzel, A. C. (2014). Teacher emotions. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 494–519). Routledge, Taylor & Francis Group.

Frenzel, A. C., Becker-Kurz, B., Pekrun, R., & Goetz, T. (2015). Teaching this class drives me nuts! - Examining the person and context specificity of teacher emotions. *PLOS ONE*, *10*(6), Article e0129630. https://doi.org/10.1371/journal.pone.0129630

Frenzel, A. C., Becker-Kurz, B., Pekrun, R., Goetz, T., & Lüdtke, O. (2017). Emotion transmission in the classroom revisited: A reciprocal effects model of teacher and student enjoyment. *Journal of Educational Psychology*, *110*(5), 628–639. https://doi.org/10.1037/edu0000228

Frenzel, A. C., Goetz, T., Lüdtke, O., Pekrun, R., & Sutton, R. E. (2009). Emotional transmission in the classroom: Exploring the relationship between teacher and student enjoyment. *Journal of Educational Psychology*, *101*(3), 705–716. https://doi.org/10.1037/a0014695

Frenzel, A. C., Goetz, T., Stephens, E. J., & Jacob, B. (2009). Antecedents and effects of teachers' emotional experiences: An integrated perspective and empirical test. In P. A. Schutz & M. Zembylas (Eds.), *Advances in Teacher Emotion Research* (pp. 129–151). Springer US. https://doi.org/10.1007/978-1-4419-0564-2_7

Frenzel, A. C., Pekrun, R., Goetz, T., Daniels, L. M., Durksen, T. L., Becker-Kurz, B., & Klassen, R. M. (2016). Measuring teachers' enjoyment, anger, and anxiety: The Teacher Emotions Scales (TES). *Contemporary Educational Psychology*, *46*, 148–163. https://doi.org/10.1016/j.cedpsych.2016.05.003

Frenzel, A. C., & Stephens, E. J. (2017). Emotionen. In T. Götz (Ed.), *Emotion, Motivation und selbstreguliertes Lernen* (2., aktualisierte Auflage, pp. 15–77). Ferdinand Schöningh.

García-Carmona, M., Marín, M. D., & Aguayo, R. (2019). Burnout syndrome in secondary school teachers: A systematic review and meta-analysis. *Social Psychology of Education*, *22*(1), 189–208. https://doi.org/10.1007/s11218-018-9471-9

Garris, C. P., & Fleck, B. (2020). Student evaluations of transitioned-online courses during the COVID-19 pandemic. *Scholarship of Teaching and Learning in Psychology*. Advance online publication. https://doi.org/10.1037/stl0000229

Gaspard, H., Dicke, A.-L., Flunger, B., Schreier, B., Häfner, I., Trautwein, U., & Nagengast, B. (2015). More value through greater differentiation: Gender differences in value beliefs about math. *Journal of Educational Psychology*, *107*(3), 663–677. https://doi.org/10.1037/edu0000003

Goetz, T., Sticca, F., Pekrun, R., Murayama, K., & Elliot, A. J. (2016). Intraindividual relations between achievement goals and discrete achievement emotions: An experience sampling approach. *Learning and Instruction*, *41*, 115–125. https://doi.org/10.1016/j.learninstruc.2015.10.007

Greifeneder, R., Alt, A., Bottenberg, K., Seele, T., Zelt, S., & Wagener, D. (2010). On writing legibly: Processing fluency systematically biases evaluations of handwritten material. *Social Psychological and Personality Science*, *1*(3), 230–237. https://doi.org/10.1177/1948550610368434

Greifeneder, R., Zelt, S., Seele, T., Bottenberg, K., & Alt, A. (2012). Towards a better understanding of the legibility bias in performance assessments: The case of gender-based inferences: Handwriting legibility and gender-based inferences. *British Journal of Educational Psychology*, *82*(3), 361–374. https://doi.org/10.1111/j.2044-8279.2011.02029.x

Hagenauer, G., Hascher, T., & Volet, S. E. (2015). Teacher emotions in the classroom: Associations with students' engagement, classroom discipline and the interpersonal teacher-student relationship. *European Journal of Psychology of Education*, *30*(4), 385–403. https://doi.org/10.1007/s10212-015-0250-0

Hagenauer, G., & Volet, S. (2014). 'I don't think I could, you know, just teach without any emotion': Exploring the nature and origin of university teachers' emotions. *Research Papers in Education*, *29*(2), 240–262. https://doi.org/10.1080/02671522.2012.754929

Hakanen, J. J., Bakker, A. B., & Schaufeli, W. B. (2006). Burnout and work engagement among teachers. *Journal of School Psychology*, *43*(6), 495–513. https://doi.org/10.1016/j.jsp.2005.11.001

Hargreaves, A. (1998). The emotional practice of teaching. *Teaching and Teacher Education*, *14*(8), 835–854. https://doi.org/10.1016/S0742-051X(98)00025-0

Holzer, J., Lüftenegger, M., Käser, U., Korlat, S., Pelikan, E., Schultze-Krumbholz, A., Spiel, C., Wachs, S., & Schober, B. (2021). Students' basic needs and well-being during the COVID-19 pandemic: A two-country study of basic psychological need satisfaction, intrinsic learning motivation, positive emotion and the moderating role of self-regulated learning. *International Journal of Psychology*. Advance online publication. https://doi.org/10.1002/ijop.12763

Hughes, D. C., & Keeling, B. (1984). The use of model essays to reduce context effects in essay scoring. *Journal of Educational Measurement*, *21*(3), 277–281.

Hughes, D. C., Keeling, B., & Tuck, B. F. (1980). The influence of context position and scoring method on essay scoring. *Journal of Educational Measurement*, *17*(2), 131–134.

Ilott, I., & Murphy, R. (1997). Feelings and failing in professional training: The assessor's dilemma. *Assessment & Evaluation in Higher Education*, *22*(3), 307–316. https://doi.org/10.1080/0260293970220304

James, H. W. (1927). The effect of handwriting upon grading. *The English Journal*, *16*(3), 180–185. https://doi.org/10.2307/803599

Jang, H., Reeve, J., Ryan, R. M., & Kim, A. (2009). Can self-determination theory explain what underlies the productive, satisfying learning experiences of collectivistically oriented Korean students? *Journal of Educational Psychology*, *101*(3), 644–661. https://doi.org/10.1037/a0014241

Janke, S., & Dickhäuser, O. (2018). A situated process model of vocational achievement goal striving within members of the academic staff at university. *Motivation and Emotion*, *42*(4), 466–481. https://doi.org/10.1007/s11031-017-9657-z

Jarosz, A. F., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting Bayes factors. *The Journal of Problem Solving*, *7*(1), 2–9. https://doi.org/10.7771/1932-6246.1167

Kajfez, R. L., & Matusovich, H. M. (2017). Competence, autonomy, and relatedness as motivators of graduate teaching assistants. *Journal of Engineering Education*, *106*(2), 245–272. https://doi.org/10.1002/jee.20167

Kanning, U. P., & Ohlms, M. (2021). Hochschullehre in Zeiten von Corona. *Wirtschaftspsychologie*, *2*, 44–55.

Keller, M. M., Frenzel, A. C., & Goetz, T. (2014). Exploring teacher emotions: A literature review and an experience sampling study. In P. W. Richardson, S. A. Karabenick, & H. M. G. Watt (Eds.), *Teacher motivation: Theory and practice* (pp. 69–82). Routledge.

Klassen, R. M., Perry, N. E., & Frenzel, A. C. (2012). Teachers' relatedness with students: An underemphasized component of teachers' basic psychological needs. *Journal of Educational Psychology*, *104*(1), 150–165. https://doi.org/10.1037/a0026253

Kordts-Freudinger, R. (2017). Feel, think, teach – Emotional underpinnings of approaches to teaching in higher education. *International Journal of Higher Education*, *6*(1), 217–229. https://doi.org/10.5430/ijhe.v6n1p217

Kuladinithi, K., Fisser, L., Fuger, K., Stolpmann, D., Vatandas, Z., Timm-Giel, A., & Dürkop, A. (2020). Online teaching of project-based learning courses—Issues, challenges and outcomes. *ACM Special Interest Group on Data Communication (SIGCOMM)*. https://doi.org/10.15480/882.2863

Lahtinen, A. (2008). University teachers' views on the distressing elements of pedagogical interaction. *Scandinavian Journal of Educational Research*, *52*(5), 481–493. https://doi.org/10.1080/00313830802346363

Larson, L. M., Seipel, M. T., Shelley, M. C., Gahn, S. W., Ko, S. Y., Schenkenfelder, M., Rover, D. T., Schmittmann, B., & Heitmann, M. M. (2019). The academic environment and faculty well-being: The role of psychological needs. *Journal of Career Assessment*, *27*(1), 167–182. https://doi.org/10.1177%2F1069072717748667

Laybourn, S., Frenzel, A. C., & Fenzl, T. (2019). Teacher procrastination, emotions, and stress: A qualitative study. *Front. Psychol.*, 10:2325. https://doi.org/10.3389/fpsyg.2019.02325

Leiner, D. J. (2019). *SoSci Servey* (3.1.06) [Computer software]. https://www.soscisurvey.de

Löfström, E., & Nevgi, A. (2013). Giving shape and form to emotion: Using drawings to identify emotions in university teaching. *International Journal for Academic Development*, *19*(2), 99–111. https://doi.org/10.1080/1360144X.2013.819553

Loh, C. E., & Liew, W. M. (2016). Voices from the ground: The emotional labour of English teachers' work. *Teaching and Teacher Education*, *55*, 267–278. https://doi.org/10.1016/j.tate.2016.01.016

Lowenthal, P. R. (2010). The evolution and influence of social presence theory on online learning. In T. T. Kidd (Ed.), *Online education and adult learning: New frontiers for teaching practices* (pp. 124–139). Information Science Reference.

Malouff, J. M., & Thorsteinsson, E. B. (2016). Bias in grading: A meta-analysis of experimental research findings. *Australian Journal of Education*, *60*(3), 245–256. https://doi.org/10.1177/0004944116664618

Marchand, G. C., & Gutierrez, A. P. (2012). The role of emotion in the learning process: Comparisons between online and face-to-face learning settings. *The Internet and Higher Education*, *15*(3), 150–160. https://doi.org/10.1016/j.iheduc.2011.10.001

Marinoni, G., Van't Land, H., & Jensen, T. (2020). *The impact of COVID-19 on higher education around the world. IAU Global Survey Report*. International Association of Universities.

Markham, L. R. (1976). Influences of handwriting quality on teacher evaluation of written work. *American Educational Research Journal*, *13*(4), 277–283. https://doi.org/10.2307/1162390

Meanwell, E., & Kleiner, S. (2014). The emotional experience of first-time teaching: Reflections from graduate instructors, 1997–2006. *Teaching Sociology*, *42*(1), 17–27. https://doi.org/10.1177/0092055X13508377

Meishar-Tal, H., & Levenberg, A. (2021). In times of trouble: Higher education lecturers' emotional reaction to online instruction during COVID-19 outbreak. *Education and Information Technologies*. Advance online publication. https://doi.org/10.1007/s10639-021-10569-1

Milyavskaya, M., & Koestner, R. (2011). Psychological needs, motivation, and well-being: A test of self-determination theory across multiple domains. *Personality and Individual Differences*, *50*(3), 387–391. https://doi.org/10.1016/j.paid.2010.10.029

Montoya, A. K., & Hayes, A. F. (2017). Two-condition within-participant statistical mediation analysis: A path-analytic framework. *Psychological Methods*, *22*(1), 6–27. https://doi.org/10.1037/met0000086

Murtonen, M., & Vilppu, H. (2020). Change in university pedagogical culture – The impact of increased pedagogical training on first teaching experiences. *International Journal of Learning, Teaching and Educational Research*, *19*(3), 367–383. https://doi.org/10.26803/ijlter.19.3.20

Naylor, D., & Nyanjom, J. (2020). Educators' emotions involved in the transition to online teaching in higher education. *Higher Education Research & Development*. https://doi.org/10.1080/07294360.2020.1811645

Nie, Y., Lau, S., & Liau, A. K. (2012). The Teacher Efficacy Scale: A reliability and validity study. *The Asia-Pacific Education Researcher*, *21*(2), 414–421.

OECD. (2014). "Indicator D4: How much time do teachers spend teaching?" In *Education at a Glance 2014: OECD Indicators*. OECD Publishing. http://dx.doi.org/10.1787/888933120005

Otter, R. R., Seipel, S., Graeff, T., Alexander, B., Boraiko, C., Gray, J., Petersen, K., & Sadler, K. (2013). Comparing student and faculty perceptions of online and traditional courses. *The Internet and Higher Education*, *19*, 27–35. https://doi.org/10.1016/j.iheduc.2013.08.001

Padrón, I., Fraga, I., Vieitez, L., Montes, C., & Romero, E. (2021). A study on the psychological wound of COVID-19 in university students. *Front. Psychol.*, 12:589927. https://doi.org/10.3389/fpsyg.2021.589927

Peixoto, F., Sanches, C., Mata, L., & Monteiro, V. (2017). "How do you feel about math?": Relationships between competence and value appraisals, achievement emotions and academic achievement. *European Journal of Psychology of Education*, *32*(3), 385–405. https://doi.org/10.1007/s10212-016-0299-4

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, *18*(4), 315–341. https://doi.org/10.1007/s10648-006-9029-9

Pekrun, R. (2018). Control-value theory: A social-cognitive approach to achievement emotions. In G. A. D. Liem & D. M. McInerney (Eds.), *Big theories revisited 2: A volume of research on sociocultural influences on motivation and learning* (pp. 162–190). IAP - Information Age Publishing, Inc.

Pekrun, R. (2019). Inquiry on emotions in higher education: Progress and open problems. *Studies in Higher Education*, *44*(10), 1806–1811. https://doi.org/10.1080/03075079.2019.1665335

Pekrun, R., Frenzel, A. C., Goetz, T., & Perry, R. P. (2007). The control-value theory of achievement emotions: An integrative approach to emotions in education. In *Emotion in education* (pp. 13–36). Elsevier.

Pekrun, R., Goetz, T., Daniels, L. M., Stupnisky, R. H., & Perry, R. P. (2010). Boredom in achievement settings: Exploring control–value antecedents and performance outcomes of a neglected emotion. *Journal of Educational Psychology*, *102*(3), 531–549. https://doi.org/10.1037/a0019243

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology*, *36*(1), 36–48. https://doi.org/10.1016/j.cedpsych.2010.10.002

Pekrun, R., Goetz, T., & Perry, R. P. (2005). Academic Emotions Questionnaire (AEQ): User's manual. *Department of Psychology, University of Munich*.

Philipp, A., & Kunter, M. (2013). How do teachers spend their time? A study on teachers' strategies of selection, optimisation, and compensation over their career cycle. *Teaching and Teacher Education*, *35*, 1–12. https://doi.org/10.1016/j.tate.2013.04.014

Pierson, R. (2013). *Every kid needs a champion* [Video]. TED Conferences. https://www.ted.com/talks/rita_pierson_every_kid_needs_a_champion

Postareff, L., & Lindblom-Ylänne, S. (2011). Emotions and confidence within teaching in higher education. *Studies in Higher Education*, *36*(7), 799–813. https://doi.org/10.1080/03075079.2010.483279

Postareff, L., & Lindblom-Ylänne, S. (2015). What triggers emotions in university teaching? *Ammattikasvatuksen Aikakauskirja*, *2*, 83–96. https://journal.fi/akakk/article/view/90107

Qualtrics (Version 05-2019). (2019). [Computer software]. Qualtrics. https://www.qualtrics.com

R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rasch, D., Kubinger, K. D., & Moder, K. (2011). The two-sample t test: Pre-testing its assumptions does not pay off. *Statistical Papers*, *52*(1), 219–231. https://doi.org/10.1007/s00362-009-0224-x

Reber, R., & Greifeneder, R. (2017). Processing fluency in education: How metacognitive feelings shape learning, belief formation, and affect. *Educational Psychologist*, *52*(2), 84–103. https://doi.org/10.1080/00461520.2016.1258173

Regan, K., Evmenova, A., Baker, P., Jerome, M. K., Spencer, V., Lawson, H., & Werner, T. (2012). Experiences of instructors in online learning environments: Identifying and regulating emotions. *The Internet and Higher Education*, *15*(3), 204–212. https://doi.org/10.1016/j.iheduc.2011.12.001

Reyna, C., & Weiner, B. (2001). Justice and utility in the classroom: An attributional analysis of the goals of teachers' punishment and intervention strategies. *Journal of Educational Psychology*, *93*(2), 309–319. https://doi.org/10.1037/0022-0663.93.2.309

Reynolds, C. R., Livingston, R. B., & Willson, V. L. (2009). *Measurement and assessment in education* (2nd ed). Pearson.

Rodrigo-Ruiz, D. (2016). Effect of teachers' emotions on their students: Some evidence. *Journal of Education*, *3*(4), 73–79.

Rusly, N. H. M., Vijayaratnam, P., & Sivarajah, A. (2021). COVID-19 pandemic and online learning: The challenges of instructors in tertiary institutions. In S. Mangir (Ed.), *2nd International Conference on Education, Social Science, Supply Chain, Engineering, Technology and Tourism (ESSET)* (pp. 38–48). https://submit.confbay.com/download/Conference_Proceeding_Book_ESSET2021_UvSKqHm59L.pdf#page=43

Russo, J., Bobis, J., Downton, A., Feng, M., Hughes, S., Livy, S., McCormick, M., & Sullivan, P. (2021). Characteristics of high enjoyment teachers of mathematics in primary schools. *Mathematics Education Research Journal*. https://doi.org/10.1007/s13394-021-00372-z

Ruxton, G. D. (2006). The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test. *Behavioral Ecology*, *17*(4), 688–690. https://doi.org/10.1093/beheco/ark016

Salikhova, N. R., Lynch, M. F., & Salikhova, A. B. (2020). Psychological aspects of digital learning: A self-determination theory perspective. *Contemporary Educational Technology*, *12*(2), Article ep280. https://doi.org/10.30935/cedtech/8584

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information*, *44*(4), 695–729. https://doi.org/10.1177/0539018405058216

Scherer, R., Howard, S. K., Tondeur, J., & Siddiq, F. (2021). Profiling teachers' readiness for online teaching and learning in higher education: Who's ready? *Computers in Human Behavior*, *118*, Article 106675. https://doi.org/10.1016/j.chb.2020.106675

Schulz, P., & Schlotz, W. (1999). Trierer Inventar zur Erfassung von chronischem Streß (TICS): Skalenkonstruktion, teststatistische Überprüfung und Validierung der Skala Arbeitsüberlastung. *Diagnostica*, *45*(1), 8–19. https://doi.org/10.1026//0012-1924.45.1.8

Schwarzer, R. (1993). *Streß, Angst und Handlungsregulation.(3. Aufl.) Stuttgart: Kohlhammer*.

Seipel, M. T., & Larson, L. M. (2018). Supporting non-tenure-track faculty well-being. *Journal of Career Assessment*, *26*(1), 154–171. https://doi.org/10.1177/1069072716680046

Shackelford, J. L., & Maxwell, M. (2012). Sense of community in graduate online education: Contribution of learner to learner interaction. *The International Review of Research in Open and Distributed Learning*, *13*(4), 228–249. https://doi.org/10.19173/irrodl.v13i4.1339

Sheldon, K. M., & Filak, V. (2008). Manipulating autonomy, competence, and relatedness support in a game-learning context: New evidence that all three needs matter. *British Journal of Social Psychology*, *47*(2), 267–283. https://doi.org/10.1348/014466607X238797

Sheldon, K. M., & Hilpert, J. C. (2012). The balanced measure of psychological needs (BMPN) scale: An alternative domain general measure of need satisfaction. *Motivation and Emotion*, *36*(4), 439–451. https://doi.org/10.1007/s11031-012-9279-4

Shuman, V., & Scherer, K. R. (2014). Concepts and structures of emotions. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 13–35). Routledge, Taylor & Francis Group.

Song, H., Kim, J., & Luo, W. (2016). Teacher–student relationship in online classes: A role of teacher self-disclosure. *Computers in Human Behavior*, *54*, 436–443. https://doi.org/10.1016/j.chb.2015.07.037

Steinberg, C. (2008). Assessment as an "emotional practice." *English Teaching*, *7*(3), 42–64.

Stough, L. M., & Emmer, E. T. (1998). Teachers' emotions and test feedback. *International Journal of Qualitative Studies in Education*, *11*(2), 341–361. https://doi.org/10.1080/095183998236809

Stupnisky, R. H., BrckaLorenz, A., Yuhas, B., & Guay, F. (2018). Faculty members' motivation for teaching and best practices: Testing a model based on self-determination theory across institution types. *Contemporary Educational Psychology*, *53*, 15–26. https://doi.org/10.1016/j.cedpsych.2018.01.004

Stupnisky, R. H., Hall, N. C., & Pekrun, R. (2019a). The emotions of pretenure faculty: Implications for teaching and research success. *The Review of Higher Education*, *42*(4), 1489–1526. https://doi.org/10.1353/rhe.2019.0073

Stupnisky, R. H., Hall, N. C., & Pekrun, R. (2019b). Faculty enjoyment, anxiety, and boredom for teaching and research: Instrument development and testing predictors of success. *Studies in Higher Education*, *44*(10), 1712–1722. https://doi.org/10.1080/03075079.2019.1665308

Stupnisky, R. H., Pekrun, R., & Lichtenfeld, S. (2016). New faculty members' emotions: A mixed-method study. *Studies in Higher Education*, *41*(7), 1167–1188. https://doi.org/10.1080/03075079.2014.968546

Sulea, C., van Beek, I., Sarbescu, P., Virga, D., & Schaufeli, W. B. (2015). Engagement, boredom, and burnout among students: Basic need satisfaction matters more than personality traits. *Learning and Individual Differences*, *42*, 132–138. https://doi.org/10.1016/j.lindif.2015.08.018

Sun, J. C.-Y., & Rueda, R. (2012). Situational interest, computer self-efficacy and self-regulation: Their impact on student engagement in distance education: Student engagement in distance education. *British Journal of Educational Technology*, *43*(2), 191–204. https://doi.org/10.1111/j.1467-8535.2010.01157.x

Sung, E., & Mayer, R. E. (2012). Five facets of social presence in online distance education. *Computers in Human Behavior*, *28*(5), 1738–1747. https://doi.org/10.1016/j.chb.2012.04.014

Thies, K., & Kordts-Freudinger, R. (2019). University academics' state emotions and appraisal antecedents: An intraindividual analysis. *Studies in Higher Education*, *44*(10), 1723–1733. https://doi.org/10.1080/03075079.2019.1665311

Townsend, M. A., Kek, L. Y., & Tuck, B. F. (1989). The effect of mood on the reliability of essay assessment. *British Journal of Educational Psychology*, *59*, 232–240. https://doi.org/10.1111/j.2044-8279.1989.tb03094.x

Trigwell, K. (2012). Relations between teachers' emotions in teaching and their approaches to teaching in higher education. *Instructional Science*, *40*(3), 607–621. https://doi.org/10.1007/s11251-011-9192-3

Valverde-Berrocoso, J., Garrido-Arroyo, M. del C., Burgos-Videla, C., & Morales-Cevallos, M. B. (2020). Trends in educational research about e-learning: A systematic literature review (2009–2018). *Sustainability*, *12*(12), Article 5153. https://doi.org/10.3390/su12125153

Van den Broeck, A., Ferris, D. L., Chang, C.-H., & Rosen, C. C. (2016). A review of self-determination theory's basic psychological needs at work. *Journal of Management*, *42*(5), 1195–1229. https://doi.org/10.1177/0149206316632058

van Hooff, M. L. M., & van Hooft, E. A. J. (2017). Boredom at work: Towards a dynamic spillover model of need satisfaction, work motivation, and work-related boredom. *European Journal of Work and Organizational Psychology*, *26*(1), 133–148. https://doi.org/10.1080/1359432X.2016.1241769

von der Embse, N., Jester, D., Roy, D., & Post, J. (2018). Test anxiety effects, predictors, and correlates: A 30-year meta-analytic review. *Journal of Affective Disorders*, *227*, 483–493. https://doi.org/10.1016/j.jad.2017.11.048

Wang, C., Hsu, H.-C. K., Bonem, E. M., Moss, J. D., Yu, S., Nelson, D. B., & Levesque-Bristol, C. (2019). Need satisfaction and need dissatisfaction: A comparative study of online and face-to-face learning contexts. *Computers in Human Behavior*, *95*, 114–125. https://doi.org/10.1016/j.chb.2019.01.034

Wasilik, O., & Bolliger, D. U. (2009). Faculty satisfaction in the online environment: An institutional study. *The Internet and Higher Education*, *12*(3–4), 173–178. https://doi.org/10.1016/j.iheduc.2009.05.001

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070.

Watts, J., & Robertson, N. (2011). Burnout in university teaching staff: A systematic literature review. *Educational Research*, *53*(1), 33–50. https://doi.org/10.1080/00131881.2011.552235

Westrick, P. A., Le, H., Robbins, S. B., Radunzel, J. M. R., & Schmidt, F. L. (2015). College performance and retention: A meta-analysis of the predictive validities of ACT® scores, high school grades, and SES. *Educational Assessment*, *20*(1), 23–45. https://doi.org/10.1080/10627197.2015.997614

Whitehill, J., Serpell, Z., Lin, Y.-C., Foster, A., & Movellan, J. R. (2014). The faces of engagement: Automatic recognition of student engagement from facial expressions. *IEEE Transactions on Affective Computing*, *5*(1), 86–98. https://doi.org/10.1109/TAFFC.2014.2316163

Winkielman, P., Schwarz, N., Fazendeiro, T., & Reber, R. (2003). The hedonic marking of processing fluency: Implications for evaluative judgment. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 189–217). Lawrence Erlbaum.

Yarmand, M., Solyst, J., Klemmer, S., & Weibel, N. (2021). "It feels like I am talking into a void": Understanding interaction gaps in synchronous online classrooms. In Y. Kitamura, A. Quigley, K. Isbister, T. Igarashi, P. Bjørn, & S. Drucker (Eds.), *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM. https://dl.acm.org/doi/10.1145/3411764.3445240

Appendix A

Codebook Study 1

PANAS

Im Folgenden finden Sie eine Reihe von Adjektiven, anhand derer Sie beschreiben sollen, wie Sie sich *im Moment* fühlen.
Es gibt keine richtigen oder falschen Antworten. Überlegen Sie bitte nicht lange und denken daran, diejenige Antwort auszuwählen, die Ihren *allgemeinen* Gefühlszustand am besten beschreibt.

(Inofficial English translation in parantheses; Below you find a list of adjectives with which you should describe how you feel at the moment. There are no right or wrong answers. Please do not think about it long and choose the answer that describes your general emotional state best.)

Bitte tragen Sie die entsprechende Nummer neben den Adjektiven ein.
(Please add the corresponding number next to each adjective.)

Scale: 5-pt rating scale
1 = ***gar nicht*** *(not at all),* 2 = ***ein bisschen*** *(a little),* 3 = ***einigermaßen*** *(more or less)*,
4 = ***erheblich*** *(much),* 5 = ***äußerst*** *(very much)*

Im Moment fühle ich mich…*(At the moment I feel…)*					
Pan_pos1	interessiert *(interested)*	Pan_neg6	Gereizt *(irritable)*	Pan_neg8	Nervös *(nervous)*
Pan_neg1	besorgt *(distressed)*	Pan_pos6	Wach *(alert)*	Pan_pos8	Entschlossen *(determined)*
Pan_pos2	Freudig *(excited)*	Pan_neg7	Beschämt *(ashamed)*	Pan_pos9	Aufmerksam *(attentive)*
Pan_neg2	Verärgert *(upset)*	Pan_pos7	Angeregt *(inspired)*	Pan_neg9	Zittrig *(jittery)*
Pan_pos3	Stark *(strong)*	Pan_neg5	Feindselig *(hostile)*	Pan_pos10	Aktiv *(active)*
Pan_neg3	Schuldig *(guilty)*	Pan_pos4	Begeistert *(enthusiastic)*	Pan_neg10	Bekümmert *(afraid)*
Pan_neg4	Erschrocken *(scared)*	Pan_pos5	Stolz *(proud)*		

PANAS positive affect: pan_pos (includes items pan_pos1-10)

PANAS negative affect: pan_neg (includes items pan_neg1-10)

Note ***(Grade)***

Note: German university grading system (grades can be modified by .3 and .7) applied to the German school grading system (grades 1 – 6; 1 = very good, 6 = not sufficient)

Bitte vergeben Sie eine Gesamtnote für Lösung 1/Lösung 2! Kreisen Sie dazu die passendste Note ein.
(Give an overall grade for solution 1/solution 2. Circle the most appropriate grade.)

1,0 1,3 1,7 2,0 2,3 2,7 3,0 3,3 3,7 4,0 4,3 4,7 5,0 5,3 5,7 6,0

Kontrollvariablen ***(Control Variables)***

Fluency

Bitte schauen Sie sich noch einmal Lösung 1/Lösung 2 an. Wie haben Sie diese während der Korrektur wahrgenommen?
(Please look at solution 1/solution 2 again. How did you perceive it when grading?)
1 = *schwierig zu lesen (hard to read),* 6 = *einfach zu lesen (easy to read)*

Experience

Wie viel Erfahrung haben Sie mit der Korrektur von Aufsätzen
(How much experience do you have with grading essays)?
1 = *gar keine (none)*, 6 = *sehr viel (very much)*

Knowledge

Wie gut kannten Sie den bearbeiteten Fall bereits?
(How well did you already know the case study?)
1 = *gar nicht (not at all),* 6 = *sehr gut (very good)*

Emotionen und Kompetenz ***(Emotions and Competence)***

Scale: 4-pt rating scale

1 = *trifft gar nicht zu* *(does not apply at all),* 2 = *trifft eher nicht zu* *(does rather not apply),*

3 = *trifft eher zu* *(rather applies),* 4 = *trifft völlig zu* *(fully applies)*

Bitte geben Sie im Folgenden an, wie sehr die jeweiligen Aussagen auf Sie zutreffen. *(Please indicate to what extent the following statements apply to you)*
Freude ***(Enjoyment)***
Es hat mir Spaß gemacht, den Text zu korrigieren. *(I had fun grading the text)*
Das Korrigieren hat mir Freude bereitet. *(I enjoyed grading)*
Das Korrigieren machte mir so viel Spaß, dass ich richtig motiviert war, daran zu arbeiten. *(I had so much fun grading that I was really motivated to work on it)*
Ärger ***(Anger)***
Beim Korrigieren war ich verärgert. *(I was annoyed when grading)*
Am liebsten hätte ich während der Korrektur dem Verfasser des Gesamtbefundes aus Verärgerung ordentlich die Meinung gesagt. *(When grading I was so angry that I'd have liked to tell the author what I thought)*
Das Korrigieren hat mich genervt. *(Grading annoyed me)*
Langeweile ***(Boredom)***
Vor Langeweile gingen mir immer wieder Gedanken durch den Kopf, die mit der Korrektur nichts zu tun hatten. *(I was so bored that I had thoughts, which were not related to grading)*
Beim Korrigieren war mir so langweilig, dass die Zeit kaum zu vergehen schien. *(I was so bored when grading that time seemed to not pass at all)*
Beim Korrigieren war ich gelangweilt. *(When grading I was bored)*
Kompetenz ***(Competence)***
Beim Korrigieren habe ich mich unsicher gefühlt, ob ich den Text überhaupt adäquat bewerten kann. *(When grading I felt insecure, if I was able to judge the text adequately) *recoded*
Ich fühlte mich kompetent, den Gesamtbefund zu korrigieren. *(I felt competent to grade the essay)*

Appendix B

Codebook Study 2a (English only)

Emotions in Three Different Contexts (Same Items Reframed for RESEARCH/TEACHING/GRADING)

In the following questions, by "grading student papers", we mean assessing the quality of students' work on assigned papers, reports, or other writing tasks (submitted as hardcopy or electronically). The paper(s) should include open-ended constructed-response essays and long-answer questions that require considerable student writing (NOT selected-response multiple choice and true/false questions).

Scale: 5-pt rating scale
0 = *not at all*, 1 = *rarely*, 2 = *sometimes*, 3 = *often*, 4 = *very often*

How often do you typically experience the following emotions when *conducting research, creative, or scholarly activities/ teaching/ grading student papers*?
Rated for: Enjoyment, Anger, Boredom, Anxiety, Frustration, Pride

Codebook Study 2b (constructed in parallel in English and German)

English	German
With this survey, we are studying faculty experiences GRADING STUDENT PAPERS. What do we mean by "student papers"? This may include homework assignments, term papers, or exams involving… - a judgement (letter grade or pass/fail) - open-ended constructed-response essays and long-answer questions which require considerable student writing - NOT selected-response multiple choice and true/false questions We will refer to such assignments or exams as STUDENT PAPERS (regardless if they are submitted as hardcopy or electronic). For the remainder of this survey, we would ask you to think of one particular student paper as defined above.	In dieser Studie untersuchen wir die Erfahrungen von Hochschullehrenden beim KORRIGIEREN VON SCHRIFTLICHEN ARBEITEN (unabhängig von der Abgabe in gedruckter oder elektronischer Form). Was ist mit „Arbeit/en" gemeint? Es handelt sich um Arbeitsaufträge, Hausarbeiten oder Klausuren mit - Bewertung (Note oder Bestanden/Nicht bestanden) - offener Fragestellung, die längere schriftlich ausgearbeitete Antworten der Studierenden erfordern - aber NICHT geschlossene Fragenformate wie Multiple oder Single Choice Aufgaben. Für den Rest dieses Fragebogens möchten wir Sie bitten, an eine bestimmte Arbeit der oben beschriebenen Art zu denken.

Emotions – Single Items

Scale: 5-pt rating scale:

0 = g*ar nicht (not at all)*, 1 = *selten (rarely)*, 2 = *manchmal (sometimes)*, 3 = *oft (often)*, 4 = *sehr oft (very often)*

English	German
Please indicate how often you typically experience the following emotions **when grading the student paper you specified above**.	Bitte geben Sie an, wie oft Sie normalerweise die folgenden Emotionen erleben, **wenn Sie die von Ihnen benannte Arbeit korrigieren.**
Enjoyment	Freude
Anger	Ärger
Relief	Erleichterung
Guilt	Schuld
Boredom	Langweile
Anxiety	Angst
Frustration	Frustration
Hope	Hoffnung
Pity	Mitleid
Pride	Stolz
Shame	Scham

Emotions – Multi-Item Scales

Scale: 5-pt rating scale

0 = *gar nicht (not at all)*, 1 = s*elten (rarely)*, 2 = *manchmal (sometimes)*, 3 = o*ft (often)*, 4 = *sehr oft (very often)*

English	German
Please indicate how often you typically feel as described below when grading the student paper you specified above.	Bitte geben Sie an, wie oft Sie sich normalerweise so wie unten beschrieben fühlen, wenn Sie die von Ihnen benannte Arbeit korrigieren.
Enjoyment	
For this student paper grading is an enjoyable challenge for me.	Bei dieser Arbeit ist das Korrigieren für mich eine Herausforderung, die mir Spaß macht.
For this student paper I gladly do my grading.	Bei dieser Arbeit bereitet mir das Korrigieren Freude.
For this student paper grading is no fun for me.	Bei dieser Arbeit ist das Korrigieren für mich echt kein Spaß.

Pride	
For this student paper I take pride in being able to fulfill my grading standards.	Bei dieser Arbeit bin ich stolz darauf, dass ich meine Korrekturansprüche erfüllen kann.
For this student paper I am proud of how well I handle the grading.	Bei dieser Arbeit bin ich stolz darauf, wie gut ich mit dem Korrigieren zurecht zu komme.
Boredom	
For this student paper I get so bored I have problems concentrating.	Bei dieser Arbeit werde ich so gelangweilt, dass ich Probleme habe mich zu konzentrieren.
For this student paper I'm so bored that I get tired.	Bei dieser Arbeit werde ich vor Langeweile ganz müde.
For this student paper grading is so boring that my mind begins to wander.	Bei dieser Arbeit ist das Korrigieren so langweilig, dass ich gedanklich abschweife.
Anxiety	
For this student paper I get a queasy feeling in my stomach when I have to grade.	Bei dieser Arbeit bekomme ich vor Nervosität ein flaues Gefühl im Magen, wenn ich korrigieren muss.
For this student paper I worry whether I'm able to cope with grading.	Bei dieser Arbeit mache ich mir Sorgen, ob ich die Korrektur bewältigen kann.
For this student paper I worry that my grades will be challenged.	Bei dieser Arbeit mache ich mir Sorgen, dass meine Noten angefochten werden.
For this student paper I am anxious about making mistakes.	Bei dieser Arbeit habe ich Angst, Fehler zu machen.
Anger	
For this student paper I get so angry I feel like throwing the papers in the trash.	Bei dieser Arbeit ärgere ich mich beim Korrigieren so sehr, dass ich die Arbeiten am liebsten in die Ecke werfen möchte.
For this student paper I'm so angry that I get restless.	Bei dieser Arbeit werde ich vor Ärger ganz unruhig.
For this student paper I wish I could tell off some students while grading their papers.	Bei dieser Arbeit würde ich manchen Studierenden beim Korrigieren ihrer Arbeit am liebsten ordentlich die Meinung sagen.
Frustration	
For this student paper grading frustrates me because it takes so much time.	Bei dieser Arbeit frustriert mich das Korrigieren, weil es so viel Zeit in Anspruch nimmt.
For this student paper just seeing the stack of papers frustrates me.	Bei dieser Arbeit frustriert mich schon der Anblick des Stapels an Arbeiten.

Control

Scale: 1 = *ja (yes)*, 0 = *nein (no)*

Sum scores were calculated to estimate the extent of control over the grading process (min = 0, max = 5) and the paper's content (min = 0, max = 3).

Control over the grading process	
For this student paper, do you get to grade within a timeframe you determine?	Dürfen Sie bei dieser Arbeit innerhalb eines von Ihnen bestimmten Zeitrahmens korrigieren?
For this student paper, do you get to decide whether or not to use a grading rubric?	Dürfen Sie bei dieser Arbeit selbst entscheiden, ob Sie ein Bewertungsschema verwenden oder nicht?
For this student paper, do you get to decide the type of questions (e.g., essay, long answer, multiple choice, true/false, etc.)?	Dürfen Sie bei dieser Arbeit über das Prüfungsformat entscheiden (z.B. Hausarbeit, Multiple Choice Klausur, …)?
For this student paper, do you get to determine the format (e.g., duration of exam, length of paper to be written, formatting, etc.)?	Dürfen Sie bei dieser Arbeit die formalen Anforderungen festlegen (z.B. Dauer der Klausur, Zeichenzahl der Arbeit, Formatierung, etc.)?
For this student paper, do you get to choose the type of grade (letter grade, percentage, pass/fail, etc.)?	Dürfen Sie bei dieser Arbeit die Art der Bewertung aussuchen (Benotung vs. Bestanden/Nicht-Bestanden)?
Control over the paper's content	
Is this student paper given in a course you teach?	Halten Sie die Veranstaltung, deren Stoff in dieser Arbeit geprüft wird?
For this student paper, do you get to pick the subject matter?	Dürfen Sie bei dieser Arbeit das Thema aussuchen?
For this student paper, do you get to formulate the questions yourself?	Dürfen Sie bei dieser Arbeit die Fragen selbst formulieren?

Competence

Scale: 5-pt rating scale

0 = *stimme überhaupt nicht zu (strongly disagree)*, 1 = *stimme nicht zu (disagree)*, 2 = w*eder noch (neutral)*, 3 = s*timme zu (Agree)*, 4 = s*timme voll und ganz zu (strongly agree)*

English	German
Please indicate your agreement with the following statements with respect to grading the student paper you specified above.	Bitte geben Sie an, wie sehr Sie den folgenden Aussagen bezüglich der Korrektur der von Ihnen benannten Arbeit zustimmen.
Content Competence / Content Knowledge	
For this student paper I have good knowledge of the content area.	Bei dieser Arbeit habe ich gute Kenntnisse des Themenbereichs.
For this student paper I am familiar with the topic.	Bei dieser Arbeit bin ich mit dem Thema vertraut.
Diagnostic Competence	
For this student paper I can ensure that my grades are fair.	Bei dieser Arbeit kann ich sicherstellen, dass meine Noten fair sind.
For this student paper I can accurately judge the quality of each paper.	Bei dieser Arbeit kann ich die Qualität jeder Arbeit genau beurteilen.
For this student paper I can appropriately handle factors that may bias my grading results.	Bei dieser Arbeit kann ich Einflussfaktoren, die meine Korrekturergebnisse verfälschen könnten, angemessen handhaben.
For this student paper I can make transparent how my grades are given.	Bei dieser Arbeit kann ich transparent darstellen, wie meine Noten zustande kommen.

Value

Scale: 5-pt rating scale

0 = *stimme überhaupt nicht zu (strongly disagree)*, 1 = *Stimme nicht zu (disagree)*, 2 = w*eder noch (neutral)*, 3 = *stimme zu (agree)*, 4 = *stimme voll und ganz zu (strongly agree)*

English	German
Please indicate your agreement with the following statements with respect to grading the student paper you specified above.	Bitte geben Sie an, wie sehr Sie den folgenden Aussagen bezüglich der Korrektur der von Ihnen benannten Arbeit zustimmen.
Diagnostic Value - Importance of „Grading Standards“	
For this student paper it is important to me to ensure that my grades are fair.	Bei dieser Arbeit ist es mir wichtig sicherzustellen, dass meine Noten fair sind.
For this student paper it is important to me that I accurately judge the quality of each paper.	Bei dieser Arbeit ist es mir wichtig, dass ich die Qualität jeder Arbeit genau beurteile.
For this student paper it is important to me to appropriately handle factors that may bias my grading results.	Bei dieser Arbeit ist es mir wichtig, dass ich Einflussfaktoren, die meine Korrekturergebnisse verfälschen könnten, angemessen handhabe.
For this student paper it is important to me that I make transparent how my grades are given.	Bei dieser Arbeit ist es mir wichtig, dass ich transparent darstelle, wie meine Noten zustande kommen.
Cost	
Grading this student paper keeps me away from more meaningful tasks.	Das Korrigieren dieser Arbeit hält mich von sinnvolleren Aufgaben ab.
Grading this student paper is a thankless task.	Das Korrigieren dieser Arbeit ist eine undankbare Aufgabe.
Utility Value	
Grading this student paper is important to me because it tells me if certain students are struggling in class.	Das Korrigieren dieser Arbeit ist mir wichtig, weil ich so erfahre, ob sich bestimmte Studierende im Kurs schwertun.
Grading this student paper is important to me because the results tell me how to adjust my teaching.	Das Korrigieren dieser Arbeit ist mir wichtig, weil mir die Ergebnisse darüber Aufschluss geben, wie ich meine Lehre anpassen kann.
Social Value	
Grading this student paper is important to me because the judgements impact my relationships with the students.	Das Korrigieren dieser Arbeit ist mir wichtig, weil die Urteile meine Beziehung zu den Studierenden beeinflussen.

Appendix C

Codebook Study 3

(Inofficial English translation in parantheses)

Basic Questionnaire (BQ)

***Arbeitsbedingungen** (Working Conditions; BQ)*

Scale: 8-point rating scale

1 = *trifft überhaupt nicht zu* *(strongly disagree)*, 8 = *trifft voll und ganz zu* *(strongly agree)*

Üblicherweise in meiner Lehre ... *(Typically, in my teaching ...)*
Autonomie *(Autonomy)*
... bin ich frei, die Dinge auf meine Weise zu tun. *(... I am free to do things my way.)*
... drücken meine Entscheidungen mein „wahres Selbst“ aus. *(... my decisions express my "true self".)*
... tue ich das, was mich persönlich interessiert. *(... I do what interests me personally.)*
... habe ich viel Druck, auf den ich gut verzichten könnte. *(... I have a lot of pressure that I could easily do without.) *recoded*
... sagen mir Andere, was ich zu tun habe. *(... others tell me what I have to do.) *recoded*
... muss ich einiges gegen meinen Willen tun. *(... I have to do some things against my will.) *recoded*
Kompetenzerleben *(Competence)*
... schließe ich schwierige Aufgaben und Projekte erfolgreich ab. *(... I successfully complete difficult tasks and projects.)*
... stelle ich mich schweren Herausforderungen und bewältige diese. *(... I face and overcome difficult challenges.)*
... meistere ich auch schwierige Dinge gut. *(... I also master difficult things well.)*
... erlebe ich Misserfolge oder kann einige Dinge nicht gut erledigen. *(... I experience failures or cannot handle some things well.) *recoded*
... stelle ich mich dumm an und fühle mich deshalb inkompetent. *(... I act stupid and therefore feel incompetent.) *recoded*
... mühe ich mich mit etwas ab, bei dem ich eigentlich gut sein sollte. *(... I struggle with something I'm supposed to be good at.) *recoded*
Soziale Eingebundenheit *(Relatedness)*
... fühle ich mich mit Kolleginnen und Kollegen, die sich um mich kümmern und um die ich mich kümmere, verbunden. *(... I feel connected to colleagues who care about me and whom I care about.)*
... fühle ich mich Kolleginnen und Kollegen, die mir wichtig sind, nahe und verbunden. *(... I feel close and connected to colleagues who are important to me.)*

... fühle ich mich mit Kolleginnen und Kollegen, mit denen ich Zeit verbringe, sehr vertraut. *(... I feel very familiar with colleagues with whom I spend time.)*

... fühle ich mich einsam. *(... I feel lonely.) *recoded*

... fühle ich mich von einer oder mehreren mir wichtigen Kolleginnen und Kollegen nicht beachtet. *(... I feel ignored by one or more colleagues who are important to me.) *recoded*

... habe ich Unstimmigkeiten oder Konflikte mit Kolleginnen und Kollegen, mit denen ich sonst gut auskomme. *(... I have disagreements or conflicts with colleagues with whom I otherwise get along well.) *recoded*

***Selbstwirksamkeitsüberzeugungen** (Self-Efficacy; BQ)*

Scale: 8-point rating scale

1 = *überhaupt nicht (not at all)*, 8 = *sehr gut (very good)*

Wie gut gelingt es Ihnen üblicherweise in Ihrer Lehre ... *(Typically, in your teaching, how well do you accomplish to...)*

... für besonders leistungsfähige Studierende herausfordernde Aufgaben anzubieten? *(... provide challenging assignments in your teaching for particularly high-achieving students?)*

... abwechslungsreiche Lehrmethoden einzusetzen? *(... use varied teaching methods?)*

... alternative Erklärungen oder Beispiele zu präsentieren, wenn die Studierenden etwas nicht sofort verstehen? *(... present alternative explanations or examples when students do not immediately understand something?)*

... den Studierenden zu vermitteln, wo der Sinn und der Nutzen der Veranstaltungsinhalte liegen? *(... convey to the students what the meaning and benefits of the course content are?)*

... uninteressierte Studierende zu motivieren? *(... motivate uninterested students?)*

... auch die gelangweilten Studierenden zu erreichen? *(... reach also the bored students?)*

... dafür zu sorgen, dass die Studierenden sich angemessen verhalten? *(... ensure that students behave appropriately?)*

... Unruhe und Störungen zu unterbinden? *(... prevent agitation and disturbances?)*

... zu vermeiden, dass Studierende durch ihr Verhalten den Sitzungsablauf stören? *(... prevent students from disrupting the course of the session through their behavior?)*

Vergleich Allgemeiner Erfahrungen in der Online Lehre und Lehre vor der Pandemie

(Comparison of General Emergency Online Teaching Experiences and Teaching Experiences Before the Pandemic; BQ)

Scale: 9-point semantic differential

−4 = w*eniger (less)*, 0 = g*leich (equal)*, +4 = m*ehr (more)*

Im Vergleich zu meinen bisherigen typischen Erfahrungen in der nicht-digitalen Lehre habe ich das Gefühl, ... *(Compared to my typical experiences in non-digital teaching so far, I feel like ...)*
Autonomie ***(Autonomy)***
... selbst bestimmen zu können wie ich meine Lehre gestalte. *(... I can determine how I design my teaching.)*
... autonom handeln zu können. *(... I can act autonomously.)*
Kompetenzerleben ***(Competence)***
... kompetent zu sein. *(... I am competent.)*
... meine Lehre gut und kompetent bewältigen zu können. *(... I can handle my teaching well and competently.)*
Relatedness ***(Soziale Eingebundenheit)***
... sozial eingebunden zu sein. *(... I am socially connected.)*
... meinen Studierenden nahe und verbunden zu sein. *(... I am close and connected to my students.)*
Zufriedenheit ***(Satisfaction)***
... zufrieden mit meiner Lehre zu sein. *(... I am satisfied with my teaching.)*

Stress (BQ)

Scale: 5-pt rating scale

0 = *nie (never)*, 1 = s*elten (rarely)*, 2 = *manchmal (sometimes)*, 3 = *häufig (often)*,

4 = s*ehr häufig (very often)*

Wie oft haben Sie seit Beginn der Corona-Krise Folgendes erlebt? *(How often have you experienced the followings since the Corona crisis has started?)*
Zeiten, in denen ich dringend benötigte Erholung aufschieben muss. *(Times when I had to postpone much needed recreation.)*
Ich habe zu wenig Zeit, um meine täglichen Aufgaben zu erfüllen. *(I have too little time to perform my daily tasks.)*
Zeiten, in denen sich die Termine so häufen, dass Sie kaum zu bewältigen sind. *(Times when appointments pile up in such a way that you can hardly cope with them.)*
Zeiten, in denen ich unter Termindruck/Zeitnot arbeiten muss. *(Times when I have to work under deadline pressure/time constraints.)*
Zeiten, in denen ich zu viele Verpflichtungen zu erfüllen habe. *(Times when I have too many commitments to fulfill.)*
Zeiten, in denen mir die Arbeit über den Kopf wächst. *(Zeiten, in denen mir die Arbeit über den Kopf wächst.)*
Ich habe zu viele Aufgaben zu erledigen. *(I have too many tasks to do.)*
Erfahrung, dass alles zu viel ist, was ich zu tun habe. *(Experience that everything is too much for me to do.)*

Technische Probleme (Technical Problems; BQ)

Scale: 8-point rating scale

1 = *stimmt gar nicht (strongly disagree)* to 8 = *stimmt genau (strongly agree)*

Ich habe das Gefühl, die Technik steht mir bei meiner Lehre im Wege. *(I feel like technology is getting in the way of my teaching.)*
Es gibt ständig technische Probleme. *(There are technical problems all the time.)*
Die Funktionalität des von mir verwendeten Online-Meeting Tools ist schlecht. *(The functionality of the online meeting tool I use is poor.)*
Der Support zur Lösung technischer Probleme in meiner digitalen Lehre ist mangelhaft. *(The support for solving technical problems in my digital teaching is poor.)*

Session-Specific Questionnaire (SQ)

***Emotionen innerhalb einer Sitzung** (Emotions Within a Teaching Session; SQ)*

Scale: 8-point rating scale;

1 = *stimme überhaupt nicht zu (strongly disagree)*, 8 = *stimme voll und ganz zu (strongly agree)*

In der heutigen Sitzung empfand ich ... *(In today's session, I experienced ...)*
... Freude. *(... enjoyment.)*
... Stolz. *(... pride.)*
... Langeweile. *(... boredom.)*
... Ärger. *(... anger.)*
... Angst. *(... anxiety.)*
... Scham. *(... shame.)*

Bedürfnisbefriedigung und Zufriedenheit mit der Lehre innerhalb einer Sitzung

(Basic Need Satisfaction and Teaching Satisfaction Within a Teaching Session; SQ)

Scale: 8-point rating scale

1 = *trifft überhaupt nicht zu (strongly disagree)*, 8 = *trifft voll und ganz zu (strongly agree)*

In der heutigen Sitzung hatte ich das Gefühl... *(In today's session, I felt like ...)*
Autonomie ***(Autonomy)***
... selbst bestimmen zu können wie ich meine Lehre gestalte. *(... I could determine for myself how I design my teaching.)*
... autonom handeln zu können. *(... I could act autonomously.)*
Kompetenzerleben ***(Competence)***
... kompetent zu sein. *(... I was competent.)*
... meine Lehre gut und kompetent bewältigen zu können. *(... I could handle my teaching well and competently.)*
Soziale Eingebundenheit ***(Relatedness)***
... sozial eingebunden zu sein. *(... I was socially connected.)*
... meinen Studierenden nahe und verbunden zu sein. *(... I was close and connected to my students.)*
Zufriedenheit mit der Lehre (***Teaching Satisfaction***)
Insgesamt bin ich mit der heutigen Sitzung zufrieden. *(Overall, I am satisfied with today's session.)*