3-2022

Grade Determination: An Exploration of High School Teacher Cognitive Processes

Bonnie Robbins

GRADE DETERMINATION: AN EXPLORATION OF HIGH SCHOOL TEACHER COGNITIVE PROCESSES

by

Bonnie Robbins

FACULTY RESEARCH COMMITTEE:

Chair: Dane Joseph, PhD

Member: Karen Buchanan, EdD

Member: Gary Sehorn, EdD

Presented to the Faculty of the

Doctor of Educational Leadership Department

in partial fulfillment for the degree of

Doctor of Education

GEORGE FOX UNIVERSITY

March 9, 2022

GEORGE FOX UNIVERSITY

COLLEGE OF EDUCATION | EDD

GRADE DETERMINATION: AN EXPLORATION OF HIGH SCHOOL TEACHER COGNITIVE PROCESSES, a Doctoral research project prepared by BONNIE ROBBINS in partial fulfillment of the requirements for the Doctor of Education degree in Educational Leadership.

This dissertation has been approved and accepted by:

		Committee Chair
April 1, 2022	Dane Joseph, Ph.D.	Associate Professor of Education
	Karen Buchanan	
April 1, 2022	Karen Buchanan, Ed.D.	Professor of Education
	Gary Sehorn	
April 1, 2022	Gary Sehorn, Ed.D.	Associate Professor of Education

ABSTRACT

This think-aloud study closely examined the cognitive processes of four high school teachers as they made grade determinations for two hypothetical students. This study serves to give insight into the often veiled process of how secondary teachers make final decisions about students' summative grades and what cognitive biases and heuristics they rely on to make such grade decisions and, if relevant, grade changes. Each teacher participated in a two-part interview: in part one, teachers were presented with two student vignettes detailing academic, extracurricular, and background information and were directed to think aloud their process as they determined each student's ultimate course grade; in part two, teachers were asked reflective questions. Both vignettes presented, despite being different in many facets, were crafted to prompt teachers to engage in thinking surrounding whether or not they would bump a grade that may be considered a cusp grade. Overall, the study found that teachers are generally consistent in their overarching cognitive processing across students, but differ in which heuristics they may commit from student to student. Additionally, while all teachers were shown to have engaged in heuristic thinking and System 1 and System 2 thinking, teachers vary greatly from one another in the complexity of their cognitive processes and the extent to which they rely on heuristics to determine grades. Furthermore, as they progress in the profession, teachers seem to become more flexible in their grade determination cognitive processes and become more candid about the emotional tolls and inequities of current grading practices.

Keywords: high school teachers, grade determination, cognitive process, heuristics, think-aloud

ACKNOWLEDGEMENTS

I would like to first thank Dr. Gary Sehorn for prompting me to embark on this doctoral journey; without his faith and guidance, I would not be where I am today. I would also like to thank Dr. Sehorn and Dr. Karen Buchanan for their invaluable guidance and feedback on my dissertation. Additionally, I would like to thank Dr. Dane Joseph for his unwavering support, inspiration, and brilliance throughout my dissertation journey. I could not have dreamt of a better chair. Most of all, however, I would like to thank my family and, more specifically, my husband for their patience, love, and support throughout my doctoral work and dissertation process. It is not lost on me that this is the work of a collective.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Grade Determination: An Exploration of High School Teacher Cognitive Processes

Chapter 1: Introduction

Background and Problem of Practice

Teachers are tasked with providing students in the American public education system with letter grades ideally representing students' knowledge and skills in a given content area. However, with emphasis often placed on assignment completion, participation, attendance, and the like, what these grades often represent may be biased, and subjective. Therefore, students may be passing courses without actually being proficient in the respective knowledge and skill domains contained within the courses. Likewise, students who are failing courses who are indeed proficient in content are not moving on to the next course or, worse, not graduating--despite having the prerequisite skills. At times this is dependent on the final, gate-keeping decision of the teacher as grade practitioner.

In many schools, teachers are given very little, if any, direct instruction themselves on how to make these important grading decisions. They are, under some administrations, given overall pedagogical guidance focusing on vision or general grading methodology (Richards, 2014), while under others they are given no guidance at all as to the *how* of grading or ultimate grade decision-making (Stitt & Pula, 2014). While such faith in the teacher, as expert, may be well-intentioned, the lack of transparency and training can result in a system which relies profoundly upon what each individual teacher deems essential, and is, thus, influenced by each teacher's own proclivities and biases.

Brookhart et al. (2016) surveyed over one hundred years of research on grading, using ERIC and ProQuest databases, and concluded that grades are typically multidimensional and represent a variety of cognitive and non-cognitive factors that teachers value. For example,

"teachers recognize the important role of effort in achievement and motivation (Aronson, 2008; Cizek et al., 1995; Cross & Frary, 1999; Duncan & Noonan, 2007; Guskey, 2002, 2009a; Imperial, 2011; S. Kelly, 2008; Liu, 2008b; McMillan, 2001; McMillan et al., 2002; McMillan & Lawson, 2001; McMillan & Nash, 2000; Randall & Engelhard, 2009, 2010; Russell & Austin, 2010; Sun & Cheng, 2013; Svennberg et al., 2014; Troug & Friedman, 1996; Yesbeck, 2011). They differentiate academic enablers (McMillan, 2001) like effort, ability, improvement, work habits, attention, and participation" (p. 834).

It is such multidimensionality that Brookhart (2016) argues makes grades "better predictors of future success in school than tested achievement (Atkinson & Geiser, 2009; Barrington & Hendricks, 1989; Bowers, 2014; Cairns et al., 1989; Cliffordson, 2008; Ekstrom et al., 1986; Ensminger & Slusarcick, 1992; Finn, 1989; Fitzsimmons et al., 1969; Hargis, 1990; Lloyd, 1974, 1978; Morris et al., 1991; Rumberger, 1987; Troob, 1985; Voss et al., 1966), especially given known limitations of achievement testing" (p. 834). Thorsen and Cliffordson (2012) reinforced this conclusion in their study of over 300,000 students. The study retrieved data on three different birth cohorts in Sweden: the cohorts being born in 1987, 1988, and 1989. The data, which was collected from the Gothenburg Educational Longitudinal Database, found that grades have better predictive validity than test scores alone and are, thus, perhaps the most valid instruments for predicting educational success. More specifically:

> The results suggest that teachers take account of aspects other than knowledge and skills when assigning grades (e.g., Klapp Lekholm & Cliffordson, 2009; Randall & Engelhard, 2010). Such aspects could either exert a direct influence on the assignment of grades in that teachers award higher grades to students who display a higher level of effort or interest (Brookhart, 1991, 1993; Pilcher, 1994), or an indirect influence in that a higher

> level of effort could result in the acquisition of additional knowledge and skills. Irrespective of whether such aspects influence grades directly or indirectly, they nevertheless prove to be of importance for future achievement (p. 167).

However, despite the multidimensional nature of grading superseding strict standardized testing as a measure, researchers have found grades to not always be assigned in fair and equitable ways. Focusing primarily on students in Germany, Westphal et al. (2016, 2020) published two papers in which they expanded upon the relationship between socioeconomic classroom composition and grade earned. In their 2020 study of 5,919 seventh grade students' enrolled in 457 classes, they examined final report card grades in conjunction with student characteristics, and noted the relationship between "individual characteristics of students, such as extraversion, academic self-efficacy, and conscientiousness" and student achievement level--finding that such characteristics "may guide teachers' evaluations of student achievement, resulting in more appropriate judgements and a stronger alignment of assigned grades with students' actual achievement level" (p. 1). A number of researchers have also found race and ethnicity (van Ewijk, 2011; Irizarry, 2015), language-learner status (Blanchard & Muller, 2015), and parent-teacher relationships (Hughes et al., 2005) to influence teacher perception and grade outcome.

A number of researchers have delved more deeply into what factors specifically influence teachers in their grading practices and assignments, and have sought to better understand teacher perception (Guskey & Link, 2019; Riley & Ungerleider, 2019; Brookhart et al., 2016). Guskey and Link, in a study of 943 K-12 teachers, found differences in what teachers measured and valued as the grade level increased, but also found evident differences at the same grade level "in teachers' consideration of both cognitive and non-cognitive factors of students' performance"

(para 1). However, there were limitations to the reach of Guskey and Link's conclusions given the methodology: they gathered their data by collecting questionnaires from teachers with questions that were more limited in scope, such as yes/no questions, multiple choice, and "check all that apply" questions. The questionnaire also focused primarily on cognitive achievement factors, and did not delve into the non-cognitive nor into the rationale behind teachers' decision-making processes.

Teachers have themselves been found to be weary of the grading process, having "reservations about the grading process as well as the impact of their beliefs and values upon their decision making" (Riley & Ungerleider, 2019, p. 220), but there remain gaps in research regarding such teacher sentiments and solutions. Riley and Ungerleider interviewed 21 teachers in grades 6-12, asking them to assess 24 fictional students as remedial, average, or advanced based upon their fourth through seventh grade report cards. The hypothetical exercise revealed biases and such teacher reservations, but was limited in sample size, limited in that it sampled only white teachers, and limited in that it was focusing on fictionalized students.

While much research has been done in the areas of teacher perception and grading, there is still much work to do to better understand some of the gaps, biases, and processes evidenced. Guskey and Link (2019), for example, contend that "studying the reasons why teachers prioritise particular types of assessment evidence at different grade levels" is a needed next step, and that "focus groups and in depth interviews of verbal protocols would be particularly useful in providing insights into different weighting strategies and to the reasons behind teachers' grading practices" (p. 316). Riley and Ungerleider (2019) reiterate this need and, more specifically, assert that research could "benefit from attention on the moral components of the process and the responsibility inherent in making unbiased judgments" (p.225); while Jewell and McPherson

(2012) recommend a similar approach in order to gain better insight into what motivates particular teachers to be more inclined to assign higher grades overall. Blanchard and Muller (2015) encapsulate much of the need whilst they contend: “we call for further research to deepen our understanding of the causes and consequences of differences in teachers’ perceptions” (p. 273).

Purpose Statement

There is much in the literature focusing on grade composition, the multidimensionality of grades, and different grading paradigms. Underpinning teacher practices and what may drive perceptions and actions is certainly heavily influenced by the multifaceted and complex methodological backdrop upon which they, as practitioners, are thrown--with some teachers subscribing to more standards-based methodologies and others subscribing to multidimensional ones with behavioral components, with others still subscribing to the more nouveau “going gradeless” ideology.

The purpose of this study was to address a need identified by Guskey and Link (2019) and investigate what factors teachers weigh most heavily in processing their final course grades, including what, if anything, prompts an active, teacher-manipulated change of a student’s final grade in a course. Furthermore, the cognitive processes and biases that teachers evidence in such weighing was analyzed utilizing psychological frameworks including System 1 versus System 2 thinking.

Teacher perception regarding what factors to weigh and how to judge student competency was also explored. More specifically, research regarding secondary teacher perception was examined with a focus on what teachers at the secondary level value when

determining grades. Furthermore, the study explored how teachers view and perceive the grading process overall, (including concerns, pressures, etc.).

Research Questions

This study investigated why secondary teachers weigh certain factors most heavily in their grading process, and what, if anything, motivates a grade change at the end of a term. More specifically, the following essential questions were of focus:

- Question 1: How do secondary teachers make final decisions about students' summative grades?
- Question 2: What cognitive biases do they rely on to make grade decisions and, if relevant, grade changes?

The research questions explore what psychological processes and biases teachers employ when making final marking decisions. Heuristics were considered in the analysis to frame the results of the study. The specific heuristics of focus in this study were algorithmic thinking, axiomatic rationality, halo effect, fundamental attribution error, intuitive judgement, availability heuristic, confirmatory bias, affect heuristic, and recency heuristic. These were chosen as they encompass the complex nature of the teacher as judge: teachers at the secondary level often get to know their students well before their marking decision at the end of the term--making them more prone to committing heuristic thinking related to emotional responses, character, etc. On a grander scale, whether teachers subscribe to more axiomatic, analytical processes versus more ecological, biased ones was explored. This entailed assessing cognitive processing for more System 1 versus System 2 thinking operations.

The study sought to also understand what influencers outside of the teacher-student realms exist and most readily impact the teacher's grading process. Such themes arose and were

coded in the interviews, and included pressure, college plans, English language learner status, etc. These particular influencers may readily affect decision-making processes and perhaps bias the instructor to be more prone to inflate grades. Teachers who are more prone to committing the affect heuristic, for example, might take into account future college plans of a student and not want to assign a failing or lower mark if the consequence of doing so is that the student will not be accepted.

Significance Statement

Understanding the relationship amongst teacher perception, drivers of action, and ultimate student grade outcomes is paramount to not only better understand a typically veiled process, but to move pragmatically and ethically forward. More specifically, incongruities amongst teacher-perception may illuminate ethical failures of our current system; what factors are weighed most heavily may unearth biases; and drivers of action may showcase a lack of fairness as well as systemic pressures. Such inaccuracies may validate the need for a clear discussion amongst stakeholders surrounding what, for example, a grade in a course should and should not take into account (such as attendance, completion, participation, work ethic, skill, knowledge, growth). This may necessitate the consideration of culturally responsive practices when weighing certain non-cognitive factors. And, ultimately, under what conditions teachers alter a student's grade at the end of a term is highlighted. This may lead into an exploration of whether teachers are consistently, fairly, and ethically applying criteria in order to make ultimate grading decisions. In turn, administrators and stakeholders may then examine whether they are placing too much pressure on teachers to pass all students.

Overarchingly, the aim of the aforementioned is to ensure that the grades assigned to students are not just that--assigned-- but are, rather, accurate marks of students' abilities so that

they will be aptly prepared for subsequent courses and life applications. Grades are of weight in that they are not only indicators of potential future success, but they are also used by various stakeholders as measures for entry. College admittance, scholarships, job applications, and many more use grades as a criterion. Thus, exploring and further understanding what teachers weigh when determining such a mark is paramount.

Ultimately, the study provides insight into what drives teacher perception and action when it comes to student grades and provides insight into what training and discussions are necessary. The study fills a gap in the literature regarding understanding *why* teachers do what they do when it comes to grades: their motivations and catalysts.

We can expect that, when given autonomy and without detailed training, teachers will evidence a variety of cognitive processes when assigning grades: some will rely on more axiomatic processes and adhere strictly to prescripted measures while others will evidence more System 1 thinking and commit the fundamental attribution error or other such biased methods. Such variety may be, in part, due to variables such as years of service in education, gender, subject matter taught, etc., but given the limited scope of the study, this was not explored in depth. However, the knowledge gleaned from the variety of processes can inform professional development and will open the door to discussions that will aid teachers in the metacognitive examination and perhaps alteration of their current grading processes and practices. The results may also shed light upon systemic influences and provide administrators insight into the impact of their verbiage and rhetoric surrounding grades and student pass rates. This may allow administrators to make more informed future judgements about how to discuss such topics with their staff.

Definition of Terms

The Affect Heuristic

Slovic and his colleagues, in 2002, posited that "a basic affective reaction can be used as the heuristic attribute for a wide variety of more complex evaluations, such as the cost-benefit ratio of technologies, the safe concentration of chemicals, and even the predicted economic performance of industries" (Slovic, 2002 as cited in Kahneman, 2003, p.710).

Algorithmic Thinking

Algorithmic thinking involves a reliance on and application of algorithms to determine an outcome (Ciobanu & Popovici-Bujor, 2020).

Assessment Literacy

Knowledge and training in the art and effectiveness of assessment practices.

Availability Heuristic

Judges are prone to weigh most heavily the information that first comes to mind or that seems most pertinent. Showing deference for that which comes to mind most easily is defined as the availability heuristic (Klein, 2005).

Axiomatic Rationality

Axiomatic rationality too falls under System 2 thinking in that it is analytical and reasoning based. It is defined as "conformity to abstract axioms" and is, by some, seen as the "norm for how human beings should reason, arguing in addition that violations would lead to real costs" (Gigerenzer, 2021, p.3547).

Cognitive Biases

For the purposes of this study, cognitive biases will be any process by which a teacher uses information other than performance data-driven information to make decisions.

Confirmatory Bias

Confirmatory bias is "the tendency to look for, notice, and remember information that fits with our pre-existing expectations. Similarly, information that contradicts those expectations may be ignored or dismissed as unimportant" (Klein, 2005, p.782).

English Language Learner (ELL)

Student whose native language is not English, and is currently classified as needing English language supports via educational services, such as teacher accommodations.

Fundamental Attribution Error, Correspondence Bias

The fundamental attribution error (FAE) is defined as "the tendency for people to make dispositional, rather than situational, attributions for an actor's behavior" (Jones, 1979 as cited in Gilbert et al., 1988). It is also termed correspondence bias--as it is not always committed in error and is not as pervasive as it was once thought to be (Miller, 1984; Harvey, Town & Yarkin, 1981 in Howell & Shepherd, 2011).

Grade Inflation

Over time, the easing of standards or expectations such that student grades and GPAs are higher. More specifically, this may include individual instructors raising individual students' grades at the end of a term (i.e. rounding up or raising grades).

Halo Effect

The halo effect is a form of cognitive bias in which the judge is influenced by their existing impression of a person's character when making decisions or judgements (Thorndike, 1920; Beehr et al., 2001; Lambart et al., 1997; Becker & Cardy, 1986; Feeley, 2002 as cited in Gweon et al., 2017).

Heuristics

"Heuristics are efficient cognitive processes, conscious or unconscious, that ignore part of the information" (Gigerenzer & Gaissmaier, 2011)

Intuitive Judgement, Expertise

Simon defines intuition as "the recognition of patterns stored in memory" (Simon, 1973 as cited in Kahneman & Klein, 2009, p.516). Such intuitive judgement can be based in skill or expertise, or not.

Recency Heuristic

That which is recent is often most easily recalled and prioritized in decision-making (Lee et al., 1998).

Special Education Student (SPED Student)

Students who are classified as needing additional educational supports and services via a 504 plan or an IEP (Individualized Education Plan). Teacher accommodations and/or modifications may vary for such students.

System 1 Thinking

System one decision-making processes are often quick, associative, and intuitive in nature being at times emotionally driven and habitual (Kahneman, 2003). Heuristics are often entailed within System 1 thinking, as they are often seen as antithetical to logic (Gigerenzer, 2008).

System 2 Thinking

System 2 decision-making is, on the other hand, slower and more deliberate. These processes are seen as more analytical and logical in nature and are often rule or algorithm governed (Kahneman, 2003).

Teacher Autonomy

Work autonomy, according to a classical definition, refers to 'the degree of control or discretion a worker is able to exercise with respect to work methods, work scheduling, and work criteria' (Breaugh, 1985, p. 556; 1999 as cited in Kwok, 2020)

Limitations

Many of the limitations of this study relate to its limited scope. Due to the in-depth nature of the study's methodology, the study has only four participants; this limits the generalizability of the study. All of the participants also work at the same high school site in the same location. The study's findings may not, then, reflect cognitive processes or structural features present in other parts of the country or world, or in other levels of academia. Further studies would need to be undertaken to validate whether the study's findings transfer to elementary educators, middle school educators, and college educators in differing locales. Furthermore, should researchers want to determine which variables, (such as gender, years of teaching, etc.), correlate with System 1 versus System 2 thinking or heuristic surfacings, further studies would be needed.

Another limitation of this study relates to researcher and subject bias. The researcher is employed at the site that is the subject of this study and works with the study's participants. This may influence the outcomes of the study. While the existing researcher-participant relationship may have fostered more open and honest communication regarding a perhaps taboo and obfuscated process, it may have also yielded inaccurate data due to social desirability bias. This, though, was taken into account in the selection of participants, and the researcher was not in the same department, nor does she have any personal or financial relationships with any of the participants.

Delimitations

One potential limitation of the study is the fact that at the secondary level, there are four different grade levels being taught and assessed--each with their own potential outlier considerations. For example, ninth grade teachers may take into consideration that freshmen are new to the high school environment and rigor, and, thus, grade students more leniently in order to allow for acclimatization. Likewise, senior teachers are often influenced by pressures to pass twelfth graders because if they do not the students may not graduate. As such, this study has controlled for these grade-specific influencers by narrowing the study to tenth and eleventh grade teachers.

Other potential limitations of studying teachers' cognitive processes are systemic or school mandates related to grading policies. This would disallow teachers from having the freedom to weigh many factors freely in assigning grades at the end of a term. This has been controlled by selecting the specific school site, as this setting allows for choice and allows teacher grading autonomy.

Conclusion

Grading is a complex and often inconsistent process that is under-studied. Teachers, as practitioners, are called to make holistic judgments in the form of final course grades for their students. Given that teachers are humans judging humans, heuristics readily occur in the grade-determination process. This study closely examines the inner workings of this cognitive process in all of its complexity.

Chapter 2: Review of the Literature

Introduction

This study explores what cognitive processes secondary teachers exhibit when making ultimate letter grade decisions. An area of particularity was the teacher manipulation, or changing of student grades at the end of a term and how teachers make such decisions, including what they weigh as most important. Heuristics and biases, (including recency heuristics, the fundamental attribution error, the halo effect, homo heuristics, etc.), were used as theoretical frameworks for the study and, as such, comprise a majority of the literature review. The literature review has four overarching foci: an introduction and background of heuristics, research related to heuristics in the educational field, analogous studies, and an overview of specific heuristics hypothesized to surface in the participants' responses. While there is much in the literature regarding heuristics more broadly and some pertinent specifically to the education field, most existing literature focuses on instructional pedagogy or single assessment decisions with very little research yet done on the decision-making process teachers utilize to make culminating assessment/marking decisions. Hence, analogous studies are noted and anticipated results are hypothesized.

Heuristics Background

Heuristics is a field of cognitive psychology that has been of study since the 1970s when the term underwent a shift from being associated with computation and artificial intelligence to explaining intelligence and cognitive processes in humans (Gigerenzer & Brighton, 2009). At this time, Daniel Kahneman and Amos Tversky conducted and published a "series of experiments in which people's reasoning was interpreted as exhibiting fallacies. 'Heuristics and biases' became one phrase. It was repeatedly emphasized that heuristics are sometimes good and

sometimes bad, but virtually every experiment was designed to show that people violate a law of logic, probability, or some other standard of rationality" (Gigerenzer & Brighton, 2009, p.109). While the term heuristic is of Greek origin, translating to "serving to find out or discover", the widely accepted cognitive psychology definition is as follows:

> Heuristics are efficient cognitive processes, conscious or unconscious, that ignore part of the information. Because using heuristics saves effort, the classical view has been that heuristic decisions imply greater errors than do "rational" decisions as defined by logic or statistical models. However, for many decisions, the assumptions of rational models are not met, and it is an empirical rather than an a priori issue how well cognitive heuristics function in an uncertain world. (Gigerenzer & Gaissmaier, 2011, p.451)

In its early inception, the field of heuristics as applied to humans focused on fallacious reasoning, or errors in logic, and thus had a negative connotation. Gigerenzer and Brighton, (2009) articulate the widely-held beliefs regarding heuristics by the end of the 20th century: "1. Heuristics are always second-best, 2. We use heuristics only because of our cognitive limitations, and 3. More information, more computation, and more time would always be better" (p.109). However, much of the subsequent work done in the 21st century has disproven these conceptions, instead finding that there is value and merit in the utilization of heuristics (Kahneman, 2003; Gigerenzer, 2008; Gigerenzer & Brighton, 2009; Kahneman & Klein, 2009; Gigerenzer & Gaissmaier, 2011; Gigerenzer, 2021). Despite the trend in validating heuristic decision-making, there remains the recognition that relying on heuristics may "lead to unintended bias in decision-making(Gilovich & Griffin, 2002)"(Crisp, 2010), and although being

generally “quite useful,...sometimes they lead to severe and systematic errors. (Tversky & Kahneman, 1974, p. 1124)” (Kahneman, 2003, p.707).

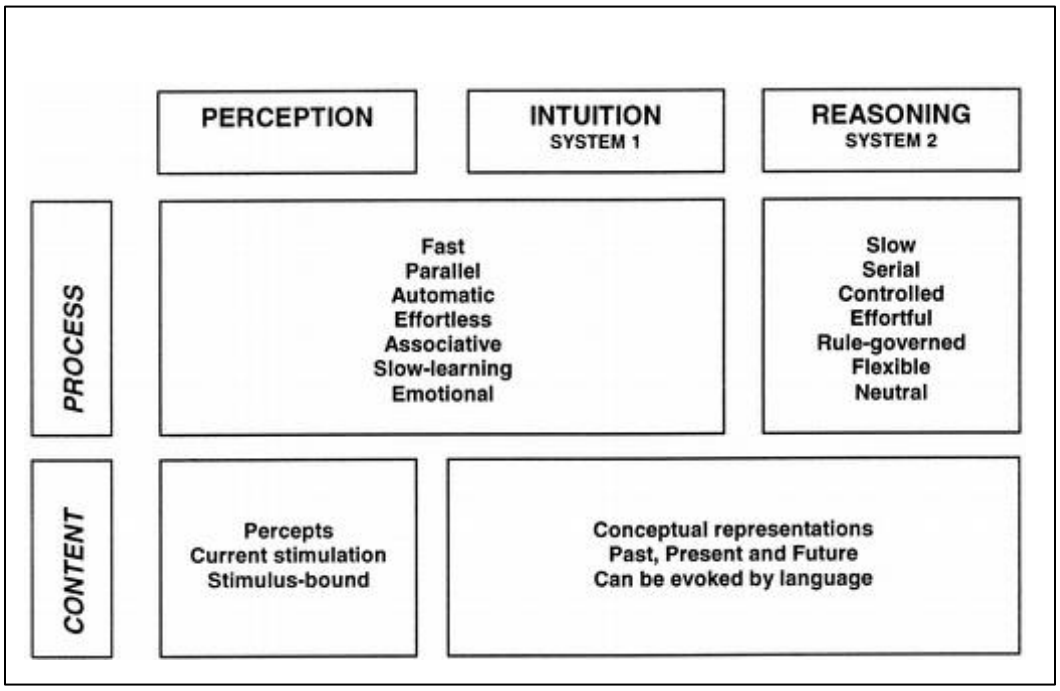
People use heuristics in order to simplify complex tasks or assessments; in essence, heuristics aid in the reduction of complexity as well as cognitive and biological resources needed for the decision-making process (Kahneman, 2003). Gigerenzer (2008) articulates that the simplicity of heuristics does not necessitate its inaccuracy, as “heuristics can be more accurate than complex procedures” by “successfully exploiting evolved mental abilities and environmental structures” (Gigerenzer & Brighton, 2009, p.110). In fact, “ignoring part of the information can lead to more accurate judgments than weighting and adding all information, for instance for low predictability and small samples” (Gigerenzer & Gaissmaier, 2011). Such has held true in various fields, including business, organizations, health care, and law (Gigerenzer & Gaissmaier, 2011). Specifically, heuristics can be beneficial in situations when people need to act quickly and in situations of complexity where “probabilities or utilities are unknown, and multiple goals and ill-defined problems prevent logic or probability theory from finding the optimal solution” (Gigerenzer, 2008, p.20). They too “exploit evolved capacities..., and thus they can provide solutions to problems that are different from strategies of logic and probability. In addition, they are tools that have been customized to solve diverse problems” (Gigerenzer, 2008, p.22). This is most certainly the case in an educational context as well with teachers often having to make grading decisions for over one hundred students within a short time frame--often with complex data input.

While there are many specifically identified heuristics and processes on which individuals often rely, overarchingly, decision-making processes are often classified into two more broad categories: System 1 and System 2 thinking. System 1 thinking is generally

understood to rely on associations and intuition, while System 2 thinking is more computational, probabilistic, and logical in nature (Gigerenzer & Regier, 1996; Gigerenzer, 2008; Kahneman, 2003). Kahneman (2003), details more clearly the distinctions in cognitive processes between System 1 and System 2 thinking in the following infographic:

Figure 1

System 1 versus System 2 Thinking



Note. From "A Perspective on Judgment and Choice" by Kahneman, D., (2003), The American Psychologist, 58(9), 697–720, p. 698 (https://doi.org/10.1037/0003-066X.58.9.697).

System 1 thinking and decision-making, according to Kahneman, is fast, parallel, automatic, effortless, associative, slow-learning and emotional; while System 2 thinking is slow, serial, controlled, effortful, rule-governed, flexible, and neutral. However, System 1 and System 2 thinking are not necessarily mutually exclusive: they do not occur in isolation and often individuals will exhibit both System 1 and System 2 thinking processes--engaging in both

intuitive decision-making processes and reasoning based ones concurrently or in succession. Kahneman explains that a judgement or choice may be made in the following ways:

- An intuitive judgment or intention is initiated, and
    - Endorsed by System 2;
    - Adjusted (insufficiently) for other features that are recognized as relevant;
    - Corrected (sometimes overcorrected) for an explicitly recognized bias; or
    - Identified as violating a subjectively valid rule and blocked from overt expression.
- No intuitive response comes to mind, and the judgment is computed by System 2.

(Kahneman, 2003, p.717)

Heuristics and System 1 thinking, then, are arguably the default cognitive process that humans undertake. Kahneman indicates that it is only when System 1 thinking is not fruitful that one then makes a decision using System 2 processes. Not all researchers agree, however, and others have posited that "the 2 types of cognitive processes are embedded in different naive metaphysical systems and tacit epistemologies," and, thus, default cognitive processes are not the same for all situations and all people in all cultures (Nisbett, 2001, p.291). Nisbett argues that "epistemology dictates the development and application of some cognitive processes at the expense of others…[and] social organization and social practices can directly influence the development and use of cognitive processes such as dialectical versus logical ones" (p.292).

In the explication and evaluation of System 1 thinking, there are two sometimes contradicting modes of thought: heuristics and biases (HB) and naturalistic decision making (NDM). Both models focus on intuition and expertise. The NDM approach focuses on expert

intuition and has its etymological roots in the work of deGroot (1946/1978) and Chase and Simon (1973) in which they conducted research on master chess players' ability to make the most beneficial moves quickly (Kahneman & Klein, 2009):

> A central goal of NDM is to demystify intuition by identifying the cues that experts use to make their judgments, even if those cues involve tacit knowledge and are difficult for the expert to articulate. In this way, NDM researchers try to learn from expert professionals. Many NDM researchers use cognitive task analysis (CTA) methods to investigate the cues and strategies that skilled decision makers apply (Crandall, Klein, & Hoffman, 2006; Schraagen, Chipman, & Shalin, 2000). (Kahneman & Klein, 2009, p.516)

While the NDM approach exhibits a positive deference for the intuitive decision-making of experts, the HB approach is more critical of expertise and the judgement of experts. The HB approach, according to Kahneman and Klein (2009), finds that "inconsistency is a major weakness of informal judgment" and "human judgments are noisy to an extent that substantially impairs their validity" (p.517). Both Meehl (1954) and Goldberg (1970) as cited in Kahneman and Klein evidenced such when researching the decisions of judges, with judges exhibiting a bootstrapping effect and inconsistency when given the same case information at different times (Kahneman & Klein, 2009). As such, consistency, according to the HB approach, has been identified as the primary advantage of algorithms over humans (Karelaia and Hogarth, 2008 as cited in Kahneman & Klein, 2009). Grove, Zald, Lebow, Snitz, and Nelson (2000) further evidenced that there exist conditions under which mechanical and analytical judgments outperform human judgment (Kahneman & Klein, 2009). However, HB does not affirm that heuristics and intuitions are always incorrect, but rather that "they are less trustworthy than

intuitions that are rooted in specific experiences" (Kahneman & Klein, 2009, p.522). Such trustworthy intuition is not easily trained into people, though, as "Tversky and Kahneman (1971) ...concluded that faulty statistical intuitions survive both formal training and actual experience" (Kahneman & Klein, 2009, p.518). Despite this survival of faulty intuition and inconsistency, there do remain researchers, such as Gigerenzer (2021), who posit that "There is lack of evidence that violations of logical rationality have detrimental consequences on people's wealth, health, happiness, proportion of true versus false beliefs, or some other measurable cost" (p.3553).

Decision-Making, Heuristics and Education

Decision-making and heuristics have been heavily studied in the fields of gaming, law, and medicine. And, while there are some studies that have been done in the field of education, most pertain to pedagogical decision-making or hyper-specific grading decisions such as singular exam markings. Additionally, general theoretical suppositions regarding educational decision-making based on analogous reasoning, for example, have been made by individuals such as Popham, Snygg and Combs, and Crisp.

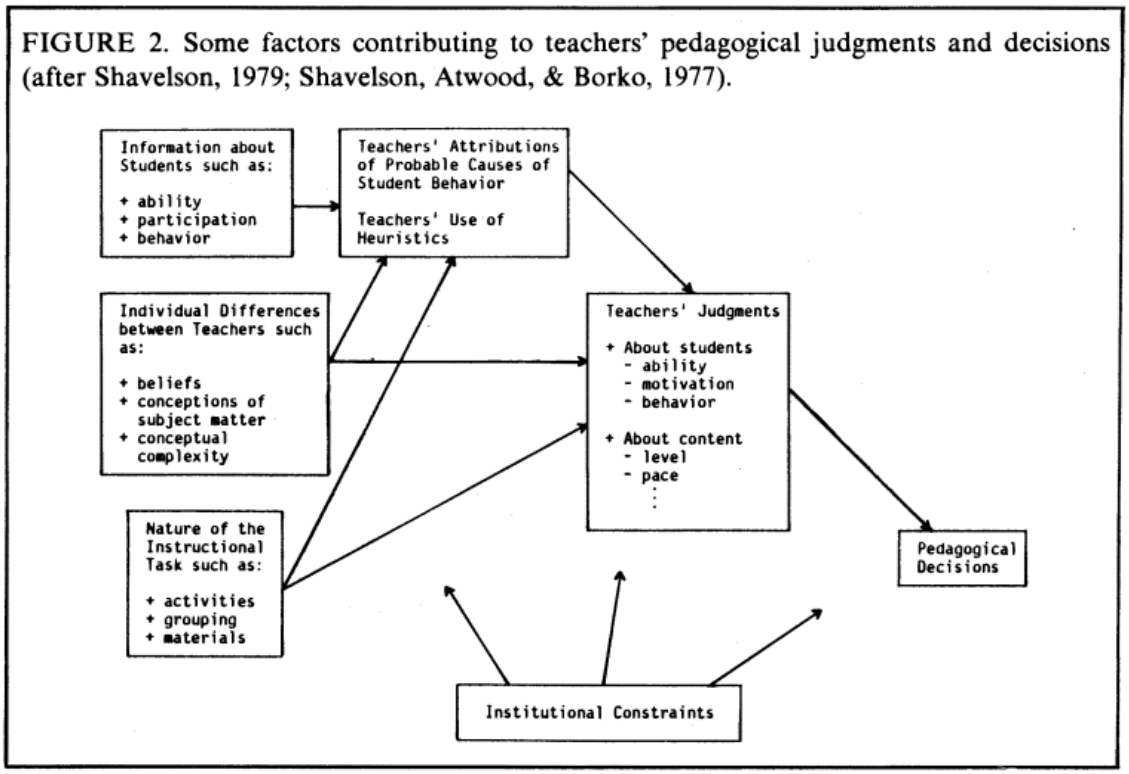
In discussing the thinking of educators, W. James Popham (2018) notes that "moving from informal evaluation as a child to formal evaluation as a professional involve[s] shifting emphasis from self-interest to objective evidence" (p.39). And, while there is merit in this aim of objectivity, Popham argues, in alignment with Snygg and Combs, that "people are primarily motivated by whatever enhances the phenomenological self" (p. 39). As humans grow into professionals, or educators (as is Popham's emphasis), Popham suggests that "systematic evaluative thinking, which involves identifying the factors of consequence, collecting evidence regarding those factors, synthesizing them, and evaluating their importance, does not happen

much of the time" (p.40). This indicates that there is, what Popham deems, 'gut-based evaluation' occurring in education, which should be feared because such "choices are more likely to be self-serving, not in the best interest of children, and as a consequence children end up being short changed" (p.40). Popham, thus, suggests that educators make decisions that are not, in fact, defensible or in the best interest of children due to their fast, gut decision-making processes.

In contrast to Popham (2018), Shavelson and Stern (1981) explicate what they identify as teachers' complex cognitive processes when it comes to pedagogical decisions, per Figure 2 below.

Figure 2

Some Factors Contributing to Teachers' Pedagogical Judgements and Decisions



Note. From "Research on Teachers' Pedagogical Thoughts, Judgments, Decisions, and Behavior" by Shavelson, R. J., & Stern, P. (1981). Review of Educational Research, 51(4), 455–498, p.472 (https://doi.org/10.3102/00346543051004455).

Shavelson and Stern detail the differing antecedent conditions, teacher characteristics, and cognitive processes that influence and yield teachers' pedagogical decisions. They, too, suggest that there are certain heuristics teachers commit when making complex pedagogical decisions. Attributions, for example, "refer to teachers' estimates of student ability because they deal with the processes by which people integrate information to arrive at causal explanations for events (Borko & Shavelson, 1978)" (Shavelson & Stern, 1981, p.470). Such inferential attributions by teachers may be "to reduce the complexity of the original information available to them as

antecedents to their judgments, decisions and behavior (Shavelson, 1976; Shavelson & Atwood, 1977; Shavelson, Cadwell & Izu, 1977). And teachers seem to make decisions and carry them out on the basis of these judgments rather than on the basis of the original information (Borko, 1978; Cone, 1978; Russo, 1978; Shavelson & Borko, 1979)" (Shavelson & Stern, 1981, p.470). Beyond ability, it is suggested that teachers are perhaps the most affected by concern for their students and their personal lives outside of the educational context, and the potentially forecasted deleterious consequence(s) of their decisions (McNair, (1978-1979), as cited in Shavelson & Stern, 1981). In response to a lack of research in this particular area, Shavelson and Stern recommend that future research should focus on "teachers' thoughts, decisions and behaviors in studying how students (e.g., class composition, conflicting goals), classroom context (e.g., social relations) and organizational context (e.g., textbook adoptions, assignment of students to teachers) influence these decisions and behaviors" (Shavelson & Stern, 1981, p.491).

Such research is validated as needed by Cain et al. (2019), whose framework asserts that "research can inform bounded decision-making, teachers' reflection and organisational learning" (p.1072). They propose that education serves multiple functions, including aiding in the development of students as individuals, socializing students, and imparting subject-specific knowledge. The complexity of aims in education makes decision-making, too, a complex process. Teachers are not simply tasked with mechanistically approaching students and algorithmically outputting scores, but rather approaching students as humans in a multifaceted web of influences (socially, societally, systematically, etc.). While a simplified version of such a process may appear as follows in Figure 3, Figure 4 offers a more comprehensive view.

Figure 3

Simplified Teacher Decision-Making Process

assumptions→ understanding evidence → discussion→

decisions→ actions

(Cain, Brindley, Brown, Jones & Riga, 2019).

Figure 4

More Comprehensive Teacher Decision-Making Process

assumptions (spoken and unspoken) → understanding evidence (of differing weight, and usually involving comparisons) → discussion (within social contexts) → decisions → actions

(Cain, Brindley, Brown, Jones & Riga, 2019).

Both the complexity of teacher decision-making and that teacher decision-making may be error-prone necessitates action (Popham, 2018; Cain et al., 2019). “Several different types of errors, common to teaching, are possible to attribute to the weaknesses of ‘fast’ (‘System 1’) thinking....They can jump to conclusions too readily; they notice evidence that supports their existing beliefs whilst ignoring contradictory evidence; and they overestimate the extent of their pupils’ existing knowledge (Shavelson, 1983; Calderhead, 1984)” (Cain et al., 2019, p. 1079). Existing and future research can hope to unobfuscate the complex decision-making process as well as create a foundation on which educational stake-holders may build conversations surrounding practitioner cognitive processes (Schuck et al., 2018 as cited in Cain et al., 2019).

Existing research has found that teachers weigh students’ emotions in their decision-making. In 2015, Sheppard and Levy (2019) interviewed 17 social studies teachers about how

emotions affected their classroom pedagogy. In doing so, they found the following: “teachers conceptualize the classroom as an emotional space... teachers make decisions based on students' real or anticipated emotional responses, [and] teachers prioritize the maintenance of a safe classroom space (p. 193). Similarly, Bourke et al. (2018) found, in a three-year longitudinal study in New Zealand, that teachers prioritize and engage with their students’ informal learning and think of their roles not as mechanistic and assessors, but rather as focusing on the learning and student voice. “One teacher commented ‘what is the role of the teacher? Is it about us knowing our learners, or our ability to assess them, test them and analyse the data? And at the moment the weighting is on data and analysis’. Another teacher expressed how she was cautious in her assessment practices, not feeling safe to go ‘out of the box’. Teachers felt that National Standards had brought about an increased focus on data” (p. 46). This study, much like Shavelson and Stern (1981), illuminates that teachers’ organizational or institutional contexts impact their decision-making and practice.

Teachers’ decision-making in their culminating assessments of students has not been readily studied, but there is research pertaining to more localized assessment. One more meta-study by DeLuca et al. (2018) interviewed 404 teachers from North America on their assessment purposes, processes, and fairness. One key finding was that “teachers overwhelmingly identified that they respond to assessment scenarios by altering the scoring or administration of assessments instead of attending to assessment design or communicating assessment results” (p. 363). Such score altering and the cognitive process that underpins it is not present in the literature. The study additionally found that teachers supported an equitable approach to assessment, as opposed to a standardized one. However, DeLuca et al. note the following:

> One significant difference was observed based on teachers' career experience. Specifically, teachers' approaches to Assessment Fairness became more differentiated with increased years of teaching experience. Early career teachers tend towards a standardised approach with mid-career teachers tending to employ an equitable approach. The most experienced teachers in this sample generally favoured a differentiated approach. Congruent with previous research (Cowan, 2009; Smith et al., 2014), we understand this finding to suggest that teachers adopt a more complex, student-focused approach to assessment fairness as they gain in-service experience. (p. 370)

DeLuca et al. conclude by proposing that future research is needed not only in assessment literacy, but also in "how teachers approach various classroom assessment scenarios" (p. 372).

More specifically, studies have examined teachers' assessment practices and cognitive processes in relation to project-based learning courses and single examination marking. Gweon et al. (2017) observed instructors' group assessments in a graduate level engineering project-based learning course as exhibiting "cognitive biases and judgments made using incomplete information in the context of an engineering design education classroom. More specifically, we hypothesize that instructors would be susceptible to human errors that are well known in social psychology, the halo effect and the fundamental attribution error, because they have a limited view of group work when they facilitate distributed and remote groups" (p. 165).

Wieman and Welsh (2016) additionally evidenced that instructors exhibit the fundamental attribution error: this time it held for university math and science instructors more generally. Their data was gathered by way of instructor surveys focusing on student learning at a North American public university. While some research has shown that eliminating cognitive biases, such as the fundamental attribution error, is challenging (Tetlock, 1985; Burger, 1991 as cited in

Gweon, 2017), Gweon et al. (2017) recommend that further research is needed and that instructors should use an explicit assessment tool that differentiates assessment categories in order to counter the halo effect and fundamental attribution error evidenced. Such explicit assessment tools could prevent instructors from allowing pre-existing character judgements or pre-existing assessments of students to influence their grading.

Crisp (2010) also contends that instructor assessment requires more research attention: "the judgement processes underpinning examination marking are central to achieving fair assessment but are under-researched" (p.1). Crisp sought to examine one such examination process when he collected data from six examiners whilst they scored narrative responses on geography exams--asking them to think aloud as they scored each exam. Crisp outlines and observes multiple judgement processes and heuristics. Configurational assessment, stemming from Kaplan (1973) and Sadler (1989) as cited in Crisp (2010), occurs when judges first make a holistic impressionistic assessment and then substantiate this assessment with criteria as they assess (Crisp, 2010). Milanovic *et al.* (1996) coined an assessment model consisting of pre-marking, scanning (with evaluation), reading quickly, rating, and modifying. (Crisp, 2010). Lastly, Freedman and Calfee (1983) proposed a three-stage model of the cognitive process of assessment: the reading or comprehension, the evaluation, and the articulation of the evaluation (Crisp, 2010). Ultimately, Crisp discerned that "these studies have started to provide a sequential (or looping) model of the phases in the marking process…. [and] where judgements are difficult, heuristics may be particularly likely to come into play at an unconscious level or a greater extent of conscious deliberation (e.g. via explicit comparisons) may be involved" (p. 5). Throughout the examination marking process, Crisp found raters to have relied on three primary heuristics: "'availability' (to do with the ease with which examples can be brought to mind),

‘representativeness’ (assessing similarity to a prototype) and ‘anchoring and adjustment’ (picking an anchor point and then adjusting from this)” (p. 4).

While there is not much existing research regarding instructors’ cumulative assessment of students and the cognitive processes therein, there exists in the literature research surrounding the assignment of failing grades, specifically. Halligan (2011) sought to explore “no fail” and “social promotion” policies in both the United States and Thailand by interviewing and surveying educators, parents and teachers. Teachers in both locales indicated they believed that students are de-motivated to learn if they receive an “f” or a failing grade” and that “there should be some system of holding back students until they are able to achieve an adequate level of mastery” (p. 88). However, the study found that Thai teachers were more likely than their US counterparts to favor changing students’ grades in order to help low achieving students.

Analogous Studies in Medicine and Medical Education

While research in heuristics and decision-making processes in elementary, middle, and secondary contexts is limited, there have been analogous studies in the medical education field, specifically, that may inform future studies. Feufel (2019) asserts that medical education should embrace heuristic thinking. Likewise, Green (2019) argues that “like clinical decisions, decision making about learners takes place in similarly 'noisy' contexts in which a learner's workload, physical and emotional well-being, educational background and training, interpersonal skills and knowledge base all contribute to his or her performance. Learners, like patients, may not be amenable targets for purely rational, rules-based decision making” (p. 322). Green expands by reasoning that if education and assessment were meant to be purely logical in nature and devoid of nuance that there would be no need for structures that we have in place in medical education, such as promotion or clinical competence committees. “Instead, the truth is that many of our

academic and behavioural standards include evaluative grey areas at their edges, and our judgements of competence and achievement are nuanced" (p. 323). While Green was speaking here to medical education, her reasoning and claims may transfer into other academic contexts such as secondary education. However, while Green recognized the 'noisy' nature of academic contexts and assessing learners, she goes on to warn of the dangers of 'pathologizing' learners. She suggests that "the heuristic strategies used by educators and those in student and trainee support roles are deeply rooted in our beliefs about learners' capacity for change, about intellect and ability as relatively stagnant or fluid concepts, and in our perception of unprofessional behaviour as an issue of character, a symptom of being unwell, or some combination of the two. Our 'rules of thumb' when it comes to learners are as ingrained and persistent as those we use in patient care, yet refining them receives far less professional development time and attention" (Green, 2019, p.323). Accordingly, Green recommends that faculty development should focus on identifying and 'refining' heuristics that benefit learners and "should include strategies to actively guard against the discriminatory power of those biases (Green, 2019, p.323).

The human element of overall assessment in medical education is further explored via Couper (2018). Couper sampled 390 nursing faculty with an open-ended questionnaire--delving into "failure to fail". While "failure to fail" has been studied and identified in the United Kingdom and Canada (Duffy, 2003; Luhanga et al., 2008 as cited in Couper, 2018), "educators in the United States have not been willing to openly talk about the struggle to assign a failing clinical grade" (Couper, 2018, p. 135). Couper, in his study, found role strain, organizational support, and faculty stress to be significant factors in the decision-making process to assign a failing grade. This resulted in 17.4 percent of nursing faculty not assigning an earned failing grade. Couper calls for more research to better understand "the dynamics of the process" (p.

136). This holds true not only for a hesitance to assign failing grades in medical educational contexts, but in more primary and secondary level academic contexts as well.

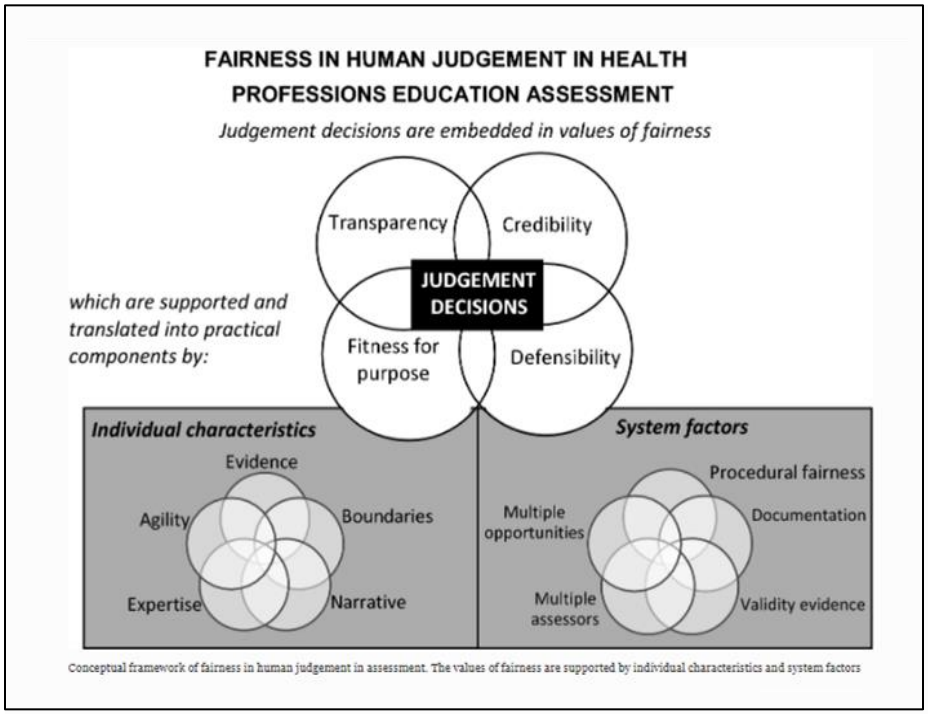
Klein (2005) details cognitive processes and biases that he has identified as affecting medical decision-making, and reinforces Couper's (2018) idea that professional development is needed. Many of these same biases may be present in educational contexts and be committed by teachers as well due to some situational constraint overlaps. Klein (2005) suggests that many of the cognitive biases are evidenced due to the rapidity of the decision-making process required in medical contexts. This also holds in education, as teachers are often tasked with adapting to the needs of their learners and classroom in the moment, and are tasked with assessing hundreds of students with very little prep time--thus requiring them to make quick decisions regarding marks. Klein points out that highly trained professionals, such as doctors (and teachers) are not immune to pitfalls and errors in decision-making. In fact, "it is common for people who are particularly prone to cognitive biases to believe that they are good decision makers" (p. 781). Three of the pitfalls on which Klein expounds are relevant for this study: the representativeness heuristic, the availability heuristic, and confirmatory bias. Representativeness was exhibited in doctors' tendency to stereotype and assume correspondence, availability was exhibited when doctors placed undue weight on that which came to mind easily, and confirmatory bias was exhibited by doctors' searching for evidence to fit their predisposed or predetermined judgements and ignore that information which counters them. This would lead practitioners to dismiss critical information. Klein concludes with the import of awareness and training in order to overcome potential biases that might result in flawed decision-making and affect patient care.

With a focus on determining what was fair in a health assessment and to inform future research, Valentine et al. (2021) completed a hermeneutic literature review in which they

synthesized the existing literature related to fairness in human judgement. They argue that "changing focus to look at what is 'fair' human judgement in assessment, rather than what is 'objective' human judgement in assessment allows for the embracing of many different perspectives, and the legitimising of human judgement in assessment" (p. 714). They distinguished between fairness being reinforced at the individual level via values, such as evidence and expertise, and at the system level via procedures, such as documentation and multiple assessors. These influencers are shown in more detail in Figure 5 below.

Figure 5

Fairness in Human Judgement in Health Professions Education Assessment



Note. From "Fairness in human judgement in assessment: a hermeneutic literature review and conceptual framework" by Valentine, N., Durning, S., Shanahan, E. M., & Schuwirth, L. (2021). Advances in Health Sciences Education : Theory and Practice, 26(2), 713–738, p.719 (https://doi.org/10.1007/s10459-020-10002-1).

For judgement decisions to be fair, according to Valentine et al., they need to be transparent, credible, defensible, and fit purpose; but they also need to be supported by individual and system factors in practice. In crafting a recommendation, they suggest that their literature review (with special emphasis on Lucey and Souba, 2010 as cited in Valentine et al., 2021) showcases the complexity of making fair judgements and the influencing factors; thus, recognizing this complexity and not attempting to overly simplify professional development is paramount (Valentine, Durning, Shanahan & Schuwirth, 2021).

Theoretical Framework and Anticipated Heuristics

System 1 and System 2 Thinking

In the field of cognitive psychology, System 1 and System 2 thinking have been identified as overarching processes that govern decision-making (Stanovich & West, 2000). System 1 decision-making processes are often quick, associative, and intuitive in nature being at times emotionally driven and habitual (Kahneman, 2003). Heuristics are often entailed within System 1 thinking, as they are often seen as antithetical to logic (Gigerenzer, 2008). System 2 decision-making is, on the other hand, slower and more deliberate. These processes are seen as more analytical and logical in nature and are often rule or algorithm governed (Kahneman, 2003). Overall, System 1 is associated with intuition while System 2 is associated with reasoning.

In this study, it is likely that interviewees will evidence one or both of these modes of thinking. Since grading is such a complex, multifaceted process with humans involved on both sides of the decision, it is probable that there will be elements of analysis and logic underpinning decisions, but there may also be cases in which educators rely on their intuition, expertise, or emotion in order to ultimately determine whether a student's grade is accurate. For example,

grades may be initially calculated using a rule-governed process or algorithm in a computer, but teachers may or may not decide to overrule that computed mark if it does not align with what they believe to be the student's deserved mark. Different heuristics may then be committed in the process of evaluating the computed mark's validity.

Algorithmic Thinking and Axiomatic Rationality

Falling under the umbrella of System 2 thinking are algorithmic thinking and axiomatic rationality. Algorithmic thinking involves a reliance on and application of algorithms to determine an outcome (Ciobanu & Popovici-Bujor, 2020). This may be evidenced by teachers who adhere to their prescribed computational models of grades, and are devout in their thinking that such grades are the logical and fair mark. Such teachers would not modify grades and would see the grading process as a more objective process.

Axiomatic rationality too falls under System 2 thinking in that it is analytical and reasoning based. It is defined as "conformity to abstract axioms" and is, by some, seen as the "norm for how human beings should reason, arguing in addition that violations would lead to real costs" (Gigerenzer, 2021, p. 3547). However, Gigerenzer posits that there is not substantial evidence of such costs. Cost aside, some researchers agree that a strict adherence to norms or axioms is only feasible in 'small worlds', or in more simplistic contexts (Savage, 1954 as cited in Gigerenzer, 2021; Gigerenzer, 2021). In more complex, intractable, or uncertain situations heuristics may be more appropriate and effective than axiomatic rationality (Gigerenzer, 2021). There may be evidence of an adherence to axioms in this study as well as a violation of logical rationality: it is anticipated that many teachers may have rules or principles that govern how they assess and when, for example, they will round a grade, but there may also be situations in which

teachers take into consideration situational complexities when making their ultimate grading decision.

Halo Effect

In making grading decisions, it is anticipated that teachers may rely on System 1 thinking and heuristics and, more specifically, commit the halo effect. The halo effect is a form of cognitive bias in which the judge is influenced by their existing impression of a person's character when making decisions or judgements (Thorndike, 1920; Beehr et al., 2001; Lambart et al., 1997; Becker & Cardy, 1986; Feeley, 2002 as cited in Gweon et al., 2017). While the halo effect has been studied and found to be present in military and some professional contexts, the halo effect has not yet been studied in teachers' evaluations of students (Gweon et al., 2017). This study seeks to fill this gap and explore whether some instructors are, in fact, biased by their preexisting judgements of students' characters in the grading process. For example, teachers may be more or less amenable to boosting, or bumping, grades depending on how they view the student's character.

Fundamental Attribution Error

Another System 1 heuristic that may be observed in this study in the fundamental attribution error, also termed correspondence bias--as it is not always committed in error and is not as pervasive as it was once thought to be (Miller, 1984; Harvey et al., 1981 as cited in Howell & Shepherd, 2011). It is defined as "the tendency for people to make dispositional, rather than situational, attributions for an actor's behavior" (Jones, 1979 as cited in Gilbert et al., 1988). The judge in such cases shows deference for people's dispositions as the cause of behavior and they, thus, neglect other factors that may influence behavior, performance, etc. Wieman and Welsh (2016) explain that "an example of this would be a teacher attributing educational

outcomes entirely to the characteristics of the student, rather than considering any factors such as the educational environment or the quality of the instruction" (p.645). According to Gilbert, Pelham, & Krull (1988), "the correspondence bias has several possible causes, including (a) a lack of awareness of the situational influences, (b) unrealistic expectations of others, (c) inflated categorizations of behavior, and (d) incomplete corrections for bias (Gilbert & Malone, 1995)".

In committing the fundamental attribution error in the context of this study, a teacher might see the grade as a direct result of a student's internal disposition or abilities and neglect to consider external factors that might influence that student's performance. Weieman and Welsh (2016) showed that "it is quite common for university math and science instructors to exhibit fundamental attribution error when considering limitations in student learning," but this heuristic has not been extensively studied at the high school level (p.648).

Intuitive Judgement and Expertise

It is also quite probable that interviewees in this study will evidence System 1 intuitive judgements perhaps justified by their professional expertise. Such intuitive judgement is complex and has been found to yield accurate and skilled judgements as well as erroneous ones (Morewedge & Kahneman, 2010). The Naturalistic Decision-Making Approach (NDM) highlights the successes of such expert intuition (deGroot, 1946; Chane & Simon, 1973; Klein et al., 1986 as cited in Kahneman & Klein, 2009) with Simon defining intuition as "the recognition of patterns stored in memory" (Kahneman & Klein, 2009, p. 516). Kahneman and Klein (2009) reaffirm that "intuitive judgments can arise from genuine skill…, but that they can also arise from inappropriate application of the heuristic processes on which students of the HB tradition have focused" (p.524). They also recognize some complexities of expert intuition: for example, experts are not always aware of the cues that guide their decision-making and non-experts are

even worse. This is problematic in that "subjective confidence is therefore an unreliable indication of the validity of intuitive judgments and decisions" (p. 524). Thus, while the decision research community, the sociology community, the heuristics and biases community, the evidence-based practices community, and the computer science community may challenge expertise (Klein et al., 2017), intuitive thinking has also been found to be beneficial and accurate (Simon & Chase, 1973; Gawande, 2002; Klein, 1998; Klein, 2003; Wilson & Schooler, 1991 as cited in Kahneman, 2003).

In the context of this study, intuitive expertise may dominantly be evidenced with more tenured teachers. Such experience may give teachers confidence since they have seen many patterns of student performance and outcomes throughout their career and in their particular subject matter and are, as a result, using those patterns to make current judgements. For example, a teacher may have confidence that they "just know" that a particular student is an "A student" or a "B student" based on their intuitive overall assessment of that student's performance throughout the course; they may not feel the need to necessarily cite specific evidentiary support for such intuitive judgements.

The Availability Heuristic

There are times when decision-makers, including educators, are faced with complex tasks with a great deal of informational input that could be weighed in decision-making. At times, there is almost too much information and judges are prone to weigh most heavily the information that first comes to mind or that seems most pertinent. Showing deference for that which comes to mind most easily is defined as the availability heuristic (Klein, 2005). This may also overlap with the recency heuristic due to that which is recent often being most easily recalled in decision-making (Lee et al., 1998). Klein (2005) argues that "in general, this guides us in the right

direction, as things that come to mind easily are likely to be common, but it may also mislead" (p.782). Interviewees in this study may evidence the availability heuristic if they, in evidencing or justifying their grading decisions, rely on evidence that comes to mind easily in lieu of looking back at student grades throughout the course.

Confirmatory Bias

Confirmatory bias is also a System 1 thinking heuristic relevant to this study. Confirmatory bias is "the tendency to look for, notice, and remember information that fits with our pre-existing expectations. Similarly, information that contradicts those expectations may be ignored or dismissed as unimportant" (Klein, 2005, p.782). Klein studied how such confirmatory bias surfaced in doctors' diagnostic decision-making and ultimately concluded that it is "critical to remain constantly vigilant for any information that may contradict [an] existing diagnosis, and to give any such information careful consideration, rather than dismissing it as irrelevant" (p. 783). This could similarly apply in an educational context--as it would also be crucial for teachers to ensure that they are examining all pertinent information in their diagnosis of students.

Confirmatory bias may be present in the interviewing of teachers regarding their ultimate grading assignments if they, for example, intuitively believe that a given student is an "A student" and then they recall and cite evidence to confirm this pre-existing notion. It may also surface if, for example, an educator knows that a student has always earned As in their previous courses and then reinforce this pre-existing knowledge of the student as an "A student" by relying on information from their course that validates such. This may also mean that said instructors ignore pertinent information that invalidates their pre-existing conceptions.

The Affect Heuristic

The affect heuristic is one of the more recently identified heuristics: only being coined in 2002 by Slovic and his colleagues (Slovic et al., 2002). There had been previous evidence to suggest that evaluations may be affective--whether conscious or not (Bargh, 1997; Zajonc, 1980, 1998 as cited in Kahneman, 2003). Slovic et al. (2002) built upon this idea and "discussed how a basic affective reaction can be used as the heuristic attribute for a wide variety of more complex evaluations, such as the cost-benefit ratio of technologies, the safe concentration of chemicals, and even the predicted economic performance of industries" (Kahneman, 2003, p.710). Kahneman and Ritov (1994) and Kahneman et al. (1999) expanded Slovic's findings and "proposed that an automatic affective valuation—the emotional core of an attitude—is the main determinant of many judgments and behaviors" (as cited in Kahneman, 2003, p.710).

Since educators are often assessing students whom they have come to know well, it could be expected that the affect heuristic might be committed in the grading process. In other words, educators might be swayed by their emotions in their ultimate assessment of their students. There are many ways in which this might specifically surface. For example, a teacher may be reluctant to assign a student a failing grade as they may feel badly about doing so, especially if they have positive feelings toward that student. The opposite could be true: educators might not be as willing to give a higher grade to a student for whom they have ill feelings.

Conclusion

Stiggins, in 1991, first introduced the concept of assessment literacy and "advanced the argument that teachers and school leaders absolutely must understand the basic principles of sound assessment practice. Specifically, if [they] are to develop truly effective schools, educators must understand how to gather dependable evidence of student achievement and use the

assessment process and its results either to support or to certify student achievement depending on the context" (Stiggins, 2014, p.67). However, this task is not always straight-forward as educators are asked to more than objectively assess students with standardized tests that might align with clear-cut, logical, System 2 thinking. Educators are now tasked with promoting "higher levels of achievement than ever, while narrowing achievement gaps, assuring lifelong learner competencies for all students, and aspiring to universal high school graduation" (Stiggins, 2014, p.69). Additionally, educators are now asked to consider whether their pedagogical and assessment practices are equitable and whether they should focus on assessment for accountability versus assessment for improvement (Banta, 2007). Suskie (2007) contends that, in response to such complex forces, "faculty today are grappling with understanding the nature, methodology, and value of assessment. Fortunately, there is a growing body of literature on the assessment of student learning in higher education, and the emphasis is on simple, flexible approaches" (p.102). This, then, requires educators to consider more than the axioms and algorithmic outputs of grading systems. As a result, teachers are placed in the position of judge and must make complex and crucial cumulative assessment decisions. However, as Stiggins points out, "very few practicing teachers and almost no practicing school leaders have been trained to develop quality assessments or to use them in effective ways regardless of the purpose" (p. 69). This, it is expected, yields a landscape in which teachers may make biased decisions and rely on heuristics in their decision-making processes.

Chapter 3: Methodology

Design

This multiple case study using a think-aloud protocol had a primary purpose of investigating why secondary teachers weigh certain factors most heavily in their grading process, and what, if anything, motivates a grade change at the end of a term. More specifically, the following essential questions were of focus:

- How do secondary teachers make final decisions about students' summative grades?
- What cognitive biases do they rely on to make grade decisions and, if relevant, grade changes?

These questions were explored via a qualitative, multiple case study of secondary teachers' decision-making processes around assigning summative grades. This study sought to add to an understanding of the cognitive processes and motivations behind why teachers do what they do in terms of grading. Such an understanding requires in-depth answers to questions that may not be easily gathered in some formats, such as surveys. Hence, interviews were undertaken to better understand and gather such decision-making insights. Furthermore, participants may need to elaborate on their answers, requiring a back-and-forth, in-person methodology.

Sampling and Population

The study sample was drawn from a single high school in Oregon. A pseudonym, Oak High School, will be used to protect study participants. Oak High School's most recent high school enrollment (as of 2019), per Oregon Department of Education, is 1,207 students. Demographic composition of the student population is as follows: Asian (4%), Black/African American (1%), Hispanic/Latino (19%), Multiracial (7%), Native American/Pacific Islander (1%), and White (68%). 95% of students earn their diplomas within four years. There are 51 teachers at Oak High School with a 95% retention rate and 91% of teachers being licensed with

over three years teaching experience. Demographic composition of the teacher population is as follows: Asian (7%), Black/African American (1%), Hispanic/Latino (8%), Multiracial (0%), Native American/Pacific Islander (0%), and White (83%). This school serves as an interesting case study in that staff are given autonomy when it comes to grading pedagogy and ultimate grade determination. Thus, there exists variance in teacher grading practices and cognitive processes to explore (Oregon, 2020).

Convenience sampling was utilized in this study with four participants being selected: two 10th grade teachers and two 11th grade teachers. Teachers of grade levels 9 and 12 were omitted from this study to eliminate grade level-specific variables that might impact teacher decision-making, such as leniency for freshman students and concern for senior graduation. Teachers were chosen from a variety of disciplines to try to capture subject matter variability. Additionally, the teachers selected were chosen to ensure a range of years of experience. Participant selection included two white participants and two participants of color.

Administration and Procedures

The four teacher participants, teachers of grades 10 and 11 varying in subject matter and teaching experience, were interviewed in 30-60 minute sessions. As a matter of convenience for the participants, interviews were conducted at Oak High School at a time of the participants choosing-- before school, after school, or at lunch. Should a participant feel more comfortable speaking openly off-campus, alternate locations were offered, but all chose to be interviewed at Oak High School. Interviews were recorded using two devices: VoiceMemo app on an iPhone, and using an Evistr 16GB digital voice recorder. Neither device is synced with a public domain, nor will the recordings be uploaded to the public domain or to any account associated with the school in the case study. Questions for the interviews were determined based on the gaps in the

current literature, questions gleaned from decision-making articles and research, and based on knowledge of the case study site and culture. Interviews were semi-structured, while simultaneously being flexible and allowing for follow-up questions and opportunities for expansion.

A think-aloud approach in response to hypothetical student vignettes was used, and was followed by open-ended reflective questions. The specific iterview protocol is appended for reference in Appendix A. The instrument consists of an introduction, a review of informed consent, and a two-part interview. Part 1 was a think aloud design, mirroring that of Zeijlmans, K., López López, M., Grietens, H., & Knorth, E. J. (2019). Participants were asked to think aloud in this part their process in determining two hypothetical students' letter grades at the end of a term. These think-alouds were in response to hypothetical student profiles given to the participants in the form of vignettes. Each student was designed to prompt teachers to think about and consider different student and academic factors, but both academic profiles were designed to be borderline grade scenarios (i.e. both students were close to the next, higher letter grade and could result in some teachers assigning the lower overall grade and some assigning the higher grade). The specific vignettes may be seen in Appendix A. In Part 2, participants were asked five reflective questions. Participants were given the opportunity to ask questions. After the interviews, the transcripts were coded and analyzed for cogntive processes, heuristics and biases, and emergent themes.

Analysis

Multiple coding approaches were utilized in the analysis of the data gathered. Top down coding using heuristics and biases was specifically used to analyze the data gathered to answer the second research question: what cognitive biases do they rely on to make grade decisions and, if relevant, grade changes? Because there exists significant research on cognitive biases, there is

a reasonable map of biases to be utilized in coding respondents' biases. The heuristics and cognitive processes that were anticipated to arise were identified and coded throughout the interviewees responses. More specifically, the following were coded: System 1 versus System 2 thinking, algorithmic thinking, axiomatic rationality, halo effect, fundamental attribution error, intuitive judgement, availability heuristic, confirmatory bias, affect heuristic, and recency heuristic. For a more detailed understanding of how each was identified in the coding process, please see Appendix C.

In addition to top down coding for cognitive processes and heuristics, top down and bottom up coding were both utilized to examine emergent themes. Initially, the facets of the vignettes that were given to the teachers, such as ELL status and behavior, were coded for presence in the cognitive process. Bottom up coding was also utilized, though, as it allowed for more flexibility given the unique features for each teacher, classroom, grade level, etc. Additional codes were then created in response to themes that arose in the teacher cognitive processes outside of the verbiage given in the vignettes. These included themes such as accuracy, anxiety, and effort.

Research Ethics

Confidentiality and anonymity were prioritized throughout the case study. Confidentiality will be achieved by using pseudonyms and prioritizing the explicit anonymity of all participants, but anonymity may not be completely feasible due to the limited population size. Readers of the study who know the researcher may be able to deduce who participants are based on the limited number of teachers of each subject matter at the school, for example. Additionally, there is a conflict of interest in that the researcher is a scholar-practitioner completing a dissertation and is currently a teacher at the site of the case study. Generally, there are not any ethical concerns

surrounding power dynamics in that the researcher is not an administrator or supervisor of the participants; they may agree to volunteer or withdraw at any time. Given that the researcher is, however, the Department Chair of the English Department at the site of the case study, no participants from the English Department were selected so as to avoid ethical concerns resulting from that potential power dynamic.

Chapter 4: Findings

This study analyzed the cognitive processes of four teachers as they sought to come to grade determinations for two different hypothetical students–both borderline students with complex profiles. Student A was an English Language Learner with a 3.2 GPA, on free and reduced lunch, with a class attendance of 75%, an assignment completion rate of 51%, a most recent test score of 61%, an average test score of 45%, and an overall class percentage of 54%. Student A was also noted to be positively and actively engaged in class, work part time, have a single parent, and help watch siblings after school hours. With an overall class percentage of 54%, in traditional grading scales, Student A would be assigned an F. Student B was presented as a student with an overall class percentage of an 89.3%, which under traditional grading scales would be a B+. Student B was described to have a GPA of 4.4, class attendance of 95%, assignment completion rate of 85%, a most recent test score of 88%, and an average test score or 92%. Student B was described to engage and participate when asked, sometimes distract others, volunteer after school, and to have come in before grades were due pleading for an A- as this would be their first B.

Teachers were asked in part one of the interview to think aloud their grade-determination processes as they examined the given information for both students. In part two, they were then asked follow-up questions regarding satisfaction and difficulty of grade-determination. For a more detailed account of the student vignettes and interview questions, please see the interview protocol in Appendix A. The data was then coded for heuristic thinking, System 1 versus System 2 thinking, and thematic surfacings. The findings are detailed in this chapter with a summary of interview responses, heuristics and cognitive processes evidenced, emergent themes, patterns and associations, and summary of findings.

Summary of Interviewee Responses

While the primary foci of this study were specific heuristics and cognitive processes employed by teachers as they determined letter grades of students, the basic summary of interviewee responses is noted below in Table 1. Identifying information for each teacher is first noted, including subject taught, gender, and years of service in the teaching profession. These specific identifiers were included as they were linked to correlating findings of the study. Participant ethnicity was not included in Table 1 as no notable correlations were found based on teacher ethnicity. Following the identifying information are succinct synopses of the teachers' grade determinations for each student at the conclusion of their think-alouds and their responses to the interview questions in part two of the interview. Student A was presented as a student with a complex personal profile and with a 54% in the course, while Student B was presented as a straight-A student with an 89.4% in the course. Please reference Appendix A for the complete vignettes presented. While these responses do not specifically align with the primary aim of the study, they contribute to the existing literature and may benefit researchers and educational stakeholders in further understanding teacher grading practices and processes.

Table 1

Summary of Responses

	Teacher 1 (9 years)	**Teacher 2 (22 years)**	**Teacher 3 (12-15 years)**	**Teacher 4 (5-7 years)**
Subject Taught	Science	Art	History	Spanish
Gender	female	female	female	male
Years Teaching	9	22	12-15	5-7
Grade Assigned Student A	I/ F	D (with ability to raise)	D-	D
Grade Assigned Student B	B+	A-	B+	B+
Deference to ___ in Determining Grade	overall class percentage	higher mark between average test score and recent test score	most recent test score	most recent test score
Satisfaction: Student A Grade	pretty satisfied, feel terribly	not at all satisfied	satisfied	very satisfied
Satisfaction: Student B Grade	satisfied	very satisfied	satisfied, but more anxiety	satisfied
Easier Student to Assign Grade	Student A	Student B	Student B	Student B
Other Info Wanted	-engagement throughout semester -behavior detail -ELL info	-test breakdown -more data	-ELL info, supports, accommodations, modifications -test corrections chances?	-quarter, previous grades -number of AP classes -background- family, SES, etc.

In examining the summary of responses, three out of four of the teachers assigned Student A, who had a 54% overall in the course and would under most traditional grading scales have earned an F, a D grade. Of the three who assigned Student A passing grade, Teacher 2, the teacher with the most experience, assigned the highest grade of a "soft D"--defined by Teacher 2 as a D letter grade with the ability to raise the grade higher with further work. And while three out of four assigned a D letter grade for Student A, only one out of four assigned an A- grade for Student B, who had an 89.3% overall in the course and would under most traditional grading scales have earned a B+. The two teachers who rounded up for Student A, but not for Student B may have, in part, done so as a result of showing consistent deference to the most recent test score. Teacher 2, however, rounded up for both students and did not exhibit a consistent deference to one metric or data point, but instead showed deference to the higher student marks to substantiate her higher grade determinations.

Overall, there was slightly less satisfaction with the ultimate grade determination noted for Student A with three out of four teachers identifying Student B as being the easier of the two to assign a grade. This may indicate that teachers struggle more with lower marks and with potentially failing students than they do with the broader question of whether to round up any grade. When prompted, teachers voiced wanting more information beyond the scope of the information provided in the vignette. Two out of the four teachers wanted more information for Student A on their ELL status, accommodations, etc. Two out of the four wanted more information about prior student engagement and previous performance, one wanted more detailed accounts of student behavior, one wanted more data generally, one wanted to know more about the students' AP course load, and one wanted a more complete picture of the students' background–including family information, detailed SES information, etc.

Heuristics and Cognitive Processes Evidenced

The four teachers' think-aloud grade-determination processes for both Student A and Student B were coded for heuristics and overarching cognitive processes. Heuristics that were noted in the interviewee responses are as follows: intuitive thinking, affect heuristic, algorithmic thinking, axiomatic rationality, fundamental attribution error, halo effect, and the recency heuristic. Please see Appendix C for a more detailed heuristics coding key. The overarching cognitive processes were categorized based on when in the teacher's think-aloud process they determined the letter grade that they would assign the student, as well as when and how frequently each teacher relied on System 1 versus System 2 thinking.

Heuristics

All four of the interviewed teachers evidenced heuristic thinking at some point in their think-alouds. Table 2, below, showcases examples of how each of the evidenced heuristics surfaced in the teachers' processes. For each of the coded heuristics, Table 2 includes prime examples and the subject to whom each quotation belongs. The examples in Table 2 are pulled from throughout the interviews with the teachers and contain examples from part two of the interview. Hence, some of the examples are from the think-aloud grade-determination process while others are from more reflective portions of the interview. These examples serve as models of how each heuristic was identified in the coding process.

Table 2

Examples of Heuristics/Cognitive Processes from Interviews

Heuristic, Cognitive Process	Subject	Examples from Interviews
Intuition*	1	"So clearly the, um, the ability to present the information is, is being hindered here"
Affect heuristic	1	"I feel terribly because I know that there's a lot of different aspects that this particular student is going through, but at the same time, I can't justify giving a passing grade"
	1	"So student a, although, emotionally, and like my heartstrings, it would be harder because it seems like this student has a lot more responsibilities on their plate."
	3	"I also feel more anxiety about that one. There's so much more pressure at our school to like I'm giving into the pressure when I pass student one or student A. And so it's like, I'm not gonna disappoint anyone with that decision."
Algorithmic thinking	1	"54% is not close enough to be able to give a passing grade. It's 6% between a 50 and a 60."
	2	"The grade, probably this one, (student B) was the easiest just cuz uh, all of these scores are really within a few points of each other."
Axiomatic rationality	3	"That is like basically passing depending on what you say the cutoff is. And for me it's usually about a 60, cause that's what it was teaching in New York."
	3	"If I can find some evidence of that base level of knowledge, I will bump them up to pass, but I don't bump kids up so that they have a higher GPA"

Heuristic, Cognitive Process	Subject	Examples from Interviews
Halo effect	1	"That to me, I would leave it as a B plus because they, they have the ability, they have the, all of the opportunities in front of them, but their, their behavioral decisions in the class is not enough for me to sway it to go up."
	3	"Um, especially cuz they come in for help. I can't help it, but that always as much as I try not to be like biased, I also am like, they worked so hard <laugh> and I would, that does generally like sway me"
Recency	3	"Their most recent test score is 61. If that test score was cumulative, I hate saying that word cumulative. Cumulative. If, uh, that is like somewhat cumulative, like a final, um, I would, I would bump them up to a D minus"
	4	I will give him like, actually a, a passing score based on this, because if the most recent test is a 61, for me, that is a passing score because I will decide the, the evaluation based on the recent learning targets.

*_Note_. Intuition here is being used synonymous with the recognition heuristic

While each of the heuristics, (intuitive thinking, affect heuristic, algorithmic thinking, axiomatic rationality, fundamental attribution error, halo effect, and the recency heuristic), were evidenced by at least one teacher interviewee, some were employed by more teachers than others. Table 3 details the total number of teachers who evidenced each heuristic at some point in the grade-determination process–whether it be in their grade-determination process for Student A, Student B, or for both students. Given that the interviewed teachers were evaluating both Student A and Student B, it is possible that the same teacher evidenced a particular heuristic for both Students A and B. In that case, the total number of teachers who evidenced either student

could be lower in sum to the number of teachers who evidenced that heuristic for A, plus the number of teachers who evidenced that heuristic for B. The table also includes a breakdown of how many teachers committed each heuristic for each student.

Table 3

Number of Teachers who Evidenced Each Heuristic/Cognitive Process in Grade Determination

Heuristics, Cognitive Processes	**Number of teachers evidenced for student A**	**Number of teachers evidenced for student B**	**TOTAL number of teachers who evidenced for either student**
INT	2	1	2
AFF	2	2	3
ALG	1	3	4
AX	2	1	2
FAE	0	1	1
HE	1	2	2
REC	3	3	3

Key: INT=intuition, AFF=affect heuristic, ALG=algorithmic thinking, AX=axiomatic rationality, FAE=fundamental attribution error, HE= halo effect, REC=recency

Based on the collated data in Table 3 above, the only heuristic that all four teachers evidenced at some point in their grade-determination process was algorithmic thinking, but none of the teachers evidence it for both students: one evidenced it for Student A and three for Student B. In other words, three teachers (versus one) relied in part on algorithmic thinking when determining whether to round up an 89.3% to an A-, but did not do so when considering rounding up a 54%. Three out of four teachers at some point exhibited the affect heuristic and the recency heuristic.

All three of the teachers who exhibited the recency heuristic did so for both students. However, the same did not hold for the affect heuristic. While three teachers exhibited the affect heuristic at some point in the interview, only two out of four did so for each student. Two teachers exhibited intuitive thinking, axiomatic rationality, and the halo effect. And while three teachers evidenced the fundamental attribution error at some point in the interview process, only one teacher exhibited the fundamental attribution error during their grade determination cognitive process.

While all of the teachers relied, at least in part, on heuristics in their cognitive processes to determine letter grades, and possibly round up letter grades, the extent to which each relied on such thinking is noted below in Table 4. Table 4 contains the total number of times each heuristic, or cognitive process, surfaced for each teacher in their interviews. Hence, Table 4 is more detailed than Table 3, but also includes information from the duration of the interview, including the entirety of the teachers' cognitive processes from grade determination through reflection.

Table 4

Total Number of Times Each Heuristic/Cognitive Process Surfaced for Each Teacher in Interview (Grade Determination and Reflection)

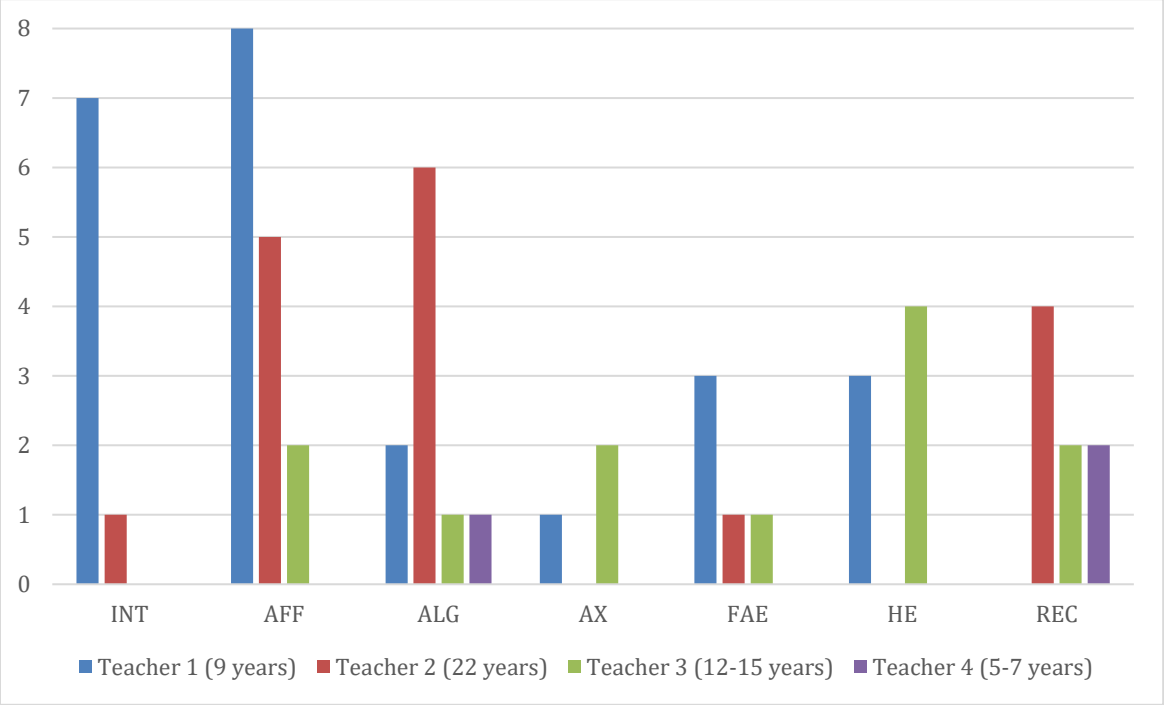
Heuristics, Cog Processes Evidenced	**Teacher 1 (9 years)**	**Teacher 2 (22 years)**	**Teacher 3 (12-15 years)**	**Teacher 4 (5-7 years)**
INT	7	1	-	-
AFF	8	5	2	-
ALG	2	6	1	1
AX	1	-	2	-
FAE	3	1	1	-
HE	3	-	4	-
REC	-	4	2	2

Key: INT=intuition, AFF=affect heuristic, ALG=algorithmic thinking, AX=axiomatic rationality, FAE=fundamental attribution error, HE= halo effect, REC=recency

Figure 6 graphically depicts the findings presented in Table 4, allowing for visual juxtaposition of the data.

Figure 6

Number of Times Each Heuristic/Cognitive Process Surfaces for Each Teacher in Interview



Key: INT=intuition, AFF=affect heuristic, ALG=algorithmic thinking, AX=axiomatic rationality, FAE=fundamental attribution error, HE= halo effect, REC=recency

Overall, Teacher 1 and Teacher 2 exhibited more heuristics in their cognitive processes, while Teacher 4 exhibited the least. Teacher 4 is also the most novice teacher, having taught for five years in the traditional classroom setting. Table 4 and Figure 6 also elucidate that even when two or more teachers both exhibit the same heuristics in their decision-making processes, the extent to which they rely on each may differ. For example, both Teacher 1 and Teacher 2 rely on intuitive thinking, but Teacher 1 had 7 occurrences throughout her interview, while Teacher 2 had one. The affect heuristic was the most apparent heuristic in total, but it similarly had a wide range of how many times it was committed by each teacher: Teacher 1 having 8 occurrences, Teacher 2 having 5, and Teacher 3 having 2. It is also discernible that teachers may commit a given heuristic only once and other heuristics may be more ingrained in their process and be

more apparent: for example, Teacher 1 committed the affect heuristic 8 times, but relied on axiomatic rationality once.

The previously included data gives insight into which heuristics are utilized by teachers in their grading decision-making processes and reflections, but it does not give insight beyond the mere presence, or lack thereof, of the coded heuristics. Table 5 delves more deeply into the cognitive processes employed by each teacher by including frequency of occurrence for each individual student. It aims to convey whether there were consistencies or discernible differences between how each teacher approached each student. Table 5 details how many times each teacher used each heuristic in their grade-determination processes for Student A versus Student B. It also includes totals of how many times each heuristic was utilized in sum for each student, so that collective consistencies and/or differences may be noted.

Table 5

Number of Times Each Heuristic/Cognitive Process Surfaced for Each Student by Teacher in Grade Determination

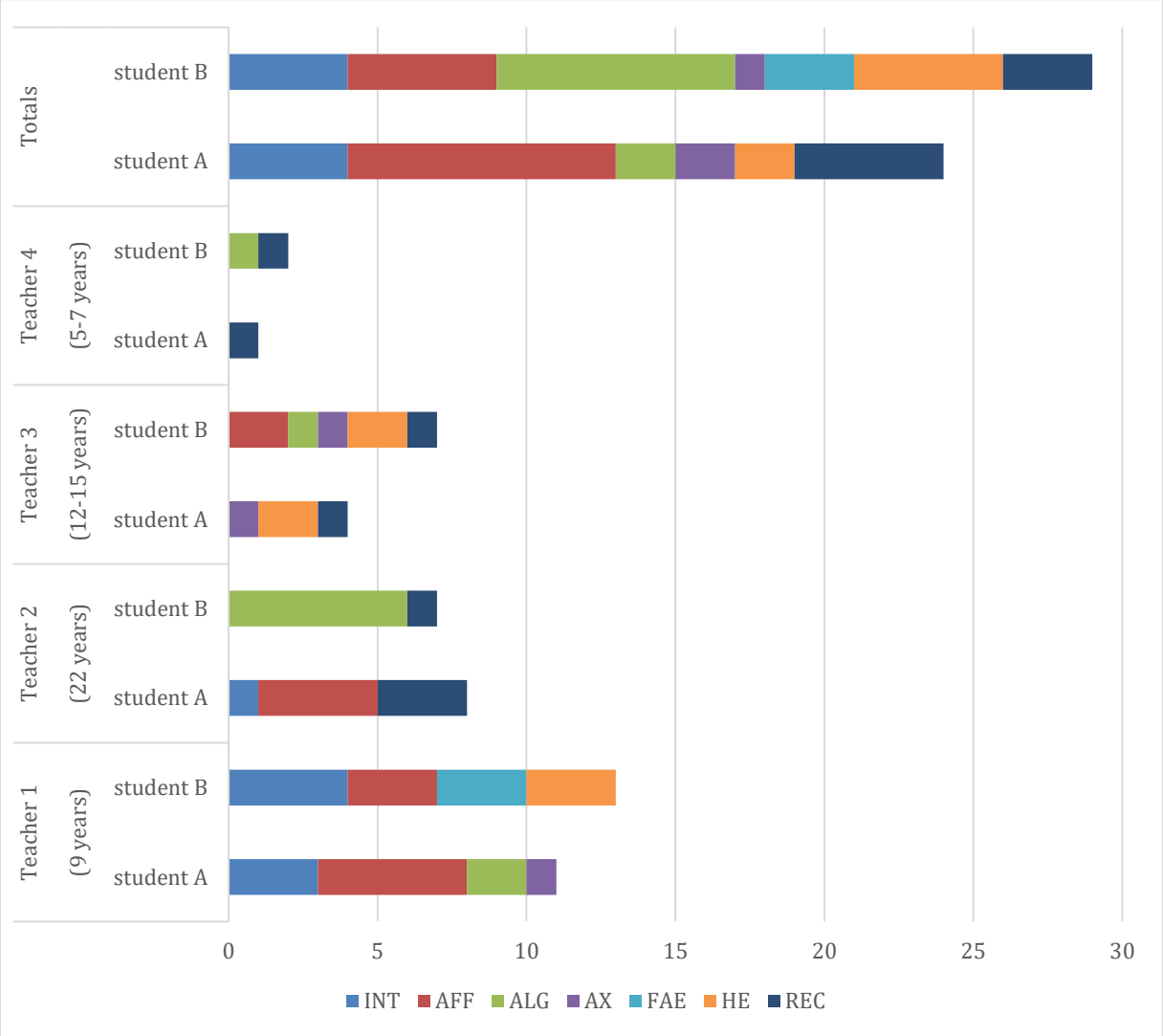
Heuristics, Cognitive Processes Evidenced	Teacher 1 (9 years)		Teacher 2 (22 years)		Teacher 3 (12-15 years)		Teacher 4 (5-7 years)		Totals	
	student A	student B	student A	student B	student A	student B	student A	student B	student A	student B
INT	3	4	1	-	-	-	-	-	4	4
AFF	5	3	4	-	-	2	-	-	9	5
ALG	2	-	-	6	-	1	-	1	2	8
AX	1	-	-	-	1	1	-	-	2	1
FAE	-	3	-	-	-	-	-	-	-	3
HE	-	3	-	-	2	2	-	-	2	5
REC	-	-	3	1	1	1	1	1	5	3

Key: INT=intuition, AFF=affect heuristic, ALG=algorithmic thinking, AX=axiomatic rationality, FAE=fundamental attribution error, HE= halo effect, REC=recency

Figure 7 graphically depicts the data presented in Figure 6 with each heuristic represented by a different color in each bar. The totals for each student are presented first, followed by a more detailed breakdown of heuristic surfacings by teacher for each student.

Figure 7

Comparing Heuristics/Cognitive Processes between Students A and B in Grade Determination



Key: INT=intuition, AFF=affect heuristic, ALG=algorithmic thinking, AX=axiomatic rationality, FAE=fundamental attribution error, HE= halo effect, REC=recency

Upon examination of Table 4 and Figure 7, the following are true for each teacher individually:

- Teacher 1 was more affected by student A, exhibited intuition and affect for both students
- Teacher 2 employed different heuristics for the two students: more intuitive and affect for Student A versus algorithmic for Student B

- Teacher 3 was relatively similar in process across students, but was more affected by Student B
- Teacher 4 was relatively similar in process across students

Collectively, the teachers were more affected by Student A–the student who is an English language learner, works, helps with siblings after school, and has a 54% overall. Teachers also relied on the recency heuristic more for Student A than for Student B. For Student B, teachers employed more algorithmic thinking and halo effect and committed the fundamental attribution error (which was not exhibited at all during the grade-determination process for Student A).

Overarching Cognitive Processes

In addition to coding, more specifically, for individual heuristics, the teachers' cognitive processes were examined with chronological and systems lenses. This included identifying when in the grade-determination process each teacher came to their grade determination. It also included examining their cognitive process when determining the grades by System 1 versus System 2 thinking. In other words, were the teachers engaging in fast thought processes with less analysis and/or engaging in processes driven by intuition, emotion, or heuristics (System 1 thinking); or were they engaging in more time-intensive, analytical processes that were more logical and rule-governed (System 2 thinking)?

Table 6, below, showcases where in each teacher's cognitive process they identified the grade that they would assign each student. The location in the process is noted for each teacher for both Student A and Student B as the two processes were done distinctly and had the potential for grade determination to occur in different locales.

Table 6

When Teachers Identified the Grade in their Cognitive Process

	Teacher 1 (9 years)	**Teacher 2 (22 years)**	**Teacher 3 (12-15 years)**	**Teacher 4 (5-7 years)**
Student A	at end of cognitive process	at end of cognitive process	at beginning of cognitive process	at beginning of cognitive process
Student B	at end of cognitive process	at end of cognitive process	at beginning of cognitive process	at beginning of cognitive process

As Table 6 indicates, all four teachers were consistent across Students A and B in where they identified the grade they would assign in their individual cognitive process. In other words, while Teacher 1 and Teacher 3 did not assign the grade at the same location in their cognitive processes, each individual teacher assigned the grade in the same location for both Student A and Student B. For example, Teacher 1 assigned the grade at the end of her cognitive process for both Student A and Student B. The teachers were evenly split in where they identified the grade that they would assign each student: Teachers 1 and 2 both assigned the grade at the end of their cognitive processes while Teachers 3 and 4 did so at the beginning of their cognitive processes. No teachers exhibited coming to a grade determination mid cognitive process, nor did any teachers begin with one grade determination and change the grade throughout or by the end of the cognitive process. Teachers 1 and 2, the teachers who came to the grade determination at the end of their processes, were also the two teachers who exhibited the most heuristic thinking (as shown in Tables 4 and 5). Teachers 3 and 4, the teachers who identified the grade at the beginning of their processes, were the only two teachers who explicitly articulated confidence in

their grading process/assigning during the interview with Teacher 3, in reference to having taught for twelve years now, stating, “no wonder I feel confident now.” Similarly, Teacher 4 noted “my tests are based on proficiency…in the last four years, I have been practicing that a lot, so I’m very, very confident.”

In addition to where the teachers identified the grade in their cognitive processes, the overarching cognitive processes were coded for System 1 versus System 2 thinking. The results are collated and presented below in Table 7. The results are again sorted by teacher and by student so as to easily discern any consistencies and inconsistencies that arose between students and across teachers. All teachers did exhibit both System 1 and System 2 thinking at some point in their grade-determination processes, but not all did so for both students and to similar degrees. Table 7 notes in what chronology and to what degree System 1 and System 2 thinking manifested for each teacher by student.

Table 7

System 1 and System 2 Thinking Chronology

	Teacher 1 (9 years)	Teacher 2 (22 years)	Teacher 3 (12-15 years)	Teacher 4 (5-7 years)
Student A				
Initial System	System 1 and System 2 even	System 1 dominant	System 1	System 1
Secondary System (at end, chronologically)	System 2 dominant	System 1 dominant, some System 2	System 2 dominant, some System 1	brief System 2
Grade ID Location	at end	at end	at beginning	at beginning
Student B				
Initial System	System 1 dominant	System 2 dominant	System 1	System 1
Secondary System (at end, chronologically)	System 1 dominant	System 2 dominant	System 2 dominant, some System 1	brief System 2
Grade ID Location	at end	at end	at beginning	at beginning

Table 7 indicates that two out of the four teachers were consistent in their cognitive processes. Teachers 3 and 4 were both consistent individually–engaging in similar thinking for both students– and they also had similarities with one another in how they processed: both began with a quick, System 1 grade determination and followed by additional System 1 or 2 thinking to validate their grade determination. Two teachers, Teacher 1 and Teacher 2, had more complex cognitive processes for Student A–involving more System 1 and System 2 thinking throughout

their grade-determination processes. And, while they were not the same system, those same two teachers, Teacher 1 and Teacher 2, had more streamlined, single system processes for Student B– Teacher 1 relying on System 1 and Teacher 2 relying on System 2.

Emergent Themes

In addition to investigating the cognitive processes of teachers via heuristics and systems lenses, the thematic lens is additionally revelatory. The interviews were coded for topics that were explicitly given in the vignettes, such as ELL status, to see the extent to which such topics influenced teachers' grade-determination processes. Furthermore, throughout their processes, additional themes and thematic patterns arose that were coded, such as teachers trying to identify the cause of student performance.

Topics/Themes from Vignettes

The vignettes for Student A and Student B contained information for the teachers to potentially consider in their grade-determination processes. The following information contained in the vignettes was coded for whether it was considered or weighed in the grade-determination process for each teacher: ELL status, completion percentage, recent score, background/student home life, student GPA, attendance, average test score, overall class percentage, and participation. If a teacher did consider the topic from the vignette, it is indicated with a Y in Table 8. Table 8 is also organized by how many teachers relied on the topics listed: the top of the table being those most frequently present and the bottom being those least frequent.

Table 8

Topical Surfacings in Cognitive Processing

Code from Vignette	Teacher 1 (9 years)	Teacher 2 (22 years)	Teacher 3 (12-15 years)	Teacher 4 (5-7 years)
ELL	Y	Y	Y	Y
completion	Y	Y	Y	Y
recent score	Y	Y	Y	Y
background, home life	Y	Y	Y	Y
GPA	Y	Y	Y	
attendance	Y		Y	Y
average test score	Y	Y		Y
overall percentage	Y		Y	
participation	Y			

Key: Y= yes, topic/theme surfaced
Yellow fill= surfaced for all 4 participants
Blue fill= surfaced for 3 participants
Green fill= surfaced for 2 participants

Of the information given in the vignettes, all teachers considered to some degree ELL status, completion rates, recent scores, and student background or home life in their grade-determination processes. Three out of four teachers weighed student GPA, attendance, and average test score in their cognitive processes. Two out of four of the teachers considered the student's overall percentage in their process, while only one teacher weighed student participation.

Emergent Themes outside of Vignette Verbiage

Topical surfacings noted in the previous section were those given to the teacher

interviewees in the verbiage of the vignettes or interview questions. However, throughout their processing of the information and determining grade assignments, additional themes emerged. These themes are presented below in Table 9, again organized by frequency with the top of the table being those most frequently present and the bottom being those least frequent. The themes included in Table 9 that only surfaced in one teacher interview were included due to their relevance to this study's foci and/or implications for future research.

Table 9

Thematic Surfacings in Cognitive Processing

Theme	Teacher 1 (9 years)	Teacher 2 (22 years)	Teacher 3 (12-15 years)	Teacher 4 (5-7 years)
why/cause	Y	Y	Y	Y
accuracy	Y	Y	Y	Y
judge student, assumptions	Y	Y	Y	Y
understanding/ knowledge (cognition)	Y		Y	Y
hard/ complicated	Y	Y	Y	
teacher supports		Y	Y	Y
inequity		Y	Y	Y
questions	Y	Y		
work with student	Y	Y		
behavior	Y		Y	
growth over time		Y	Y	
grades malleable		Y	Y	
fairness		Y	Y	
more detail		Y	Y	
pressure/ anxiety			Y	Y
confident			Y	Y
effort	Y		Y	

Theme	Teacher 1 (9 years)	Teacher 2 (22 years)	Teacher 3 (12-15 years)	Teacher 4 (5-7 years)
cumulative			Y	Y
feel badly	Y			
teacher blame		Y		
access		Y		
quick process		Y		
systemic issues		Y		
bump to pass			Y	
previous grade				Y
college	Y			

Key: Y= yes, topic/theme surfaced
Yellow fill= surfaced for all 4 participants
Blue fill= surfaced for 3 participants
Green fill= surfaced for 2 participants

At some point in the interview, all four teachers tried to identify the cause of some performance or questioned why something was occurring within the vignette. Similarly, all teachers noted or questioned the accuracy of data points and made judgements or assumptions about students beyond the scope of what was included in the vignette. This included assumptions about struggle with curriculum. about economic privilege, etc. For example, Teacher 2 articulated that Student B "breezed through school" and Teacher 1 referenced Student A as having a "very privileged background." Three teachers out of four discussed student understanding or knowledge in their processes and discussed teacher supports–either that they wanted to know if they had been implemented or suggested that they would implement them moving forward in giving opportunities for students to show increased proficiency. Three teachers also articulated how hard or complicated the grading process is and noted issues of inequity or fairness. Two out of

the four teachers posited questions in their grade-determination process and noted that they would work with students. Two also weighed behavior, growth, and effort in their grading decision-making processes. Two teachers specifically noted the malleability of grades and discussed issues of fairness. Two teachers also wanted more detail than that given in the vignettes, two alluded to the pressures and anxiety surrounding grading practices, and two referenced having confidence in their current grading practices. Of the four teachers interviewed, one teacher specifically discussed bumping up grades below the traditional scale cut-offs to pass students, but not in other circumstances. Emotional responses surfaced slightly differently for teachers, with one teacher indicating feeling badly for assigning a grade they ultimately determined and another teacher placing self-blame in their grade-determination process when analyzing the data provided. One teacher noted or considered a student's possible future plans to attend college in their decision-making and another teacher suggested there are systemic limitations on grading equitably. Finally, one teacher specifically cited that in real life and in real time the grade-determination process, including possible grade changing, is a very fast process–much faster than that for the student vignettes in the interview.

Patterns and Associations

There were notable correlations in the study between teacher tenure and cognitive processes, with the more tenured teacher exhibiting less consistency and more heuristics in her process and the least tenured teacher exhibiting the most consistency and the least heuristic thinking in his. Teacher 2, the most tenured teacher, showed the least consistency–both with which data point she deferred to (deferring to different data points in order to substantiate higher grade assignments) and with System 1 versus System 2 thinking processes (having different processes for each student). Teacher 2, having taught 22 years, also heavily relied on heuristics

in her grade-determination process when compared to the more novice Teacher 4 of 5 years. Teacher 4, the least tenured teacher in the study, showed the least heuristic thinking, was consistent in his System 1 and System 2 thinking chronology and degree, and showed deference to the most recent score for both students. It is also of note that Teacher 4, the teacher who was the most consistent in process and exhibited the least heuristic thinking, was also the only male in the study. The two teachers with the most experience, Teacher 2 of 22 years and Teacher 3 of 12-15 years, also were the two most candid in the reflective Part 2 of the interview and the only two who explicitly commented on often taboo and obfuscated grading practices such as the malleability of grades, their bumping up in certain circumstances, the speed at which this process is done in real time, and the inequities that result.

Teachers 3 and 4 had much in common, as they tangibly both relied on recency and, in doing so, bumped up Student A and did not bump up the grade for Student B. They also were both consistent in their overarching cognitive processes–both identifying the grade assignment at the beginning of their cognitive processes for both students and reinforcing their grade determinations with System 2 thinking to validate their grade assignments. These two teachers, Teachers 3 and 4, were also the only two teachers who relayed confidence in their grading practices overall in the reflective section of the interview. And while the two teachers who identified the grade at the beginning of their processes did range in terms of how heavily each relied on heuristic thinking (Teacher 3 relying much more on heuristics than Teacher 4), both teachers who identified their grade assignments for the students at the end of their cognitive processes (Teachers 1 and 2) exhibited notable heuristic thinking throughout their processes.

Summary of Findings

The study analyzed the cognitive processes of four teachers as they sought to come to

grade determinations for two different hypothetical students–both borderline students with complex profiles. Student A was an English Language Learner with a more complex familial and socioeconomic profile and had an overall class percentage of 54%, which in traditional grading scales would be assigned an F. However, three of the four teachers interviewed bumped the grade to a D of some kind. For Student A, there was also significantly more affect heuristic coded, (9 instances for Student A versus 5 for Student B), and more reliance on recency (5 for Student A versus 3 for Student B). This may indicate that teachers were more affected by Student A and that teachers are more prone to weigh factors and engage in processes that benefit students, since the recent score for Student A was favorable. And while the overarching cognitive processes were similar in complexity for Teachers 3 and 4, there were more complex processes system-wise evidenced for Student A by both Teachers 1 and 2. This may indicate that teachers engage in more complex cognitive processes when presented with students with more complex background information (as Student A has more factors outside of the classroom to consider than Student B), or it may indicate that teachers engage in more complex cognitive processes when weighing whether or not to assign a failing grade to a student.

Student B was presented as a straight-A student with an overall class percentage of an 89.3%, which under traditional grading scales would be a B+. In contrast to Student A, only 1 teacher bumped the grade for Student B. Upon reflection, three teachers indicated that Student B was easier to assign a grade than Student A, reinforcing that teachers may struggle more with lower marks and with potentially failing students than they do with the broader question of whether to round up any grade. For Student B, more algorithmic thinking was coded (8 for Student B versus 2 for Student A), more halo effect was coded (5 instances for Student B versus 2 for Student A), and more fundamental attribution error was coded (3 for Student B versus 0 for

Student A). This may indicate that for students with less complex background information and/or students who are borderline grades with higher overall percentages, teachers rely on numerical thinking with more self-reported ease, but also are more prone to bias in the form of allowing what they deem negative behaviors (such as distracting others and pleading before grades were due) to influence decision-making.

In their grade determination cognitive processes, all teachers exhibited at some point heuristic thinking, System 1 thinking, and System 2 thinking. The most common heuristics in total were algorithmic thinking, evidenced by all teachers and coded a total of ten times; affect heuristic, evidenced by three of four teachers and coded a total of fifteen times; and the recency heuristic, evidenced by three of four teachers and coded a total of eight times. This showcases that, while grade determination may be a numbers game, it is much more than that to teachers: all exhibited emotional responses throughout the process. Furthering the notion that the grading process is more than a logical, numerical feat, all teachers topically weighed ELL status, student background or home life, and sought to understand the cause of student performance. All teachers also at some point made assumptions or judgements about students beyond the scope of the information given, and three out of the four teachers noted how difficult or complicated the grading process is. And thematically, teachers noted struggles such as fairness, systemic issues, equity, pressure and anxiety, and self-blame.

Even though teachers iterated that grading is complicated, most individual teachers exhibited consistency of process: three out of four were consistent in terms of deference data point, all teachers were relatively even in the number of times they committed heuristics for each student in total (not specific heuristics), and all teachers were consistent in where they identified the letter grade in their cognitive process for students. However, teachers were not individually

consistent with which heuristics they utilized from student to student and there were disparate cognitive processes evidenced across teachers. When teachers identified the grade in their cognitive process was split, with half of the teachers identifying the grade at the beginning of their cognitive process and the other half identifying the grade at the end of their cognitive process. No teachers identified the grade in any other location other than the beginning or end of the cognitive process, and no teachers changed the grade once they had articulated their letter grade assignment. There was a tenuous correlation found between tenure and cognitive processes, with the most senior teacher exhibiting a great deal of heuristic thinking, and having the most flexible, least consistent process of the teachers interviewed; and the most novice teacher exhibiting little heuristic thinking, and exhibited the most consistent and least complex process of the teachers interviewed. The most novice teacher was also the only male in the study, so further research would be needed to strengthen the tenure correlation found. The types and frequency of heuristics committed by teachers in their grade-determination processes also varied widely: two teachers exhibited 6 of the 7 coded heuristics, while one teacher only exhibited 2 of the seven; and one teacher committed heuristic thinking 24 times, while another teacher interviewee only did 3 times in total.

Overall, this reveals that teachers are generally consistent in their overarching cognitive processing across students, but differ in which heuristics they may commit from student to student. Additionally, while all teachers were shown to have engaged in heuristic thinking and System 1 and System 2 thinking, teachers vary greatly from one another in the complexity of their cognitive processes and the extent to which they rely on heuristics to determine grades. Furthermore, as they progress in the profession, teachers seem to become more flexible in their grade determination cognitive processes and become more candid about the emotional tolls and

inequities of current grading practices.

Chapter 5: Conclusions

This study sought to fill a gap in the existing literature surrounding the cognitive process in which teachers engage as they are assigning culminating grades at the end of a term. This study was catalyzed by administrative prompting at Oak High School at the end of each term to ensure that students' grades were indeed accurate; teachers were instructed to adjust any inaccuracies, if needed. The hypothetical student vignettes presented to teachers aimed to mimic actual student profiles that teachers may encounter that might prompt an adjustment at the end of such a term. The vignettes were crafted to discern whether teachers engage in similar processes for students of varying backgrounds and students in differing academic scenarios. Both vignettes presented, despite being different in many facets, were crafted to prompt teachers to engage in thinking surrounding whether or not they would bump a grade that may be considered a cusp grade. This study ultimately answered the following research questions:

- Question 1: How do secondary teachers make final decisions about students' summative grades?
- Question 2: What cognitive biases do they rely on to make grade decisions and, if relevant, grade changes?

Discussion of Findings

Major findings suggest that all teachers may engage in complex cognitive processes including System 1 and System 2 thinking and heuristic thinking when determining ultimate grades. However, some teachers rely heavily on heuristics, are more complex, or are more or less consistent in their cognitive processes from student to student than others. Teachers seem to become more candid and more flexible as they gain years of service in the profession, and there is evidence of the concurrent dominance of algorithmic thinking and the affect heuristic–

reinforcing the complexity of the grading process. In juxtaposing the teachers' approaches to the two students, findings suggest that there may be a correlation between complex student background information and/or failing grades and higher complexity of teacher cognitive process in grade determination and more affect. Less complex student background and/or higher grades correlated with more reported ease, more numerical thinking, and student behavior/character influencing the decision-making process.

Comparisons and Connections

In addition to fulfilling the primary research questions identified, this study also substantiates, furthers, and challenges some of the existing literature. One of the significant themes in the existing literature surrounding the teacher-grading process and evaluation is that of rapidity and cognitive biases. Klein (2005) found that many of the cognitive biases in medical contexts are evidenced due to the rapidity of the decision-making process. In other words, when under time constraints, Klein found that biases were more prevalent. While such rapidity of process in real practice was noted by Teacher 2, teachers in the interview process of this study were not rushed to make their grade determinations, yet still exhibited heuristic thinking and bias.

The findings of this study also pertain to configurational assessment, stemming from Kaplan (1973) and Sadler (1989) as cited in Crisp (2010). Such assessment occurs when judges first make an overall impressionistic assessment and then substantiate this assessment with criteria as they assess (Crisp, 2010). Half of the teachers in this study conformed to such a process and did identify their holistic impressionistic assessment of a grade at the beginning of their process. Half of the teachers, however, did not. Crisp (2010) also noted that heuristics are more likely when judgements are difficult. Remember, heuristics are "efficient cognitive

processes, conscious or unconscious, that ignore part of the information" (Gigerenzer & Gaissmaier, 2011)**.** Teachers in this study dominantly reported that Student A was more difficult to assign a grade, but that did not coincide with more presence of heuristic thinking (see Figure 7); the heuristics that were present were simply different ones in the more difficult case.

More specifically, Gweon et al. (2017) and Wieman and Welsh (2016) found that instructors are prone to committing specific heuristics: the halo effect and the fundamental attribution error. Both studies were done at the university level with math, science, and engineering professors participating. This study furthers these findings as one teacher did commit the fundamental attribution error in her grade-determination process and two out of four committed the halo effect. While they were found to be existing, it was expected that the two would be more present and prevalent than they were given the existing research and given that there was not an explicit assessment tool given to the teachers to counter such cognitive biases as Gweon et al. (2017) recommends. This may, however, be as a result of the hypothetical nature of the vignettes presented to teachers; it is possible that with real students teachers may be more prone to committing the fundamental attribution error and halo effect as a result of authentic and sustained connections.

The concepts of "no fail" policies and tendencies amongst teachers was also present in the literature and relevant in this study's findings. Halligan (2011) explored "no fail" and "social promotion" policies in both the United States and Thailand and found that Thai teachers were more likely than their US counterparts to favor changing students' grades in order to help low achieving students. Similarly, Couper (2018) sampled 390 nursing faculty and found that 17.4 percent of nursing faculty did not assign an earned failing grade. While small, this study challenges these findings–as three out of four teachers (75% of teachers) in this study assigned a

student with a 54% a passing grade. Hence, this study indicates that "failure to fail" may be more prevalent than previous research suggests.

Thematically, the issues of malleability of grades and fairness arose in this study–just as they did in the literature. Green (2019) relayed that "many of our academic… standards include evaluative grey areas at their edges, and our judgements of competence and achievement are nuanced" (p. 323). Valentine et al. (2021) expands upon such nuance when they argue that shifting the focus to what is fair rather than objective "allows for the embracing of many different perspectives, and the legitimising of human judgement in assessment" (p. 714). The complexity of cognitive processes found in this study, as well as inconsistencies of cognitive processes and grade assignments, reinforce the lack of objectivity and concurrent nuanced nature of student evaluation.

Such nuance and lack of objectivity is addressed by McNair's (1978-1979) finding that teachers are perhaps most affected by concern for their students. This was validated by this study's findings as affect was one of the most prevalent heuristics noted. That being said, teachers were more affected by the student with the more complex home life/background, so this concern may not extend to all students evenly. DeLuca et al. (2018) surveyed 404 teachers from North America and found that nuanced approaches to fairness were not consistent amongst all teachers, but were linked to tenure, early career teachers being more standardized in approach and more tenured teachers being more differentiated. This was also noted in this study: the most tenured teacher of 22 years was the teacher who was the most flexible in her deference data points and most complex in her overarching cognitive processes while the most novice teacher was the most consistent and simplistic in his process and exhibited the least heuristic thinking.

Implications and Recommendations for Practice

This study helps to validate that the grading process is multifaceted, complex, and time-intensive for most teachers, especially when considering changing a grade. As such, more new teacher training in teacher preparatory programs may be warranted and continued professional development over time would be beneficial. Such training may illuminate the complexities and heuristics that surface for teachers and validate teachers' processes, but also allow teachers to be more cognizant of potential cognitive biases that may surface in their grade-determination processes. When given the time, the grade-determination process for each student was time intensive. Given that teachers are often tasked with affirming or assigning grades for more than one hundred students in a given grading period, more time is needed for grading to be meaningful, consistent, and equitable.

Additionally, due to the complex nature of the grading process, administrators should be cautious of oversimplifying or dictating processes without discussion or consideration of staff cognitive processes. Likewise, administrators should not assume that because their staff adheres to the same overarching model of student evaluation/grading (such as proficiency) that all staff have similar or the same cognitive processes when ultimately determining student grades, nor should they make the assumption that because their staff may adhere to the same overarching model of evaluation/grading that they would assign students the same grade. Teachers' cognitive processes in determining grades were found to be different, resulting in different grade outcomes for students and inconsistency. There was, specifically, variation noted related to years in the profession. These findings necessitate more collective conversations and collaboration amongst stake-holders. Vignettes similar to those used in this study could be utilized in professional development and collaborative conversations in order to ground discussions and unveil teacher

cognitive processes and biases. Placing staff in small groups with teacher participants of varying years of service; presenting them with specific, hypothetical vignettes such as those utilized in this study; and then directing them to think aloud their process in determining how and why they would ultimately assign a grade for each student vignette could result in more staff clarity, consistency, and more equitable student outcomes.

Limitations and Recommendations for Further Research

The findings of this study are limited in generalizability due to the scope of the sample population. Given that this study only interviewed four teachers, a larger study with more teachers is needed in order to validate this study's findings. More specifically, teachers of varying years of service and gender are needed in addition to teachers from varying school sites. This study interviewed teachers all from the same school site, so broadening to different sites or different levels is needed. Furthermore, this study preliminarily indicates a correlation between complexity of student background and/or failing grades and an increased complexity in cognitive process. Additionally, this study preliminarily suggests a correlation between less complex student background and/or higher scores with ease of process, a reliance on numerical thinking, and increased consideration of student behavior/character in the decision-making process. However, not enough vignettes were administered in order to determine true correlation or causation. Further studies would be needed to control for variables. Furthermore, teachers in this study articulated wanting more information than was presented in the vignettes, so further studies with more complex student profiles would add to the body of research.

Conclusion

This study delved deeply into the cognitive processes of teachers when undertaking one of the most difficult facets of their job: grading. More specifically, this study examined teachers'

cognitive processes when trying to determine letter grade assignments in complex, cusp grading circumstances. Overall, the study affirmed my hypothesis that teachers vary in how they approach and make such decisions. However, the extent to which teachers varied was unexpected with some teachers relying very heavily on heuristic thinking to guide them to an ultimate grade determination with others determining the grade at the forefront of their cognitive process and relying very little on heuristics. While grading, given that teachers are humans assessing humans, will always be complex; removing the veil, researching, and discussing such complexities can hopefully lead to more educational equity.

References

Annerstedt, C., & Larsson, S. (2010). 'I have my own picture of what the demands are ... ': Grading in Swedish PEH — problems of validity, comparability and fairness. European Physical Education Review, 16(2), 97–115. https://doi.org/10.1177/1356336X10381299

Atkinson, M. P., Wills, J. B., & McClure, A. I. (2008). The Evidence Matrix: A Simple Heuristic for Analyzing and Integrating Evidence. Teaching Sociology, 36(3), 262–271. https://doi.org/10.1177/0092055X0803600306

Banta, T. W. (2007). Can assessment for accountability complement assessment for improvement? Peer Review : Emerging Trends and Key Debates in Undergraduate Education, 9(2), 9. https://alliance-primo.hosted.exlibrisgroup.com/permalink/f/1160sj7/TN_cdi_proquest_journals_216592048

Blanchard, S., & Muller, C. (2015). Gatekeepers of the American Dream: How teachers' perceptions shape the academic outcomes of immigrant and language-minority students. Social Science Research, 51, 262–275. https://doi.org/10.1016/j.ssresearch.2014.10.003

Bourke, R., O'Neill, J., & Loveridge, J. (2018). What starts to happen to assessment when teachers learn about their children's informal learning? Australian Educational Researcher, 45(1), 33–50. https://doi.org/10.1007/s13384-018-0259-x

Bowers, A. J. (2011). What's in a grade? The multidimensional nature of what teacher-assigned grades assess in high school. Educational Research and Evaluation, 17(3), 141–159. https://doi.org/10.1080/13803611.2011.597112

Brookhart, S., 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

Cain, T., Brindley, S., Brown, C., Jones, G., & Riga, F. (2019). Bounded decision-making, teachers' reflection and organisational learning: How research can inform teachers and teaching. British Educational Research Journal, 45(5), 1072–1087. https://doi.org/10.1002/berj.3551

Chen, J.-F., Hsieh, H.-N., & Do, Q. (2014). Predicting Student Academic Performance: A Comparison of Two Meta-Heuristic Algorithms Inspired by Cuckoo Birds for Training Neural Networks. Algorithms, 7(4), 538–553. https://doi.org/10.3390/a7040538

Ciobanu, M. & Popovici-Bujor, V. (2020). Genesis of the Concept of Algorithmic Thinking and the Nature of Mathematics. Akademos (Chișinău, Moldova), Nr. 3(58), 28–37. https://doi.org/10.5281/zenodo.4269379

Couper, J. (2018). The Struggle Is Real: Investigating the Challenge of Assigning a Failing Clinical Grade. Nursing Education Perspectives, 39(3), 132–138. https://doi.org/10.1097/01.NEP.0000000000000295

Crisp, V. (2010). Towards a model of the judgement processes involved in examination marking. Oxford Review of Education, 36(1), 1–21. https://doi.org/10.1080/03054980903454181

DeLuca, C., Valiquette, A., Coombs, A., LaPointe-McEwan, D., & Luhanga, U. (2018). Teachers' approaches to classroom assessment: a large-scale survey. Assessment in Education: Principles, Policy & Practice, 25(4), 355–375. https://doi.org/10.1080/0969594X.2016.1244514

Feufel M, & Flach J. (2019). Medical Education should 'embrace' rather than overcome heuristics. Medical Education 53 (4): 334 – 44.

Gigerenzer, Gerd. (2008). Why Heuristics Work. Perspectives on Psychological Science, 3(1), 20–29. https://doi.org/10.1111/j.1745-6916.2008.00058.x

Gigerenzer, Gerd. (2020). How to Explain Behavior? Topics in Cognitive Science, 12(4), 1363–1381. https://doi.org/10.1111/tops.12480

Gigerenzer, Gerd. (2021). Axiomatic rationality and ecological rationality. Synthese (Dordrecht), 198(4), 3547–18. https://doi.org/10.1007/s11229-019-02296-5

Gigerenzer, G., & Brighton, H. (2009). Homo Heuristics: Why Biased Minds Make Better Inferences. Topics in Cognitive Science, 1(1), 107–143. https://doi.org/10.1111/j.1756-8765.2008.01006.x

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic Decision Making. Annual Review of Psychology, 62(1), 451–482. https://doi.org/10.1146/annurev-psych-120709-145346

Green, E. P. (2019). Refining heuristics for educators. Medical Education, 53(4), 322–324. https://doi.org/10.1111/medu.13807

Guskey, T. R., & Link, L. J. (2019). Exploring the factors teachers consider in determining students' grades. Assessment in Education: Principles, Policy & Practice, 26(3), 303–320. https://doi.org/10.1080/0969594X.2018.1555515

Gweon, G., Jun, S., Finger, S., & Rosé, C. P. (2017). Towards effective group work assessment: even what you don’t see can bias you. International Journal of Technology and Design Education, 27(1), 165–180. https://doi.org/10.1007/s10798-015-9332-1

Halligan, K. M. (2011). An Examination of the No Fail Policy in Thailand and the Effect on Community Relations. Language Testing in Asia, 1(2), 1–12. https://doi.org/10.1186/2229-0443-1-2-79

Hopfenbeck, T. N. (2019). Assessment reforms and grading. Assessment in Education: Principles, Policy & Practice, 26(3), 255–258. https://doi.org/10.1080/0969594X.2019.1625510

Howell, J. L., & Shepperd, J. A. (2011). Demonstrating the Correspondence Bias. Teaching of Psychology, 38(4), 243–246. https://doi.org/10.1177/0098628311421320

Hughes, J. N, Gleason, K. A., & Zhang, D. (2005). Relationship influences on teachers' perceptions of academic competence in academically at-risk minority and majority first grade students. Journal of School Psychology, 43(4), 303–320.

Irizarry, Y. (2015). Selling students short: Racial differences in teachers' evaluations of high, average, and low performing students. Social Science Research, 52, 522–538. https://doi.org/10.1016/j.ssresearch.2015.04.002

Jewell, R. T., & McPherson, M. A. (2012). Instructor-Specific Grade Inflation: Incentives, Gender, and Ethnicity. Social Science Quarterly, 93(1), 95–109. https://doi.org/10.1111/j.1540-6237.2011.00827.x

Kahneman, D. (2003). A Perspective on Judgment and Choice. The American Psychologist, 58(9), 697–720. https://doi.org/10.1037/0003-066X.58.9.697

Kahneman, D., & Klein, G. (2009). Conditions for Intuitive Expertise. The American Psychologist, 64(6), 515–526. https://doi.org/10.1037/a0016755

Klein, J. G. (2005). Five pitfalls in decisions about diagnosis and prescribing. BMJ, 330(7494), 781–783. https://doi.org/10.1136/bmj.330.7494.781

Klein, G., Shneiderman, B., Hoffman, R. R., & Ford, K. M. (2017). Why Expertise Matters: A Response to the Challenges. IEEE Intelligent Systems, 32(6), 67–73. https://doi.org/10.1109/MIS.2017.4531230

Kwok, C. (2020). Work autonomy and workplace democracy: the polarization of the goods of work autonomy in the two worlds of work. Review of Social Economy, 78(3), 351–372. https://doi.org/10.1080/00346764.2019.1690671

Lee, J. C.-M., Pong, T.-C., & Esterline, A. (1998). Enhancing Object Recognition Using Recency and Co-occurrence Heuristic. Pattern Recognition, 31(9), 1319–1336. https://doi.org/10.1016/S0031-3203(97)00163-5

Morewedge, C. K., & Kahneman, D. (2010). Associative processes in intuitive judgment. Trends in Cognitive Sciences, 14(10), 435–440. https://doi.org/10.1016/j.tics.2010.07.004

Nisbett, R. E., Pens, K., Incheol C., & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. Psychological Review, 108(2), 291–310. https://doi.org/10.1037//0033-295X.108.2.291

Oregon Department of Education. (2020). *2019-2020 Adapted At-a-Glance Profile: West Linn-Wilsonville School District*. Retrieved from https://www.oregon.gov/ode/schools-and-districts/reportcards/reportcards/Documents/Adapted1920/1920-AAAG-1922.pdf

Pattison, E., Grodsky, E., & Muller, C. (2013). Is the Sky Falling? Grade Inflation and the Signaling Power of Grades. Educational Researcher, 42(5), 259–265. https://doi.org/10.3102/0013189X13481382

Popham, W. J. (2018). Formal Evaluation Versus Seat-of-the-Jeans Thinking. New Directions for Evaluation, 2018(157), 39–41. https://doi.org/10.1002/ev.20293

Richards, J. (2014). An old chestnut revisited: Teachers' opinions and attitudes toward grading within a competency based training framework. *International Journal of Training Research, 12*(3), 182-191. https://georgefox.idm.oclc.org/login?url=https://www-

proquest-com.georgefox.idm.oclc.org/scholarly-journals/old-chestnut-revisited-teachers-opinions/docview/1642597230/se-2?accountid=11085

Riley, T., & Ungerleider, C. (2019). Imputed Meaning: An Exploration of How Teachers Interpret Grades. Action in Teacher Education, 41(3), 212–228. https://doi.org/10.1080/01626620.2019.1574246

Salokangas, M., Wermke, W., & Harvey, G. (2020). Teachers' autonomy deconstructed: Irish and Finnish teachers' perceptions of decision-making and control. European Educational Research Journal EERJ, 19(4), 329–350. https://doi.org/10.1177/1474904119868378

Scott, J., & Nichols, T. P. (2017). Learning analytics as assemblage: Criticality and contingency in online education. Research in Education (Manchester), 98(1), 83–105. https://doi.org/10.1177/0034523717723391

Shavelson, R. J., & Stern, P. (1981). Research on Teachers' Pedagogical Thoughts, Judgments, Decisions, and Behavior. Review of Educational Research, 51(4), 455–498. https://doi.org/10.3102/00346543051004455

Sheppard, M., & Levy, S. A. (2019). Emotions and teacher decision-making: An analysis of social studies teachers' perspectives. Teaching and Teacher Education, 77, 193–203. https://doi.org/10.1016/j.tate.2018.09.010

Singleton Jr, R., & Smith, E. R. (1978). Does Grade Inflation Decrease the Reliability of Grades? Journal of Educational Measurement, 15(1), 37–41. https://doi.org/10.1111/j.1745-3984.1978.tb00054.x

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? The Behavioral and Brain Sciences, 23(5), 645–665. https://doi.org/10.1017/S0140525X00003435

Stitt, J. L., & Pula, J. J. (2014). Voting for Subjectivity: Adding Some Gray Areas to Black-and-White, Objective Grading Practices. *Delta Kappa Gamma Bulletin, 80*(3), 24-27. https://georgefox.idm.oclc.org/login?url=https://www-proquest-com.georgefox.idm.oclc.org/scholarly-journals/voting-subjectivity-adding-some-gray-areas-black/docview/1509200297/se-2?accountid=11085

Stiggins, R. (2014). Improve assessment literacy outside of schools too: teaching and assessment have become separated, which has kept teachers from developing the assessment skills they need to truly enhance learning. Phi Delta Kappan, 96(2), 67–67. https://alliance-primo.hosted.exlibrisgroup.com/permalink/f/1160sj7/TN_cdi_gale_infotraccpiq_384544243

Suskie, L. (2007). Some Thoughts and Suggestions on Assessing Student Learning. PS, Political Science & Politics, 40(1), 102–102. https://doi.org/10.1017/S1049096507250164

Thorsen, C., & Cliffordson, C. (2012). Teachers' grade assignment and the predictive validity of criterion-referenced grades. Educational Research and Evaluation, 18(2), 153–172. https://doi.org/10.1080/13803611.2012.659929

Valentine, N., Durning, S., Shanahan, E. M., & Schuwirth, L. (2021). Fairness in human judgement in assessment: a hermeneutic literature review and conceptual framework. Advances in Health Sciences Education: Theory and Practice, 26(2), 713–738. https://doi.org/10.1007/s10459-020-10002-1

van Ewijk, Reyn. (2011). Same work, lower grade? Student ethnicity and teachers' subjective assessments. Economics of Education Review, 30(5), 1045–1058. https://doi.org/10.1016/j.econedurev.2011.05.008

Westphal, A., Lazarides, R., & Vock, M. (2020). Are some students graded more appropriately than others? Student characteristics as moderators of the relationships between teacher-assigned grades and test scores in mathematics. British Journal of Educational Psychology, e12397–e12397. https://doi.org/10.1111/bjep.12397

Westphal, A., Becker, M., Vock, M., Maaz, K., Neumann, M., & McElvany, N. (2016). The link between teacher-assigned grades and classroom socioeconomic composition: The role of classroom behavior, motivation, and teacher characteristics. Contemporary Educational Psychology, 46, 218–227. https://doi.org/10.1016/j.cedpsych.2016.06.004

Wieman, C., & Welsh, A. (2016). The Connection Between Teaching Methods and Attribution Errors. Educational Psychology Review, 28(3), 645–648. https://doi.org/10.1007/s10648-015-9317-3

Zeijlmans, K., López , M., Grietens, H., & Knorth, E. J. (2019). Heuristic decision-making in foster care matching: Evidence from a think-aloud study. Child Abuse & Neglect, 88, 400–411. https://doi.org/10.1016/j.chiabu.2018.12.007

Appendix A: Interview Protocol

Introduction:

Thank you so much for meeting with me today and volunteering your time to participate in my doctoral research. The aim of this study is to gain insight into the cognitive processes of high school teachers as they evaluate and determine a final grade at the end of a term. Grading and assessment have been heavily studied, including what specific elements are weighed in creating grading frameworks, but there is a lack of research into how teachers make final grading decisions. Your participation in this study will help to add to the growing body of research in this much needed area.

Informed Consent:

Before we begin, I want to ensure you that this interview will remain anonymous and only myself and my dissertation chair will have access to the original recordings and transcripts of our interview. A pseudonym is being used for both the school and you. As we are recording, please let me know if you need to pause, stop, or would like anything to be struck from the record. If you have not already done so, please review the informed consent form that was given to you prior to us beginning. Do you have any questions before we jump in and start the interview?

-Begin Recording-

Interview:

The interview will contain two parts: part one will consist of you reviewing and thinking aloud to two hypothetical student vignettes and part two will consist of reflecting upon your thinking in part one via some direct questions.

Part 1: Vignette Think-Aloud:

In a moment, you will be given two hypothetical student vignettes. These represent profiles of students that you may have in your class. I want you to imagine that you are at the end of the grading term and are reviewing this information to determine what will be each students' final letter grade in your course. As you review each student, I would like for you to process all of the information aloud, including reading the vignettes and any and all corresponding thinking as you are working toward trying to decide on their final letter grade. This is meant to be a stream of consciousness exercise, so please do not leave any of your thinking out—even if you do not think that an individual element will be weighed heavily or at all in your decision-making process, please speak aloud this fact and the reasoning. Essentially, anything and everything going on in your head as you process is of merit and of interest here. Do you have any questions before I give you the vignettes?

Great. Let's begin. Here are the two student vignettes. Let's start with Student A. Please begin reading aloud through the information for Student A and thinking aloud your process to determine what final letter grade you would ultimately assign.

Perfect. Thank you. Now, let's move on to Student B. Please begin reading through the information for Student B and thinking aloud your process to determine what letter grade you would ultimately assign.

	Vignette 1	Vignette 2
Name	Student A	Student B
Grade Level	10	11
SPED	NA	NA
ELL	ELL active	NA
Free Reduced Lunch	Yes	No
Extracurriculars	Soccer, theater	Swimming, National Honor Society, ISEF
GPA	3.2	4.4
Class Attendance	75%	95%
Assignment completion rate	51%	85%
Most recent test score	61%	88%
Average test score	45%	92%
Overall class percentage	54%	89.3%
Engagement	Lively class participant Answers questions Participates in group work Positive attitude Comes in for help	Completes all classwork Participates in group work Answers questions when called upon Sometimes distracts others
Other	Works part time Has a single parent Helps watch siblings after school hours	Volunteer tutors in after school program Emailed and came in day before grades are due pleading for an A- You will be their first possible B

Part 2: Reflection:

Thank you for your responses to the vignettes. Now, we are going to move on to part two and reflect upon the vignettes and the process.

1. Please describe how satisfied you are with your letter grade decision for student A. Explain.
2. Please describe how satisfied you are with your letter grade decision for student B. Explain.
3. For which student was it the easiest to assign a letter grade? Why was this student easier?
4. Why was Student ___ more difficult to assign a letter grade? Explain.
5. Is there any other information you wish you had had in order to determine the ultimate letter grade for either/both student(s)? Explain.

Conclusion, Thank You:

Thank you so much for participating in my research. I know that you are busy and have many commitments, and I appreciate you taking the time to speak with me.

Adopted from:

Zeijlmans, K., López, M., Grietens, H., & Knorth, E. J. (2019). Heuristic decision-making in foster care matching: Evidence from a think-aloud study. Child Abuse & Neglect, 88, 400–411. https://doi.org/10.1016/j.chiabu.2018.12.007

Appendix B: Informed Consent Form

RESEARCH SUBJECT INFORMED CONSENT FORM

Prospective Research Subject: Read this consent form carefully and ask as many questions as you like before you decide whether you want to participate in this research study. You are free to ask questions at any time before, during, or after your participation in this research.

Principal Researcher: Bonnie Robbins
Dissertation Chair: Dr. Dane Christian Joseph, Professor of Education, George Fox University

Purpose:
You are being asked to participate in a research study designed to gain insight into the cognitive processes of high school teachers as they assign final letter grades. Ultimately, the study will provide insight into what drives teacher perception and action when it comes to student grades and provide insight into what future trainings and discussions may be necessary.

Procedures:
You will be asked to participate in an interview expected to be approximately one hour in length. We may schedule two shorter interviews if that works better for your schedule. The interview(s) will be semi-structured, while simultaneously being flexible and allowing for follow-up questions and opportunities for expansion. The interview(s) will consist of two parts. Part 1 will be a think aloud design in which you will be asked to think aloud your process in determining two hypothetical students' letter grades at the end of a term. In Part 2, you will be asked five reflective questions. You will be given the opportunity to ask questions throughout.
Your responses during Parts 1 and 2 will be audio recorded using VoiceMemo on an iPhone, and using an Evistr 16GB digital voice recorder. Pseudonyms will be used for both the school site of study and for all participants. Any highly identifiable information that puts the anonymity of subjects at risk will be ommitted from publication. The original recordings and transcripts of the interviews will only be accessible by the researcher, a professional transcriber, and the dissertation chair--all of whom will maintain confidentiality. All recordings and transcripts will be housed in personal drives outside of district/school accounts and outside of the public domain. Once transcribed, the original audio recordings will be deleted from all devices and drives.

Possible Risks, Benefits, and Confidentiality:
Risks in this study are low. Confidentiality and anonymity will be prioritized throughout the case study. Confidentiality will be achieved by using pseudonyms and prioritizing the explicit anonymity of all participants, but anonymity may not be completely feasible due to the limited population size. The results of the study, including laboratory or any other data, may be published for educational purposes, but will not give your name or include any identifiable references to you.

There is no financial compensation for participation in this study. You will be helping to contribute to educational research that may positively influence future professional development of teachers and educational stake-holders.

Withdrawal

You are free to choose whether or not to participate in this study and may withdraw at any point. You will be provided with any significant new findings developed during the course of this study that may relate to or influence your willingness to continue participation.

Voluntary Consent

I have read and understand this consent form, and I volunteer to participate in this research study. I understand that I will receive a copy of this form. Please contact either the principal researcher, Bonnie Robbins, or the dissertation chair, Dr. Dane Joseph with any questions.

_____	_____	_____
Participant Name	Participant Signature	Date

Principal Researcher:
Bonnie Robbins
brobbins17@georgefox.edu

Dissertation Chair:
Dr. Dane Christian Joseph
djoseph@georgefox.edu

Appendix C: Coding Key

Color, Code	Heuristic, Cognitive Process	Indicators, Key terms
Blue 1	System 1 thinking	thought process driven by intuition, emotion, heuristics -may be faster process with less analysis -weigh by looking at coding for heuristics and cognitive processes for each vignette -heuristics will lead to System 1 thinking code -weighing of non-numerical, quantifiable factors and considerations
Green 2	System 2 thinking	logical, rule-governed thinking more analytical focus on the data, numbers presented in the vignettes -more time-intensive, thorough, analytical process -weigh by looking at coding for heuristics and cognitive processes for each vignette -algorithmic thinking, axiomatic rationality may lead to System 2 thinking code, as will showing deference to numerical thinking and/or an intensive process to validate reliability of score
Purple ALG	algorithmic thinking	relies on applying an algorithm, numerical principle or rule, in order to come to decision
Pink AX	axiomatic rationality	conforms strictly to axioms or rules (such as when to round up)
red HE	halo effect	impression of student's character influences decision-making -look for descriptions, judgments of personality, character (such as "good kid", "hard worker", etc.) -influenced by behavior of student, relies on behavioral features to influence decision-making
orange FAE	Fundamental attribution error, correspondence bias	attributes performance, scores, etc. to a tangible, identifiable trait (such as ELL status, GPA, etc.)
grey INT	intuitive judgment	relies on past experience, patterns witnessed in order to make decision or justify decision -look for statements like "this type of kid typically", "usually when this happens"
brown AV	availability heuristic	weighs heavily most pertinent information and negates other

cyan CB	confirmatory bias	looks for or weighs most information that coincides with pre-existing expectations or beliefs
yellow AFF	affect heuristic	evaluates with affect, or emotion -look for words such as “I feel…”, “I want to…” -collective pronoun “we” to show affect, connectedness
navy REC	recency	prioritizes most recent score, performance in decision-making