

GRADING FAIRLY – A REFLECTION ON EXTRANEOUS FACTORS AFFECTING ABILITY TO GRADE OBJECTIVELY

James Emmet Owens

Kanda University of International Studies

ABSTRACT

Sociological research has demonstrated that a range of factors compromise how objective people are capable of being when making decisions that affect others, e.g., a person's appearance, the hunger / blood sugar levels of the person making the decision at the time of judgement (Danziger et al., 2011). There is good reason, then, to assume that teachers are similarly compromised by extraneous factors when grading students' written work. In this paper, the author investigates the issue by inviting colleagues via a survey to reflect on what may impair their own ability to grade fairly, and on possible ways of mitigating this interference. The paper will hopefully be of use to all educators in any region who are interested in minimising their own subjectivity when awarding grades, firstly by raising self-awareness of such human tendencies, and secondly by making some useful suggestions to lessen their negative impact.

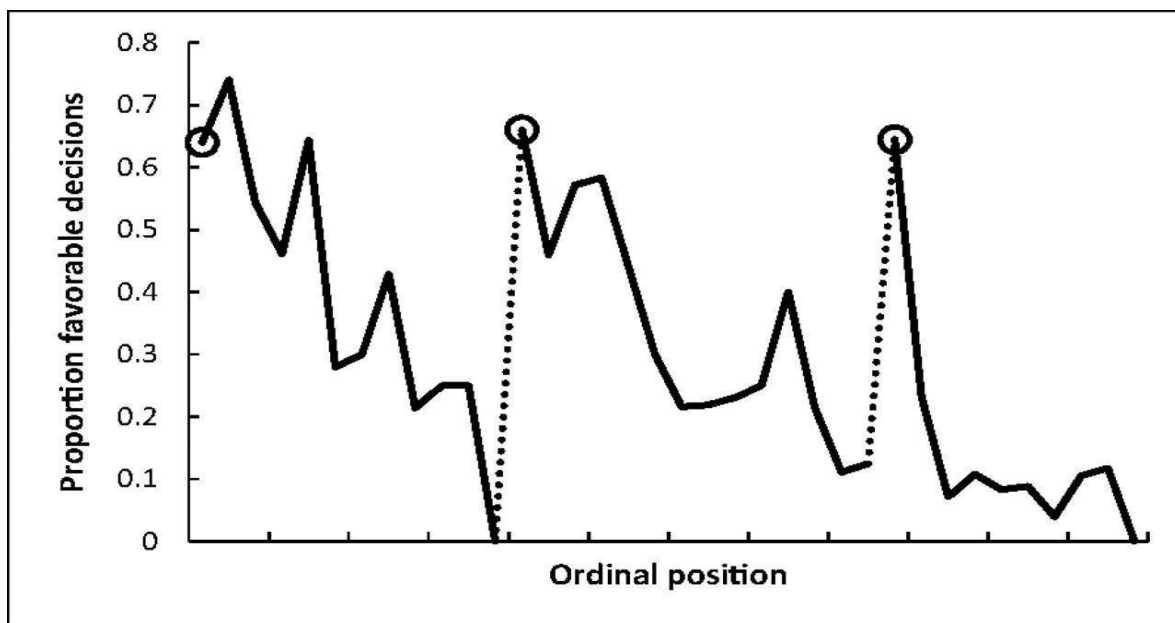
INTRODUCTION

The impetus for this paper was a study conducted by Danziger et al. (2011) that found that "...judicial rulings can be swayed by extraneous variables that should have no bearing on legal decisions" (p. 6889). While seemingly unrelated to education, it encouraged me to reflect on the degree to which a multitude of factors, some unconscious, affect our ability to make important decisions. The first televised American presidential debate, between Nixon and Kennedy in 1960, famously highlighted one such "extraneous variable": a person's appearance; those listening via radio felt Nixon had the edge, while television viewers were overwhelmingly impressed by Kennedy's visual charm (The Kennedy-Nixon debates, 2010). Kennedy went on to win the election by a tight margin (Webley, 2010). These "swayed" decisions can clearly have a consequential impact on other people's lives. More relevantly, my own experience as a language test examiner has led me to reflect on

inconsistency in grades awarded to examinees in speaking tests depending on whether the context is face-to-face or listening via audio.

Danziger et al. (2011) found that time of day and hunger levels of Israeli judges had a significant impact on the “percentage of favorable” (p. 6889) judicial rulings (see Figure 1). It can be reasonably assumed that this is also true of educators when grading student work. Thus, a study was conducted comprising a brief survey in which lecturers and advisers working in a university in Japan were asked to voluntarily reflect on what extraneous factors might affect their ability to grade written work fairly, and how they might go about mitigating such factors. While the context here is Japan, the issue is of relevance to all educators. Extraneous factors affect us all and should concern anyone serious about grading fairly. Before the survey and its results are discussed, however, a brief rudimentary summary of related literature in the field is necessary to establish the context.

Figure 1. “Proportion of rulings in favour of the prisoners by ordinal position. Circled points indicate the first decision in each of the three decision sessions ... dotted line denotes food break” (from Danziger et al., 2011, p. 6890).



LITERATURE REVIEW

Interest in the inherent flaws of teacher subjectivity is no new thing. Writing over a century ago, Chase (1914) questioned the ability of teachers to provide reliable grades to their students, concluding that, “... let us not be too sure of our judgement and our ability to grade papers

111

accurately" (p. 82). More recent research has highlighted the gender of the student submitting work as one such variable unfairly affecting grades. Lavy (2008), for example, offers evidence that "strongly suggests a bias against boys" (p. 2103) in both humanities and sciences in Israel, contending that this "bias against male students is the result of teachers', and not students', behaviour" (p. 2083). Dee (2007), on the other hand, found that in the U.S, teachers tend to allocate higher grades to students of the same gender, a conclusion supported by Ouazad and Page (2013) based on their own study in the U.K.

Gender is a thorny issue, however, intertwined with many other correlating variables, such as culture. Indeed, higher grades for females could simply be a result of lower expectations, or a conscious attempt to reverse a perceived gender gap in larger society. It is also a variable that is to a large extent outside the control of most people: in a vast majority of circumstances, neither a teacher nor a student would alter their gender to neutralise such bias. An increased awareness of potential bias on the part of the teacher is a viable goal, however, and blind marking (i.e., where the author is anonymous) is one possible solution (see Findings and Discussion section below).

Less controversial, and of arguably more relevance, are extraneous variables such as those raised by Brackett et al. (2013), who look at the role of emotion in affecting ability to mark student work objectively. They conclude that their research "...provides initial evidence that emotions may bias the grades teachers assign to their students, such that positive and negative emotions may influence grade assignment in emotion-congruent ways" (p. 643). In simple terms, teachers are more likely to award higher grades if they are in a good mood, and lower grades if they are in a bad mood. Whilst this might seem an obvious conclusion, given this evidence more consideration should be granted to rectifying the consequences. Grades that students are awarded can have important effects on their lives, for example, on their motivation to study, or on the tier or level of class they enter. In more extreme contexts, it may mean the difference between being able to study abroad or not (in the case of IELTS or TOEFL test score requirements) or finding employment. Additionally, teachers may or may not have knowledge of these consequences for their students, which may in turn affect their own mental health and stress levels. The ramifications of these extraneous factors compromising ability to grade fairly, then, are far-reaching.

With this in mind, I set out to survey colleagues at a university to ask them to reflect on what variables might compromise their ability to mark objectively, and how they have thus far attempted, or might in the future attempt, to realistically minimise the effects of such extraneous factors.

METHOD

In investigating what factors affect teachers' ability to grade objectively, it should be stressed that it is not imperfections in the instrument or task itself being used to grade students (e.g., the assignments set, or the rubrics utilised) that provide the focus of this study. There are of course rational concerns with both of the above that could make up the subject of another paper. However, in this paper I wish to focus only on the extraneous factors affecting ability to grade fairly and objectively which are within an individual teacher's relative control; factors such as health, timing, emotional state, or hunger.

With that in mind, the study uses what might be termed mixed methods. A digital survey (using Google Forms) was sent out to all colleagues working in the same building as the author, at a university specialising in languages and international studies based near Tokyo, Japan. Most of the subjects were lecturers working within the English Language Institute (ELI). However, some worked in other departments or as Learning Advisers. All those who assign and grade written work were asked to voluntarily participate in the survey, and the following explanation was offered in order to be transparent and encourage responses:

The purpose of this survey is ...invite reflection. It is not to be used to evaluate any respondent, and all answers are voluntary and anonymous, unless you would like to be contacted for follow-up interviews.

In total, the request was sent to over 70 lecturers and advisers; 17 people responded and provided consent to their answers being used for research purposes, and the research project itself was cleared for ethical approval by the relevant authority within the ELI.

A non-digital format of the questions themselves can be seen in the Appendix. As can be seen from the Appendix, questions were both closed, to gain some solid statistical data from which to find general patterns, and open, in order to provide context and allow for complexity. Teeter (2015), for example, is critical of quantitative studies that "confine... responses into decontextualized boxes that eliminate diversity of opinion" (p. 104). Although her context is specifically student responses, the point of avoiding oversimplification is an important one. The closed questions required answers, and the open questions did not. I felt that doing so would garner the greatest number of respondents by making the survey potentially brief, while allowing those with stronger opinions to make their case should they feel they wanted to. The next section presents a summary, and discussion, of these responses.

FINDINGS AND DISCUSSION

Professional Experience of Respondents

The responses to questions in Section A suggest that all respondents have a considerable amount of teaching experience, and the majority (59%) have some professional training in grading / written assessment. All respondents answered that they had been teaching for at least five years; 47% for ten years or more. This level of experience with teaching English, and writing in particular, arguably makes their responses more likely to broadly represent opinions and issues facing many teachers in similar positions.

Extraneous Factors

When asked in Section B which of the suggested extraneous factors (see Appendix) they thought might affect their own ability to mark objectively when grading written work, respondents were required to answer, "not at all", "to some degree" or "substantially" for each suggestion. Interestingly, none of these suggestions were dismissed outright by all respondents, i.e. each of the proposed extraneous factors resulted in at least some respondents admitting it might compromise their ability to grade fairly. Those carrying the lowest number of answers indicating some level of effect were "A wish to be liked by the students" (12 out of 17 respondents answering "not at all"), "Amount of sugar / caffeine I have ingested" and "Agreeing / Disagreeing with an opinion expressed by the student in an assignment" (11 "not at all" responses each).

Conversely, large numbers of teachers/advisers admitted that certain factors likely had some detrimental effect on their ability to grade fairly. As many as 15/17 respondents indicated either "substantially" or "to some degree" being affected by "Comparison to other work e.g. the assignment marked previously to this one was significantly better/worse". 14 respondents acknowledged that "Level of interest in the topic the student decided to write about (where a choice is offered)", "How overloaded with work I am", and "Time of day" could all have an unfair effect on the grade they decided to award. Some of the answers tended to stress aspects that were arguably, or to some extent, beyond the control of the responder, for example, nearly half (eight) of all subjects answered, "How overloaded with work I am" to substantially have an effect, while 12 responses indicated "Where I am marking e.g. noisy office...", to be a potential factor (to some degree or substantially). This issue will be discussed later.

When asked to explain their answers, one respondent wished to make "...a technical point - for a number of the answers above, I would probably answer 'not sure' e.g. Sugar intake. But this option wasn't

available.” This is a valid criticism of the limited choices available in the answers themselves, but I did not want to present a “not sure” option as it may have led to many teachers over-relying on this as an answer for most of the suggestions. The very nature of this survey is very subjective and relies on honesty and consideration on the part of the subjects. It is the author’s opinion that any survey that allows for indecisive answers can, depending on the context, undermine its very purpose, if that purpose is to find a coherent set of answers and/or to encourage the respondent to reflect on their own situation; thus, it is often best to “force” respondents in to offering some sort of opinion. As the intention here was to investigate leanings rather than convictions (Payne, 1951), removing this ‘not sure’ option was deemed valid.

Another respondent offered the opinion that their grading might be compromised by wishing to “...hold some [students] to a higher standard because they can attain it and should expect it of themselves, while encouraging those who are just not there yet but trying hard and making progress”. This is essentially about the human aspect of being a teacher. It is not always easy, nor necessarily expedient, to be ruthlessly consistent in grading when we know that certain students might benefit more/less from positive/negative feedback. This accounts for the 13 respondents who felt “Sympathy towards the student e.g., you know they need a high grade to pass/motivate themselves” to be a factor (substantially or to some degree). On a similar note, one respondent pointed out that “it is hard to be objective, if not impossible because there are no lines that can be drawn as to what is good or bad, since these are subjective ideas anyway.” Another respondent conceded that “content” and “ideas” carried a heavier weight for them than “mechanics” or “grammar”. All these points are valid, but in order to ensure consistency they also indicate the importance of developing effective, universally accepted rubrics for grading.

When asked in both Sections A and B (questions 5 and 7) what other factors, not listed, might affect their ability to mark fairly, the respondents volunteered many interesting suggestions, e.g. “hormones”. This is similar to emotional state or level of fatigue, and would be a difficult variable to control for in any experiment.

The “halo effect” (Thorndike, 1920) came up a lot, i.e. graders acknowledging a cognitive bias that encourages them to award higher grades to students they know who, for example, work hard or who have a good “attitude towards study”, or even “what other students or teachers say about the student”. This would account for the relatively high number of respondents (13/17) admitting that “Whether or not I have positive feelings towards the student (where the writer’s name is known)” likely

compromises, to some degree or substantially, their objectivity in grading. Marking blindly might be one effective way of dealing with some of these issues; that will be discussed later.

Another teacher admitted the “formatting” of an assignment affected their grading, but an argument here could be that this is fair; presentation of work is an important academic/life skill, therefore formatting should be included in the rubric, and thus form part of classroom instruction, too. One respondent challenged the example given in the survey that, “I might be less likely to give higher grades to the 10th assignment if I’m marking in bulk”, by asserting:

I do think timing influences my grading to some degree, but in a way opposite to the example. I think I am a more impartial grader if I grade in bulk because if I block out a large chunk to grade, I feel in a good head space to grade... I find grading in bulk allows me to be more consistent.

Clearly, there are no universal solutions to ensuring we all mark as fairly as we can. Different teachers/advisers have their own ways of securing consistency in their grading patterns, and it was the main objective of this survey to simply encourage reflection on such practice.

Finally, an additional interesting finding from this section is that 12 out of 17 respondents felt that they might be overcompensating in some way when grading i.e. they are “aware of” the potential influence of extraneous factors such as hunger or the halo effect, “and in an attempt to correct this, possibly overcompensate and award an unfair grade” (see the last suggestion in question 6, Appendix). That so many teachers/advisers demonstrated this level of metacognition (or “meta-awareness”) serves as an effectual example of the way that grading student work can be complicated, full of uncertainty, and indeed stressful.

Emotional Effect on Grader

When asked in Section C if they had ever considered how such factors may affect the objectivity of their grading before taking the survey, 100% of respondents answered “yes” (76.5% responding “Yes, I often concern myself with this”). This strongly implies that the area of extraneous variables affecting grading is an important issue that has thus far been neglected and under-researched. It has implications not only for the students but also for the mental well-being of graders. Indeed, only one respondent answered that they “mostly” enjoy grading (see question 11), 58.8% answering “mostly no”, and more than half of all respondents admitted putting off grading, 17.6% doing so “frequently” because they found it stressful (question 12).

There is an interesting variety in the answers provided to the question, "How does the idea that your decision (i.e. grade awarded) might affect someone's life make you feel?" (question 10). For some, the pressure leads to anxiety and stress. One respondent confessed to feeling "tons of pressure". Another replied, "It makes me feel very anxious. Sometimes guilty," and similar comments include, "It makes me feel nervous", "I don't like it", and "I worry that my decision could discourage some students". One respondent attempted to be stoic and pragmatic about the process:

When it comes to grading, I try to just be as fair as I can and not worry too much about the grading being a potential life-altering "decision" of mine-- that's just extra stress I don't need to put on myself.

However, they then conceded that the pressure "... can really eat away sometimes."

However, an almost equal number of respondents gave a very opposite portrayal of their feelings. For example, one person wrote that grading is "...intrinsic to the job of a good educator, and I have to work under that responsibility". In fact, the words, "responsible" and "responsibility" came up unprompted in similar contexts in five separate answers to this question. For many, it would seem, stress is part of any job, and grading work is no "big deal" unless you choose to let it become an issue. Another response exemplifying this sort of mindset is the comment, "Indifferent - it's not the end of the world - most of the time, students themselves are aware that I treat each work quite fairly."

Clearly, then, some teachers/advisers find grading and its implications very stressful, while others accept it as "part of the job". Some respondents find themselves switching between these two mindsets, or occupying a grey area somewhere in between. Where there is more consistency, however, is in the shared belief that effective rubrics and training can help in this matter. Some respondents feel this is already adequate: "By having clear rubrics... I can rest easy because I know I did what I could and the rest is on them." Others would like to see improvements in this regard, arguing that, "...a better assessment system with alternative assessment methods and an assessment support system is needed in order to deal with this issue." Similarly, another opined that they "...would like to grade with balance and objectivity, and would appreciate more training and guidelines on how to do this."

Solutions: Redressing Potential Extraneous Factors

In Section D, respondents were shown a list (see question 15) of ideas

that might mitigate the effects of extraneous factors, and asked to decide which statement best described its potential usefulness: “I have tried this but don’t think it’s effective”, “Unlikely I will try this”, “I haven’t tried this but it sounds like a good idea”, or “I have tried this and I think it’s a good idea”. The main results are presented in table format below.

Table 1. Total Responses to Question 15, Ordered by Response Type

Response	Total
I have tried this but don’t think it’s effective	8
Unlikely I will try this	47
I haven’t tried this but it sounds like a good idea	39
I have tried this and I think it’s a good idea	76

The most satisfying aspect of the results here is that the “positive” responses i.e. those indicating a willingness to proactively attempt to redress extraneous variables (“I haven’t tried this but it sounds like a good idea”, and “I have tried this and I think it’s a good idea”) far outweigh the “negative” answers (“I have tried this but don’t think it’s effective”, and “Unlikely I will try this”), as can be seen in Table 1. Only eight responses out of a total of 170 (10 suggestions and 17 respondents) were able to dismiss the effectiveness of an idea based on having already tried it, whereas 76 responses (almost 45%) were able to verify the effectiveness of ideas based on the same experience. It should be noted, however, that these responses all came from willing volunteers, and thus this positivity may not be a representative sample.

Table 2. Most Frequent Answers for “I Have Tried This but Don’t Think it’s Effective”

Response	Total
Snacking / drinking while marking	3
Marking in short bursts rather than in bulk	3

While answers indicating “I have tried this but don’t think it’s effective” numbered only eight in total, at least 3 people in each category felt that

snacking/drinking, and marking in short bursts (see Table 2) were ineffective based on their own experience. Earlier, we saw how one person opined why they felt “short bursts” to be ineffective; in the comments section here (question 17) another respondent explains that “Snacking while grading just distracts me”. However, this is clearly an issue that divides opinion as nine people have found snacking to be effective (see Table 5 below). The answers here should also be seen in the context that “How hungry I am”, “Amount of sugar / caffeine I have ingested” and “Timing” were all seen as relatively minor concerns in the answers to Section B (e.g. 11 respondents felt that sugar and caffeine levels were “not at all” a factor in undermining their ability to grade fairly).

Table 3. Most Frequent Answers for “Unlikely I Will Try This”

Response	Total
Letting students know beforehand about topics you have strong feelings about	13
Having someone else double mark a sample	9
Going back over and reading each assignment again	7

As can be seen from Table 3, a large majority of respondents are, by their own admission, unlikely to let students know topics they have strong feelings about, and no one reported any success with this idea. I practise this myself, as I feel it only fair to warn students of areas that I am likely to have my own ability to mark objectively compromised. That this is clearly such an unpopular idea with colleagues is interesting, but unfortunately no comments were written that could shed some light on why. This could warrant further investigation in any future paper; perhaps the reason is that a teacher/adviser may fear it would undermine them in the eyes of their students, or they see it as unethical to nudge students away from certain topics they may be interested in based on the feelings of the teacher. On a related note, one respondent did suggest telling “...students what the evaluation criteria are before they start their writing and evaluate their writing based on the criteria.” It certainly seems beneficial, and fair, to present rubrics, or simplified versions, to students in advance.

The other two ideas that generate the least enthusiasm (double-marking and re-reading) can be attributed perhaps to lack of time rather than them being unhelpful suggestions, as both ideas also had high numbers of advocates (see Table 5). One of the comments would seem to verify this

conclusion: "One of the realities in this job is time and a high workload. While re-reading assignments is an excellent idea, I'm more likely to only read those which I was unsure about (due to time constraints)."

Table 4. Most Frequent Answers for "I Haven't Tried This but it Sounds Like a Good Idea"

Response	Total
Marking at the same time of day each time	9
Not marking if my health or emotional state is poor	6
Marking blindly	6

When it comes to untested suggestions that people are willing to try (Table 4), "Marking at the same time of day" generated the most enthusiasm. Pleasingly, more than half of all respondents have as a result of taking this survey considered trying a new idea that could ensure they are more able to grade work objectively. It relates directly to the aforementioned work done by Danziger et al. (2011), where time of day was found to have a significant impact on the "percentage of favorable rulings" (p. 6889) by judges. It would be very interesting to find volunteers and conduct more focused research on this one variable, to investigate if it has any similarly significant impact on grading.

Table 5. Most Frequent Answers for "I Have Tried This and Think it's a Good Idea"

Response	Total
Not putting off marking / leaving it to the last minute	12
Marking blindly	11
Following the rubric strictly rather than comparing assignments	11
Snacking / drinking while marking	9
Going back over and reading each assignment	9

The idea that people found most effective based on their experience was

avoiding procrastination: 12 people out of 17 can verify that this is "a good idea". Presumably others acknowledge its potential effectiveness, but know they are unlikely to change their behaviour, or they simply do not have enough time to complete their grading early. There seems more realistic interest in the idea of "Marking blindly": in addition to the 11 people who have tried this and approve (Table 5), six respondents have not yet tried it but are willing to (Table 4). All 17 respondents believe blind marking is or could be effective in minimising bias. Nobody sees this as a "bad idea". However, there are problems with this suggestion. Most teachers/advisers see any initial first drafts of student work and offer personal feedback, thus rendering it almost impossible to not know student names when the final drafts are turned in. There can also be clues within the content and style of the work if the teacher knows the students well, and if the work is submitted electronically, the email address could be another "dead giveaway" as to authorship.

"Following the rubric strictly" also carries a lot of approval based on experience (11 respondents). This is echoed in one of the answers to question 17: "...for certain courses I think course-standardised rubrics need to be made. This will help with things like double marking and ensuring that all teachers are coherent." However, one person felt that, where the rubric fell short, their advice for teachers would be, "don't change the points for it, but do address it in the comment instead." This was echoed by another respondent: "Make a habit of a short comment for feedback in addition to the rubric." Clearly, teachers have a lot of faith in, and would like to see more effective use of, rubrics to ensure greater consistency and even allow teachers to double mark a sample of each other's work, but even the best rubrics can be imperfect and an additional comments section in the feedback is essential.

LIMITATIONS

Before drawing any broad conclusions from the above analysis, it is prudent to consider the limitations that frame the research. The sample size is small, which by itself is one factor preventing any sort of quantitative, scientific analysis from taking place that can implicate significant proof of any extraneous factors affecting objectivity. The main objective of the survey was more modest; I simply wanted to encourage teachers/advisers to consider the potential effect of extraneous variables within their control in the hope that increased self-awareness alone might mitigate some such bias.

The answers to the survey are very subjective by their very nature. The survey asked respondents to essentially guess their own tendency to be affected by extraneous factors. Some people are more self-critical than

others. The issue is further complicated by political issues; the author's position may compromise the reliability of the answers. This may partly explain the aforementioned relatively high number of people expressing dissatisfaction with large workloads, noisy offices, and lack of rubrics. This may or may not be true, but the purpose of this survey was to encourage self-reflection and a sense of personal responsibility rather than advocate institutional improvements.

Nonetheless, despite the above limitations, the research generated some very interesting points to consider, and the impetus for further work.

CONCLUSION

As seen in their responses to Section A, the teachers/advisers taking part in this study are all experienced professionals, and thus their answers are all the more compelling.

All the suggested extraneous factors in Section B were felt to some extent to compromise objectivity when grading student work, but some were felt to be more of an issue than others. Related to the work done by Danziger et al. (2011) that provided the impetus for this study, the time of day when one marks an assignment is one such example. Being more consistent in this regard also generated enthusiasm as something worth attempting for those who had not tried it before. This is an area of research that is well worth exploring in more detail. Five respondents have volunteered to be interviewed or used as case studies (Section E). One idea would be to monitor the degree to which time of day actually has an impact using quantitative methods. Unlike other, thornier issues (such as emotional state, hormone levels, or hunger) this variable may be a fairly comprehensive one to control.

This preliminary study then has provided an abundance of ideas and impetus for further study. Another direction research could take is to look in more detail at the implications for stress and mental health of graders, some of whom feel grading as a large burden. One consistent conclusion drawn from respondent answers is the need for more training and more carefully constructed rubrics. This might alleviate some of the pressure felt by teachers/advisers, and create more objective standards of grading in the process.

This paper has been effective in the modest way that it was intended: firstly, it has demonstrated that extraneous factors are a neglected issue that should be taken more seriously - as mentioned earlier, all respondents had considered this matter before, with 76.5% "often concerning" themselves with it. Secondly, it has encouraged colleagues to

reflect on what specifically might be compromising their ability to be objective and how they can go about minimising this, offering suggestions to respondents and receiving some in return. This "give and take" process has proven productive and immensely worthwhile, and I would encourage any interested party to engage in a similar discourse with fellow professionals.

REFERENCES

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

Chase, H. (1914). Are teachers' grades reliable? *The North Carolina High School Bulletin*, 5(2), 79–82. <http://www.jstor.org/stable/43799300>

Danziger, S., Levav, J., & Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences Apr 2011*, 108 (17), 6889–6892. <https://doi.org/10.1073/pnas.1018033108>

Dee, T. (2007). Teachers and the gender gaps in student achievement. *Journal of Human Resources*, XLII (3), 528–554.
<https://doi.org/10.3368/jhr.XLII.3.528>

Lavy, V. (2008). Do gender stereotypes reduce girls' or boys' human capital outcomes? Evidence from a natural experiment. *Journal of Public Economics*, 92(10–11), 2083–2105.
<https://doi.org/10.1016/j.jpubeco.2008.02.009>

Ouazad, A., & Page, L. (2013). Students' perceptions of teacher biases: Experimental economics in schools. *Journal of Public Economics*, 105, 116–130. <https://doi.org/10.1016/j.jpubeco.2013.05.002>

Payne, S. L. (1951). *The art of asking questions*. Princeton U. Press.

Teeter, J. (2015). Deconstructing attitudes towards plagiarism of Japanese undergraduates in EFL academic writing classes. *English Language Teaching*, Vol 8 (1), 95–109. <https://doi.org/10.5539/elt.v8n1p95>

The Kennedy-Nixon debates. (2010). A+E Network.

<http://www.history.com/topics/us-presidents/kennedy-nixon-debates>

Thorndike, E. (1920). A constant error in psychological ratings. *Journal of*
123

Owens, J. (2023). Grading fairly: A reflection on extraneous factors affecting ability to grade objectively. *Literacies and Language Education: Research and Practice*, Spring 2023, 110-131. English Language Institute, KUIS.

Applied Psychology, 4, 25–29. <https://doi.org/10.1037/h0071663>

Webley, K. (2010, September 23). How the Nixon-Kennedy debate changed the world. *Time Magazine*.
<http://content.time.com/time/nation/article/0,8599,2021078,00.html>

APPENDIX

Print Version of Survey: Impediments to Grading Objectively - Extraneous Factors

The purpose of this survey is to investigate emotional responses to the process of grading written work, and especially the inherent subjectivity involved in this same process, and also to invite reflection. It is not to be used to evaluate any respondent, and all answers are voluntary and anonymous, unless you would like to be contacted for follow-up interviews.

Thank you very much in advance for completing this survey! Hopefully the reflection process will be of benefit to anyone taking it. The results should prove very interesting and I will be happy to share them.

Should you have any questions, you can contact me at #####@#####.ac.jp, or come and see me in person.

*Required

1. By continuing with this survey, I give my consent for my answers to be used for research purposes. * Mark only one oval.

I give my consent

I do not give my consent

Section A: Professional experience

2. How long have you been teaching English for (in any environment/country)? (rounded up to the nearest year) * Mark only one oval.

1-2 years

3-4 years

5-6 years

6-7 years

7-8 years

8-9 years

10 years or more

3. Have you any professional training in grading / written assessment? (e.g. IELTS examiner, inhouse training) * Mark only one oval.

125

Yes
No

4. If “yes”, what training have you received?

5. What factors do you think may affect your own ability to grade objectively when it comes to marking written work? (If you’re unsure, no answer is necessary at this point. The next section will clarify the question by offering some examples)

Section B: Extraneous factors

The main focus of the study is subjectivity involved in the grading process, but not the instrument or task itself (e.g. the assignment or the rubric). Rather, I wish to look at how the emotional state or condition of the person grading has an effect.

Various sociological studies have shown that a range of factors compromise how objective people are capable of being when making decisions that affect other people. E.g. Danziger et al (2011) suggest that time of day / hunger levels have an impact on the “percentage of favorable rulings” given by Israeli judges in judicial rulings. I would like to invite teachers to reflect on the fact that they may be similarly affected by “extraneous factors”.

6. Answering as honestly as you can, which of the following do you think could affect your own ability to mark objectively when grading written work? * Mark only one oval per row.

	Not at all	To some degree	Substantially
Time of day	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
My mood in response to personal issues (e.g. recent bad/good news)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Level of interest in the topic the student has decided to write about (when a choice is offered)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Whether or not I have positive feelings towards the student (where the writer's name is known)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
How hungry I am	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Amount of sugar / caffeine I have ingested	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Where I am marking (e.g. noisy office, cramped train, at home)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
My health at the time of writing	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
How overloaded with work I am	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Timing e.g. I might be less likely to give higher grades to the 10th assignment if I'm marking in bulk	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Comparison to other work e.g. the assignment marked previously to this one was significantly better / worse	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sympathy towards the student e.g. you know they need a high grade to pass / motivate themselves	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A wish to be liked by the students	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Conversely, a wish to be seen as strict by students	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
You don't agree with the rubric's grading scheme	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Agreeing / Disagreeing with an opinion expressed by the student in an assignment	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Over-compensation: i.e. you are aware of one of the above trends (e.g. you are hungry or you like the student on a personal level) and in an attempt to correct this, possibly overcompensate and award an unfair grade	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

7. Are there any other factors, not mentioned above or in Section A, that you feel may also affect your ability to mark objectively when grading written work?

8. Please comment if you wish to further elaborate on your answers, but only if you feel it necessary.

Section C: Emotional effect on teacher/adviser

9. Have you ever considered how any of the above factors (Section B) may affect the subjectivity of your grading before? * Mark only one oval.

- NO
- SOME THOUGHT BUT NOT IN TOO MUCH DEPTH
- YES, OFTEN CONCERN MYSELF WITH THIS

10. How does the idea that your decision (i.e. grade awarded) might affect someone's life make you feel?

11. Do you enjoy grading? *

Mark only one oval.

- Mostly no
- Sometimes
- Mostly yes

12. Do you ever put off grading because you find it stressful? * Mark only one oval.

- Yes frequently
- Sometimes
- No, not usually

13. Do you grade upwards or downwards? (Grading upwards involves a starting assumption of 0% for a grade, and awards marks upwards from that position e.g. this student has shown me they can use paragraphs, that means they will qualify for at least 40%. Grading downwards involves a starting assumption of 100% for a grade, and deducts marks each time a student fails in some way e.g. this student has failed to use paragraphs, so they can get no higher than 60%) * Mark only one oval.

- Predominantly grade upwards
- Depends on the assignment
- Predominantly grade downwards
- Don't know

14. Please comment if you wish to further elaborate on your answers, but only if you feel it necessary.

Section D: Redressing potential extraneous factors

15. How do you, or might you, approach dealing with some of the issues from Section B? * Mark only one oval per row.

- I have tried this but don't think it's effective
- Unlikely I will try this
- I haven't tried this but it sounds like a good idea
- I have tried this and I think it's a good idea

	don't think it's effective	will try this	sounds like a good idea	think it's a good idea
Marking in short bursts rather than in bulk	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Not putting off marking / leaving it to the last minute	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Marking at the same time of day each time (e.g. only after lunch) to ensure consistency	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Snacking / drinking while marking	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Not marking if my health or emotional state is poor	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Letting students know beforehand about topics you have strong feelings about (e.g. "I hate Disney")	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Following the rubric strictly rather than comparing an assignment to other assignments	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Marking blindly (not being able to see the student's name)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Going back over and reading each assignment again after you have finished marking them all	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Having someone else double mark a sample of your assignments to check for consistency	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

16. Are there any other ideas, not mentioned above, that you feel may also help redressing any extraneous factors?

17. Please comment if you wish to further elaborate on your answers, but only if you feel it necessary.

Section E

Finally, please feel free to leave your name and email address if you would be willing to be interviewed or used for a case study in the future as a follow-up activity.

Or, if you would be willing to take part but do not want your answers here to be linked with your name, please send me an email separately, indicating that you are happy to volunteer

18. I am happy to leave my contact details and be contacted for interviews / case studies. Here is my email address: