Department of Policy and Planning Sciences

Discussion Paper Series

No.1325

Incentives and information order with applications

by

Michiko OGAKU

Mar 2015

UNIVERSITY OF TSUKUBA
Tsukuba, Ibaraki 305-8573
JAPAN

Incentives and information order with applications*†

Michiko Ogaku‡

Preliminary

20 March 2015

Revised 29 February 2016

This version: 13 December 2017

Abstract

One might predict students to be less motivated to work when evaluated according to a coarse rating pattern, such as pass or fail, rather than a primary score $(1, 2, \ldots, 100)$. However, coarser rating patterns often induce more effort. We ask what conditions will guarantee that the primary score is the most motivational grading. This paper provides a simple sufficient condition for this: the primary score is the most motivational grading when a higher rank is associated with greater ability by the market, and when greater effort increases the chance to achieve a higher rank at any level of effort. The proposed condition is given via two simple functions, a market belief and a likelihood ratio, which do not require knowledge for the distribution of ability. Thus, this may open an avenue for examinations of a theory on the impact of information on implicit incentives by experimental and empirical research.

Key words: career concerns; Blackwell; incentives; grading; transparency.

JEL: D86; D83.

---

*I thank participants at EBES, EARIE and BEAM Kyoto International Conference for helpful comments; and Ishii Memorial Security Research Promotion Foundation for financial support.

†Earlier versions of this paper were circulated under the title "Career concerns and negatively correlated signals conditional on garblings thereof."

‡Faculty of Economics, Nagasaki University, 4-2-1 Katafuchi, Nagasaki, 850-8506 Japan.

1 Introduction

One might predict that students would be less motivated to work when a coarse score such as pass or fail is disclosed to parties such as university committees, graduate schools or employers (hereafter, the market) rather than the primary score, $1, 2, 3, \ldots, 100$. In fact, the answer to this question is mixed.

The purpose of this paper is to explore how this garbling of information, which is a typical example of a transformation of information structure in the sense of Blackwell (1951, 1953), has an impact on incentives for individuals (e.g. students, researchers and CEOs) who care about the evaluation their ability by the market.

A prominent study by Dewatripont, Jewitt and Tirole (1999, hereafter, DJT) provided conditions for identifying the impact of garbling on effort. However, these conditions are not satisfactory for rating problems such as grading students. Because their conditions require a joint probability density function for a garbled report and ability, one cannot directly apply the conditions to a rating pattern that has a probability distribution with, for example, masses of 1/6, 1/2 and 1/3 at A, B and C, respectively.

The main contribution of this paper is to provide a sufficient condition for primary score reporting to be the most motivational grading: (i) if the market belief about ability after observing the primary score, $\mathrm{E}[\theta | X_{\hat{a}} = x]$, is an increasing function; and (ii) if the likelihood ratio of the primary score is an increasing function, which is often referred as the monotone likelihood ratio property (MLRP), at any level of effort.

The conditions presented extend the corresponding conditions given by DJT to include grading reports such as a letter grade. In addition, our conditions are more accurate with regard to equilibrium analysis.

Importantly, because this paper's conditions do not require knowledge of the distribution of ability, it is easier to test in experimental and empirical research. In the context of grading, the market belief about ability may be increasing with respect to score. Therefore, if the primary score satisfies MLRP at any level of effort, then the primary score will be the best grading for motivating students to work.

Interestingly, Dubey and Geanakoplos (2010) investigated a similar question and provided an answer that is inconsistent with this paper's conclusion: the primary score cannot be better than a properly garbled score. The reason that the difference arises is as follows. They assumed that students who compete with each other for relative ordering are ordered by ability ex ante, and no chance is given to change this order by effort. Under this assumption, each student knows his/her ability and the primary score reveals their ranking, while a coarser letter rank gives an opportunity to be mixed with higher ranking students through effort. On the other hand, this paper supposes that the ability is not known by everyone ex ante, which is perhaps in the spirit of a remark in Lazear and Rosen (1981): contests are feasible only when chance is an important factor.

Rodina (2016) independently showed a result similar to ours in a study of principal's information design.[1] Their result shows a sufficient condition under which fully revealing the output gives the best pure strategy equilibrium. The major difference with this paper is that Rodina (2016) does not assume differentiability of the distribution of signals with respect to effort, but instead assumes supermodularity of the conditional expectation of ability given effort and output with respect to effort and output.

However, it may be difficult to test the supermodularity in Rodina (2016), or whether an increase in the market belief about ability under an increase in output does not decrease when greater effort is conjectured by the market. Our sufficient condition is much easier to check. In addition, our proof is simpler and straightforward.

This paper also relates to the literature on the relationship between formal incentive contracts and information order. The important implication from the literature is that garbled information makes second-best contracts more expensive (Holmström, 1979; Grossman and Hart, 1983; Kim, 1995). Thus, more information is better for the principal if formal contracts based on the information are feasible. Such a formal contract is infeasible for the motivational grading studied in this paper.

The rest of this paper is organised as follows. The next section presents a motivating example in which pass/fail reporting is a better incentive device than full reporting. Section 3 contains the main result. Section 4 summarises the analysis and provides

[1] I would like to thank Professor Barton L. Lipman for kindly pointing out this article.

conclusions.

2 Motivating example

The example consists of a market, a professor and a representative student. The professor's problem is choosing a disclosure rule that encourages the student to work. The student's problem is choosing an optimal level of effort to raise the market belief about his/her ability as much as possible.[2] The timeline of events for this example is as follows.

1. The professor decides the disclosure rule; full or pass/fail.
2. The student decides the level of effort.

   (The student does not know his/her own ability ex ante)
3. The signal for ability is realised.
4. The signal is reported in accordance with the rule.
5. The market updates beliefs about ability according to the report.

Signals

We define a probability space (Ω, Σ, P) for signals as follows. Let $\Omega = \{w_i, 1 \leq i \leq 4\}$, where the elements w_i of $\mathbb{R}^2$ are specified by

$$
\begin{aligned}
& w_1 = (0,0),\ w_2 = (1,1),\ w_3 = (5,5), \\
& w_4 = (0,1).
\end{aligned}
$$

Let Σ be the power set of Ω, and let $P[\{w_1\}] = P[\{w_4\}] = 1/6$ and $P[\{w_2\}] = P[\{w_3\}] = 1/3$.

Next, consider a process $X_1,\ X_2$, where $X_i(\omega)$ is the ith component of w_i, so that $X_1(w_4) = 0, X_2(w_4) = 1$. Let X_1 be the grade without effort. For simplicity, let X_1 be the perfect signal for ability in the sense that $X_1 = \theta$. Let X_2 be the grade with effort.

[2] A similar model can be found in DJT. We change it so as to satisfy the condition used in Section 3: any realisation of the signal whose distribution is chosen strategically by the student is measurable using the distribution of signals that the market conjectures.

Reward and cost

The student's reward is the market's belief of his/her ability after observing the report. The report is either full disclosure of the grade in the exam or the garbled disclosure of whether the student passed or failed the exam. We assume that those whose grade is 1 or 5 pass the exam and fail otherwise. We assume that no-one knows his/her ability ex ante, and that the student and the market know the distributions of signals.

The student's effort incurs a private cost $c(\text{index})$ that satisfies

$$\frac{1}{6} < c(2) - c(1) < \frac{2}{5}.$$

Expected rewards

Table 1 indicates the student's rewards and the corresponding probabilities under the full disclosure rule. The expression $\mathrm{E}[\theta|X_{\hat{1}} = x]$ (resp. $\mathrm{E}[\theta|X_{\hat{2}} = x]$) represents the market belief about ability after observing the grade x when the market conjectures that the student shirks (resp. works). If the student does what the market has conjectured, then his/her expected reward is $\mathrm{E}[\theta] = 2$. If the student works when the market conjectures that he/she shirks, then the student's expected reward is: $\mathrm{E}[\mathrm{E}[\theta|X_{\hat{1}}]] = 1 \times 3/6 + 5 \times 2/6 = 13/6$. If student shirks when the market conjectures that he/she works, then the student's expected reward is: $\mathrm{E}[\mathrm{E}[\theta|X_{\hat{2}}]] = 1/3 \times 2/3 + 1/3 \times 5 = 17/9$.

Table 1: Rewards and probabilities under the full disclosure rule

	$x = 0$	$x = 1$	$x = 5$
$\mathrm{E}[\theta\|X_{\hat{1}} = x]$	0	1	5
$\mathrm{E}[\theta\|X_{\hat{2}} = x]$	0	$\frac{2}{3}$	5
$P[X_1 = x\|a = 1]$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
$P[X_2 = x\|a = 2]$	$\frac{1}{6}$	$\frac{3}{6}$	$\frac{2}{6}$

Rewards and probabilities corresponding to the pass/fail disclosure rule are obtained in a similar manner.

Table 2: Rewards and probabilities under the pass/fail disclosure rule

	fail: $\{0\}$	pass: $\{1,5\}$
$\mathrm{E}\left[\theta \mid X_{\hat{1}} \in \{\{0\},\{1,5\}\}\right]$	0	3
$\mathrm{E}[\theta \mid X_{\hat{2}} \in \{\{0\},\{1,5\}\}]$	0	$\frac{12}{5}$
$P[X_1 \in \{\{0\},\{1,5\}\} \mid a=1]$	$\frac{1}{3}$	$\frac{2}{3}$
$P[X_2 \in \{\{0\},\{1,5\}\} \mid a=2]$	$\frac{1}{6}$	$\frac{5}{6}$

Perfect Bayesian equilibrium

In equilibrium, the student shirks under the full disclosure rule. This is because, when the market conjectures that the student shirks and the student decides to work, then the expected reward is 13/6. However, the net benefit is less than the cost: $13/6 - 2 = 1/6 < c(2) - c(1)$. Second, when market conjectures that the student works and the student decides to work, the expected reward is 2. However, the net benefit is again less than the cost: $2 - 17/9 = 1/9 < c(2) - c(1)$.

On the other hand, the student works under the pass/fail disclosure rule in equilibrium. Because the net benefit of working is more than the cost no matter what the market conjectures. Thus, the professor would be better off if he chooses the pass/fail disclosure rule ex ante.

Remark 2.1. The model structure in this example is similar to that in Grossman and Hart (1983). However, this result contradicts one of their results which shows that a decrease in informativeness in the sense of Blackwell makes the second-best contract less efficient. This is because the market belief is not the second-best contract controlled by the principal: if the market belief were a contract, then the principal could offer 12/5 for a grade of either 1 or 5, which would induce the student to work under the full disclosure rule also.

MLRP is important

The above result is not entirely rigid. For example, suppose that those with an ability of 1 have a chance to obtain the highest score with probability 1/2;

$$\begin{aligned} &w_1 = (0,0),\ w_2 = (1,1),\ w_3 = (5,5), \\ &w_4 = (0,1),\ w_5 = (1,5). \end{aligned}$$

Then, the full disclosure rule induces more effort. In both examples, $\mathrm{E}[\theta|X_{\hat{a}} = x_i]$ is increasing and the first-order stochastic dominance holds;

$$P[X_2 \le x_i] \le P[X_1 \le x_i] \qquad \forall x_i.$$

The key difference is that MLRP is not satisfied in the former example, while it is satisfied in the latter example. To see this, let p_{ij} be the probability that $x = x_j$ when the effort is i under the full disclosure rule. Then, under the full disclosure rule, the difference between the expected reward of working and shirking when the market conjectures that the student works is

$$\begin{aligned} \left(\frac{p_{21}-p_{11}}{p_{21}}\right) \mathrm{E}[\theta|X_{\hat{2}} = x_1] \cdot p_{21} + \left(\frac{p_{22}-p_{12}}{p_{22}}\right) \mathrm{E}[\theta|X_{\hat{2}} = x_2] \cdot p_{22} \\ + \left(\frac{p_{23}-p_{13}}{p_{23}}\right) \mathrm{E}[\theta|X_{\hat{2}} = x_3] \cdot p_{23} \end{aligned} \tag{1}$$

Similarly, let q_{ij} be the probability that $z = z_j$ when the effort is i under the pass/fail disclosure rule. Then, under the pass/fail disclosure rule, the difference between the expected reward of working and shirking when the market conjectures that the student works is

$$(q_{21} - q_{11})\mathrm{E}[\theta|X_{\hat{2}} = x_1] + \mathrm{E}\left[\frac{p_{2i}-p_{1i}}{p_{2i}} \middle| X_2 \in \{x_2, x_3\}\right] (p_{22}+p_{23})\mathrm{E}[\theta|X_{\hat{2}} \in \{x_2, x_3\}]. \tag{2}$$

The difference of (1) and (2) can be written as the following conditional covariance:

$$\mathrm{Cov}\left[\mathrm{E}\left[\theta|X_{\hat{2}}=x_i\right], \left(\frac{p_{2i}-p_{1i}}{p_{2i}}\right)\middle|X_2\in\{x_2,x_3\}\right](p_{22}+p_{23}).$$

Observing the above conditional covariance, it seems that the full disclosure rule induces more effort if the market belief after observing the signal, $E[\theta|X_{\hat{a}}=x_i]$, and the likelihood ratio of the signal, $(p_{2i}-p_{1i})/p_{2i}$, are conditionally positively correlated given the result of pass or fail. This observed relationship holds more generally, which is verified in the next section.

3 Model

Consider a family of random variables, $\{\theta, X_a, a\in(\underline{a},\bar{a})\}$, on a probability space, (Ω,Σ,P). Let X_a be a signal about the agent's unknown ability θ when the unobservable effort is $a\in(\underline{a},\bar{a})$, and let $P_{X_a}(\cdot)$ be the distribution of X_a:

$$P_{X_a}(B)=P\left\{\omega: X_a(\omega)\in B\middle|\ a\ \right\},\ B\in\mathcal{B}(\mathbb{R}).$$

Suppose that for a fixed $B\in\mathcal{B}(\mathbb{R})$, P_{X_a} is a Borel-measurable function of a that is differentiable with respect to $a\in(\underline{a},\bar{a})$ and the partial derivative $\partial P_{X_a}/\partial a$ is absolutely continuous with respect to P_{X_a}. Then, by the Radon-Nikodým theorem, there is a Borel-measurable function h such that

$$\frac{\partial P_{X_a}}{\partial a}(B)=\int_B h\ dP_{X_a}\quad\text{for all}\quad B\in\mathcal{B}(\mathbb{R}).\tag{3}$$

We call h the likelihood ratio of X_a, written as $f_a/f(x|a)$.

In addition, suppose that for any $a\in(\underline{a},\bar{a})$, X_a is $\sigma(X)$-measurable and vice versa, $\sigma(X_a)=\sigma(X)$, and that for any $a\in(\underline{a},\bar{a})$, $\{X_a(B):B\in\sigma(X)\}=\{X(C):C\in\sigma(X)\}$, where X is a signal about ability when effort is not determined, and $\sigma(X)$ is the smallest σ-algebra generated by X.

The agent's reward is the market's ex post belief about ability after observing a report based on the signal. To alter the report by effort $a \in (\underline{a}, \bar{a})$, the agent privately incurs the cost of effort, $c(\cdot) : (\underline{a}, \bar{a}) \to \mathbb{R}$, which is strictly increasing and convex; $c' > 0, c'' > 0$.

The timeline is the same as in the example in Section 2.

3.1 Marginal reward

Consider the case where the authority chooses full disclosure of X_a. Then, the agent's expected reward when the market conjectures effort $\hat{a}$ is

$$\int \mathrm{E}[\theta | X_{\hat{a}} = x] dP_{X_a}(x).$$

The agent's objective function is given by

$$\int_{\mathbb{R}} \mathrm{E}[\theta | X_{\hat{a}} = x] dP_{Xa}(x) - c(a). \tag{4}$$

Assuming an interior solution, the marginal reward is given by

$$\int \mathrm{E}[\theta | X_{\hat{a}} = x] \frac{f_a}{f}(x|a) dP_{Xa}(x). \tag{5}$$

3.2 A garbled report

To define a garbled report, which includes coarse ratings such as A, B and C, we use the concept of sufficiency and measurability. First, we restate the concept of a *stochastic transformation* and *sufficiency* from Blackwell (1953).

Definition 3.1. (Stochastic transformation) Let $\mathcal{B}$, $\mathcal{C}$ be Borel fields of subsets of Ω_1, Ω_2, respectively. A stochastic transformation T is a function $Q(x, C)$ defined for all $x \in \Omega_1$ and $C \in \mathcal{C}$. For fixed C, $Q(\cdot, C)$ is a $\mathcal{B}$-measurable function of x; for fixed x, $Q(x, \cdot)$ is a probability measure on $\mathcal{C}$. For any probability measure m on $\mathcal{B}$, the function

$$M(C) = \int Q(x, C) dm(x)$$

is a probability measure on $\mathcal{C}$, denoted by Tm.

Definition 3.2. (Sufficiency) Consider two experiments, $\{m_i, i = 1, \ldots, N\}$ and $\{M_i, i = 1, \ldots, N\}$ with m_i and M_i defined on Borel fields $\mathcal{B}$ and $\mathcal{C}$ of Ω_1, Ω_2, respectively. We say that $\{m_i, i = 1, \ldots, N\}$ is sufficient for $\{M_i, i = 1, \ldots N\}$ if there is a stochastic transformation T with $Tm_i = M_i$ for all $i = 1, \ldots, N$.

In Blackwell (1951), $i = 1, \ldots, N$ denotes an unknown state of nature. This is slightly modified in our model to consider the unobservable effort $a \in (\underline{a}, \bar{a})$. Suppose that $\{m_a, a \in (\underline{a}, \bar{a})\}$ is sufficient for $\{M_a, a \in (\underline{a}, \bar{a})\}$. Then, we can consider corresponding random variables X and Z by specifying the following joint distribution. The conditional distribution of X given a is $m_a = P_{X_a}$; the conditional distribution of $Z \in C$ given a, x is $Q(x, C)$, which is a Borel-measurable function of x only. Then, M_a is the conditional distribution of Z given a, namely $M_a = P_{Z_a}$ such that

$$P_{Z_a}(C) = \int Q(x, C) dP_{X_a}(x) \tag{6}$$

holds.

In addition to the concept of sufficiency, we assume that there is a measurable function $h : \mathbb{R} \to \mathbb{R}$ such that $Z_a = h(X_a)$. By the Doob-Dynkin lemma, this holds if and only if Z_a is $\sigma(X_a)$-measurable, that is, $\sigma(Z_a) \subset \sigma(X_a)$ (see e.g. Rao and Swift, 2006). This assumption guarantees that anyone who knows the realisation of X_a also knows the rank in the ranking report for Z_a for sure. An equivalent assumption is made in DJT using the statistic $T = T(X)$ as a garbled report in (see Lemma 5.1 in Dewatripont et al., 1999).

Definition 3.3. (Garbled report) Let $\{P_{X_a}, a \in (\underline{a}, \bar{a})\}$ be sufficient for $\{P_{Z_a}, a \in (\underline{a}, \bar{a})\}$, and let Z be $\sigma(X)$-measurable. Then Z_a is a garbled report made with X_a.

Remark 3.4. Letter grading can be seen as a partition of scores, and thus is a garbled report. This is easily checked by setting $Q(x, C)$ to be an index function: For example, consider a partition $\mathcal{P} = \{A_i, i = 1, \ldots, n\}$ on $\Omega_1 = \Omega_2$, and let $Q(x, A_i) = I_{\{x \in A_i\}}, A_i \in \mathcal{P}$.

The concept of sufficiency implies the following useful lemma.

Lemma 3.5. (Martingale property) *If* $\{P_{X_a}, a \in (\underline{a}, \bar{a})\}$ *is sufficient for* $\{P_{Z_a}, a \in (\underline{a}, \bar{a})\}$, *then the martingale property:*

$$\frac{g_a}{g}(z|a) = \mathrm{E}\left[\frac{f_a}{f}(x|a)\middle|Z_a = z\right] \;\; a.e.\ [P_{Za}] \;\;\; for \;\;\; a \in (\underline{a}, \bar{a}) \tag{7}$$

holds, where $g_a/g(z|a)$ is the likelihood ratio of Z_a.

Proof. Consider $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$, and let $\Omega = \Omega_1 \times \Omega_2$, $\mathcal{F} = \mathcal{F}_1 \times \mathcal{F}_2$, $X_a(x, z) = x$ and $Z_a(x, z) = z$.

Suppose that $\{P_{X_a}, a \in (\underline{a}, \bar{a})\}$ is sufficient for $\{P_{Z_a}, a \in (\underline{a}, \bar{a})\}$. Then, the conditional distribution of $Z \in C \in \mathcal{F}_2$ given $x \in \Omega_1$, a is written as $Q(x, C)$, which is a function of x only. Let $D((z, a), B)$ be the conditional distribution of $x \in B \in \mathcal{F}_1$ given $z \in \Omega_2$, a.

Let μ_a be the unique measure on $\mathcal{F}$ determined by P_{X_a} and $Q(x, \cdot)$. Then, μ_a is determined by $D((z, a), \cdot)$ and P_{Z_a} as well, since by Fubini's theorem for $B \in \mathcal{F}_1$ and for $C \in \mathcal{F}_2$

$$\mu_a(B, C) = \int_B \int_C Q(x, dz) dP_{X_a}(x) = \int_C \int_B D((z, a), dx) dP_{Z_a}(z).$$

Applying Fubini's theorem further, the following equations hold:

$$\begin{aligned}
\int_{\{Z_a \in C\}} \frac{f_a}{f}(X_a|a) d\mu_a &= \int_\Omega \frac{f_a}{f}(X_a|a) I_{\{Z_a \in C\}} d\mu_a \\
&= \int_{\Omega_2} \int_{\Omega_1} I_C(z) \frac{f_a}{f}(X_a(x, z)|a) D((z, a), dx) dP_{Z_a}(z) \\
&= \int_C \left[\int_{\Omega_1} \frac{f_a}{f}(x|a) D((z, a), dx)\right] dP_{Z_a}(z) \qquad (8) \\
&= \int_C \mathrm{E}\left[\frac{f_a}{f}(x|a)\middle|Z_a\right] dP_{Z_a}(z) \qquad (9)
\end{aligned}$$

However, (8) can be written as

$$\begin{aligned}
\int_C \int_{\Omega_1} \frac{f_a}{f}(x|a) D((z, a), dx) dP_{Z_a}(x) &= \int_{\Omega_1} \int_C \frac{f_a}{f}(x|a) Q(x, dz) dP_{X_a}(x) \\
&= \int_{\Omega_1} \frac{f_a}{f}(x|a) Q(x, C) dP_{X_a}(x) \\
&= \frac{\partial P_{Z_a}(C)}{\partial a} \qquad (10) \\
&= \int_C \frac{g_a}{g}(z|a) dP_{Z_a}. \qquad (11)
\end{aligned}$$

Equations (10) and (11) are obtained by differentiating (3) and (6) with respect to a, respectively.

We obtain (7) from (9) and (11). $\square$

3.3 Comparison of marginal reward

Using Lemma 3.5, the difference in marginal reward under X_a and under Z_a can be written as the conditional covariance of the market belief after observing X_a and the likelihood ratio of X_a given Z_a. More precisely, we have the following result.

Lemma 3.6. *Suppose that $\{P_{X_a}, a \in (\underline{a}, \bar{a})\}$ is sufficient for $\{P_{Z_a}, a \in (\underline{a}, \bar{a})\}$, and that Z_a is $\sigma(X_a)$-measurable. Let the equilibrium effort under Z_a be a^*. Then, if the market belief after observing X_a, $\mathrm{E}[\theta|X_{\hat{a}} = x]$, and the likelihood ratio of X_a when the effort is a^*, $f_a/f(x|a^*)$, are conditionally strictly positively correlated given Z_{a^*}, then the agent will make strictly greater effort when the market observes X_a.*

Proof. Since $\{P_{X_a}, a \in (\underline{a}, \bar{a})\}$ is sufficient for $\{P_{Z_a}, a \in (\underline{a}, \bar{a})\}$, using Lemma 3.5, the marginal reward under the garbled report when the market conjectures effort $\check{a}$ is given by

$$\int \mathrm{E}[\theta|Z_{\check{a}} = z]\mathrm{E}\left[\frac{f_a}{f}(x|a)\middle|Z_a = z\right] dP_{Z_a}(z). \tag{12}$$

Since Z_a is $\sigma(X_a)$-measurable, Z_a is $\sigma(X_{\hat{a}})$- measurable as well and (12) can be written as

$$\int \mathrm{E}\Big[\mathrm{E}[\theta|X_{\hat{a}} = x]\Big|Z_{\check{a}} = z\Big]\mathrm{E}\left[\frac{f_a}{f}(x|a)\middle|Z_a = z\right] dP_{Z_a}(z), \tag{13}$$

which follows from the commutativity property of the conditional expectation operator (see e.g. Rao and Swift, 2006).

Let a^* be the perfect Bayesian equilibrium effort under Z_a. Then, the equilibrium effort is $a = \check{a} = a^*$. Setting $a = \check{a} = a^*$ and subtracting (13) from (5) yields the difference in the marginal reward under X_a and that under Z_a when the effort is $a = \check{a} = a^*$:

$$\begin{aligned}
&\int \mathrm{E}\left[\mathrm{E}[\theta|X_{\hat{a}} = x]\frac{f_a}{f}(x|a^*)\middle| Z_{a^*} = z\right] dP_{Z_{a^*}}(z) \\
&\quad - \int \mathrm{E}\left[\mathrm{E}[\theta|X_{\hat{a}} = x]\middle| Z_{a^*} = z\right] \mathrm{E}\left[\frac{f_a}{f}(x|a^*)\middle| Z_{a^*} = z\right] dP_{Z_{a^*}}(z) \\
&\quad = \int \mathrm{Cov}\left[\mathrm{E}[\theta|X_{\hat{a}} = x], \frac{f_a}{f}(x|a^*)\middle| Z_{a^*} = z\right] dP_{Z_{a^*}}(z). \qquad (14)
\end{aligned}$$

From the given hypotheses in this lemma, the sign of (14) is positive.

Let a^{**} be the perfect Bayesian equilibrium effort under X_a. Then, a^{**} maximises the agent's objective function under X_a in (4). Thus, the equilibrium marginal reward under X_a must be higher than the marginal reward under X_a when the effort is a^*. Since the marginal cost function c' is strictly increasing with respect to effort, $a^{**} > a^*$ holds. $\square$

Remark 3.7. In the construction of a garbled report, X_a and Z_a are supposed to be univariate. However, Lemma 3.6 holds if X_a and Z_a are random vectors that satisfy the martingale property.

Remark 3.8. Lemma 3.6 extends the corresponding lemma in DJT (Lemma 5.1) to include ranking reports such as letter grades. In addition, Lemma 3.6 follows from a more accurate analysis of the difference in equilibrium effort under X_a and under Z_a than Lemma 5.1 in DJT.

Using Lemma 3.9 below, the sufficient condition given in Lemma 3.6 can be simplified so that the correlation is not required to be conditional.

Lemma 3.9. (Sign of conditional covariance) *Let (Ω, Σ, P) be a probability space, and let $\mathcal{G} \subset \Sigma$ be any σ-algebra. Let X be a real valued random variable, and let g and h be two bounded increasing real functions. Then*

$$\int_G \mathrm{E}^{\mathcal{G}}[g(X)h(X)]dP - \int_G \mathrm{E}^{\mathcal{G}}[g(X)]dP \int_G \mathrm{E}^{\mathcal{G}}[h(X)]dP \geq 0 \quad G \in \mathcal{G}. \qquad (15)$$

The proof is in the Appendix.

Proposition 3.10. *Suppose that $\{P_{X_a}, a \in (\underline{a}, \bar{a})\}$ is sufficient for $\{P_{Z_a}, a \in (\underline{a}, \bar{a})\}$ and that Z_a is $\sigma(X_a)$-measurable. Let the equilibrium effort under Z_a be a^*. Then, if the market belief after observing X_a, $\mathrm{E}[\theta | X_{\hat{a}} = x]$, and the likelihood ratio of X_a when the effort is a^*, $f_a/f(x|a^*)$, are concordant functions of x, then the agent will make greater effort when the market observes X_a. Furthermore, if $\mathrm{E}[\theta|X_{\hat{a}} = x]$ and $f_a/f(x|a)$ are concordant for any effort a, then X_a will be the most motivational reporting.*

Proof. Let $\mathcal{G} = \sigma(Z_a)$. Let $h(X_a) = f_a/f(X_a|a)$, and let $g(X_a) = \mathrm{E}[\theta|X_{\hat{a}}]$, which follows from the Doob-Dynkin lemma; $\sigma(X_{\hat{a}}) \subset \sigma(X_a)$ if and only if there is a measurable function $k : \mathbb{R} \to \mathbb{R}$ such that $X_{\hat{a}} = k(X_a)$.

From Lemma 3.9, if $g(X_a)$ and $h(X_a)$ are increasing functions, then the sign of (14) is nonnegative. If $-g(X_a)$ and $-h(X_a)$ are increasing functions, then the sign of (14) is nonnegative.

Thus, if $g(X_a)$ and $h(X_a)$ are concordant for $a = a^*$, then from Lemma 3.6, the agent will make greater effort. The last part of the proposition holds because $g(X_a)$ and $h(X_a)$ are concordant for any $a \in (\underline{a}, \bar{a})$. □

Remark 3.11. The corresponding proposition in DJT (Proposition 5.2) requires the joint p.d.f. of θ and x which satisfies the concept of affiliation, and the joint p.d.f. of θ and z such that the p.d.f. of z conditional on (θ, x, a) is independent of a and θ. On the other hand, Proposition 3.10 does not specify the distribution of θ as strictly. With regard to θ Proposition 3.10 requires only the market belief after observing the signal X_a, which enable the implications to be tested more easily.

4 Conclusion

This paper studies the impact of the primary score and coarser gradings on a student's motivation to work, and provides a sufficient condition for the primary score to be the most motivational grading. Specifically, the primary score is the most motivational when a higher rank is associated with higher ability by the market, and when greater effort increases the chance to increase rank at any level of effort. If data are available, this can be tested.

Appendix

Lemma 3.9 shows that conditional covariance also satisfies Harris's inequality, which is the well known inequality for covariance.

Proof of Lemma 3.9

Proof. Let Y be another random variable which is stochastically identical to X, that is, $X =_{st} Y$, and independent of X. Then, for any $G \in \mathcal{G}$

$$\begin{aligned}
0 &\leq \int_G (g(X) - g(Y))(h(X) - h(Y))dP \\
&= \int_G g(X)h(X)dP + \int_G g(Y)h(Y)dP - \int_G g(X)h(Y)dP - \int_G g(Y)h(X)dP \\
&= 2\Big[\int_G g(X)h(X)dP - \int_G g(X)dP \int_G h(X)dP\Big] \\
&= 2\Big[\int_G \mathrm{E}^{\mathcal{G}}[g(X)h(X)]dP - \int_G \mathrm{E}^{\mathcal{G}}[g(X)]dP \int_G \mathrm{E}^{\mathcal{G}}[h(X)]dP\Big].
\end{aligned}$$

The first inequality is true since for any $\omega \in G$ either $X(\omega) \leq Y(\omega)$ or $X(\omega) \geq Y(\omega)$, and thus the two factors on the right-hand side have the same sign. From the bracket of the right-hand side of the last equation, we obtain the inequality in (15). □

References

Blackwell, David (1951) "Comparison of experiments," in *Proceedings of the second Berkeley symposium on mathematical statistics and probability*, Vol. 1, pp. 93–102.

——— (1953) "Equivalent Comparisons of Experiments," *Annals of Mathematical Statistics*, Vol. 24, No. 2, pp. 265–272.

Dewatripont, Mathias, Ian Jewitt, and Jean Tirole (1999) "The Economics of Career Concerns, Part I: Comparing Information Structures," *Review of Economic Studies*, Vol. 66, No. 1, Special Issue: Contracts, pp. 183–198.

Dubey, Pradeep and John Geanakoplos (2010) "Grading exams: 100,99,98,... or A,B,C?" *Games and Economic Behavior*, Vol. 69, No. 1, pp. 72–94.

Grossman, Sanford J. and Oliver D. Hart (1983) “An Analysis of the Principal-Agent Problem,” *Econometrica*, Vol. 51, No. 1, pp. 7–45.

Holmström, Bengt (1979) “Moral Hazard and Observability,” *Bell Journal of Economics*, Vol. 10, No. 1, pp. 74–91.

Kim, Son Ku (1995) “Efficiency of an Information System in an Agency Model,” *Econometrica*, Vol. 63, No. 1, pp. 89–102.

Lazear, Edward P. and Sherwin Rosen (1981) “Rank-Order Tournaments as Optimum Labor Contracts,” *Journal of Political Economy*, Vol. 89, No. 5, pp. 841–864.

Rao, M. M. and R. J. Swift (2006) *Probability Theory with Applications*: Springer, 2nd edition.

Rodina, David (2016) “Information Design and Career Concerns.” Northwestern University Job Market Paper.