

# TARGETS FOR HONESTY: HOW PERFORMANCE INDICATORS SHAPE INTEGRITY IN DUTCH HIGHER EDUCATION

---

SJORS OVERMAN, AGNES AKKERMAN AND RENÉ TORENVLIED

Universities across the world have recently experienced a number of serious cases of academic misconduct. In the public and academic debates, one dominant explanation exists for the fraudulent behaviour of university staff. Academic misconduct is considered to be the logical behavioural consequence of output-oriented management practices, based on performance incentives. This article puts this explanation to the test. Based on our analysis of a dataset of employees in Dutch higher education (N = 4,775) from 2010 and 2012, we argue that performance indicators have a positive impact on higher education professionals' perception of integrity in their work environment.

## INTRODUCTION

Recently, higher education institutions seem to be more prone to academic misconduct, such as falsification, fabrication, and plagiarism. Instances of severe violations of academic integrity were reported, and experts assume that these instances are only the tip of an academic iceberg (Marshall 2000; Fanelli 2009). Conservative estimates of the prevalence of such violations of integrity among scientists amount to 1–2 per cent of all published articles (Steneck 2006; Fanelli 2009). Meanwhile, the number of retracted publications has risen sharply in the last 15 years (Steen *et al.* 2013). A *Science* publication, co-authored by a leading political scientist, was retracted in May 2015 when the data were shown to be fabricated (Broockman *et al.* 2015; Carey 2015). A highly esteemed Dutch professor in social psychology fabricated his data at the kitchen table, using a vast number of top-level publications, among others in *Science* (Budd 2013).<sup>1</sup>

It is not only individual academics but also large academic institutions that are vulnerable to misconduct. Dutch institutions of higher education, for example, experienced their fair share of the phenomenon recently. Major public outrage stirred when a huge case of diploma fraud was discovered at a Dutch university of applied science in 2011. Several programmes in that university attempted to improve their diploma rate by hiring so-called 'output-improvement coaches'. These coaches initiated 'alternative programmes' with the specific aim to increase the diploma rate. The coaches allowed inadequate theses to pass, and granted credits to students for exams they had never taken (*NRC.Next* 2011). Following an official investigation, the Dutch Inspectorate of Education judged that at least 22 per cent of the diplomas in the scrutinized programmes had been awarded unjustly (Dutch Inspectorate for Education 2011).

Confronted with such extreme cases of violations of academic integrity, the general public and academia commonly respond with a singular and dominant reaction on the systemic level (Sovacool 2008). They consider the relinquishing of academic integrity norms to be a rational behavioural consequence of output-oriented management practices in higher education and academic research. Output-oriented management practices

---

Sjors Overman is at the School of Governance, Utrecht University, The Netherlands. Agnes Akkerman is at the Institute for Management Research, Radboud University, The Netherlands. René Torenvlied is at the Department of Public Administration, University of Twente, The Netherlands.

are assumed to induce perverse financial incentives (NRC.Next 2011). The Dutch professor who fabricated his data used a similar argument to explain his misconduct: he was engaged in an academic 'rat race', based on the use of quantitative performance criteria. Ultimately, a non-critical attitude of co-authors and reviewers emerged due to his successful track record. The accused professor argued that the 'contemporary university is a factory, and not a hatchery (for ideas). The modern university revolves ... around controlled production' (*De Morgen* 2013). Other researchers eagerly acknowledge that the current, performance-driven publication culture is the main cause for the lack of academic integrity (see e.g. FoliaWeb 2014; KNAW 2013).

In response to the serious cases of academic misconduct, national governments and academic interest organizations have proposed nation-wide policy measures (LaFollette 2000; Claxton 2005). The US Public Health Service Office of Research Integrity, for example, reported on the responsible conduct of research (Steneck 2003). Policy measures stimulating academic integrity make violations of academic integrity more visible: countries with integrity policies display a larger number of retractions (Fanelli *et al.* 2015).

The public debate and science policy-makers associate academic integrity and the lack thereof primarily with the use of (quantitative) performance targets in higher education. A quote from the *New York Times* is illustrative of the common discourse: 'Economists like to say there are no bad people, just bad incentives. The incentives to publish today are corrupting the scientific literature and the media that covers it. Until those incentives change, we'll all get fooled again' (Marcus and Oransky 2015). Academics themselves often refer to this argument as well. Nylenna and Simonsen stated in the *Lancet* (2006, p. 1883): 'most important is perhaps a thorough discussion of the academic system of reward and merit. How can the emphasis on productivity and the number of publications be reduced ...'. In the Netherlands, the Netherlands Academy of Arts and Sciences advised reducing the pressure on output-driven performance criteria (KNAW 2013). This resulted in the removal of 'production' as one of the main evaluation criteria in the standard evaluation protocol – used for the research assessment of the academic research programmes in the Netherlands (KNAW, VNSU and NOW 2014). The policy decision was clear and explicit: it referred to the perverse effects that the use of quantitative performance measures and targets have on academic integrity.

There is a rapidly growing body of research, with theoretical contributions to the study of academic integrity (e.g. LaFollette 2000; Sovacool 2008; Kumar 2010) as well as empirical studies that map the prevalence of scientific misconduct (for example, Martinson *et al.* 2005; Fanelli 2009; Steen *et al.* 2013). Although there is a large body of research into misconduct committed by students (for an overview see Stone *et al.* 2010), academic misconduct by faculty and staff receives far less attention (Jones and Spraakman 2011). The influence of personal and group factors also seems to be understudied, or even considered to be old-fashioned, in studies of academic integrity (Martinson *et al.* 2005). In addition, even the much broader theme of ethics in organizations focuses directly on organizations' ethical climate in explaining misconduct (Treviño *et al.* 2006).

Empirical evidence for the suggested relation between the academic system and scientific integrity is therefore hard to find (but see Martinson *et al.* 2006). In a recent study, Fanelli *et al.* (2015) tested some predictions of the 'pressure to publish' hypothesis. They show that retracted papers are less likely, and corrected papers more likely, to be authored by productive researchers. Yet, their study also shows that authors from countries coded as systems in which performance is highly incentivized do not show significantly higher retraction rates. Fanelli *et al.* (2015) conclude that 'perceived problems with pressures to

publish might be just a “shared myth”, and suggest that the explanation may be sought in the disproportionate attention paid to extreme cases of fraud. Indeed, they find that these ‘prolific retractors’ can be found among the most productive male academics.

Thus, any explanation of violations of integrity norms by the average academic may be quite different from the spectacular cases of academic fraud illuminated in the media. The present article aims to provide better insight into the theoretical and empirical relation between performance management in higher education and academic integrity, using self-reported information at the individual and group levels. The research question is: How does the use of performance targets influence academics’ integrity? Theoretically, we specify two mechanisms that explain how the use of performance targets affects academic integrity. From the explanatory mechanisms, we derive two competing hypotheses about the use of performance targets and academic integrity. One hypothesis, indeed, predicts that the use of performance indicators negatively affects academic integrity. The second hypothesis, however, predicts a reverse effect: the use of performance indicators promotes academic integrity.

Empirically, we test our hypotheses on a dataset of employees in Dutch higher education ( $N = 4,775$ ) that was collected in the years 2010 and 2012. This dataset is a sub-sample of a larger dataset collected on Dutch public sector employees and includes academics’ perceptions of integrity of others at the workplace (Den Besten *et al.* 2010). On the basis of our analyses, we reject the first hypothesis, which dominates the current discourse. Rather, in the data the use of performance indicators appears to be positively associated with academics’ self-reported perceptions of integrity in their work environment. These results have important implications for theory and practice, which we discuss in the final section of our article.

## THEORETICAL BACKGROUND

Integrity is a multifaceted concept. It is considered to be a virtue, based on professional standards (Steneck 2006, p. 56). It is assumed to reflect a positive reputation, and result in positive behaviours – both in the broader literature on ethics of science (D’Angelo 2012) and in the more specialized literature on political-administrative ethics (Butler and Cantrell 1984; De George 1993; Kernaghan 2003). Integrity is, rather than strict compliance, the disposition to be honest, accurate, and fair, and to conduct responsible research (Institute of Medicine 2002, cited in Steneck 2006, p. 56). It is the opposite of academic misconduct, which includes engaging in fabrication, falsification, or plagiarism (Fanelli 2009).

In the present article we focus on the (organizational) behavioural aspects of integrity. We conceptualize integrity as the disposition for behavioural confirmation of professional norms and principles (cf. De George 1993). Integrity in the organizational context is more specific than the concept of compliance because it refers to a person’s intention, based on a disposition. Integrity norms and regulations refer to principles of honesty, accuracy, and fairness in the profession, towards colleagues as well as clients. As with compliance, the concept of integrity compares behaviours with norms and rules, and changes in either one affect the integrity of behaviour.

Antecedents of academic integrity are discussed on three levels of analysis: the personal, the group, and the system level (Sovacool 2008). The first two narratives are far less often heard than the narrative on the system level. On a personal level, certain individual-level factors lead to deviation from the accepted norms, such as the perception of injustice

(Martinson *et al.* 2006). This is the story of the ‘few rotten apples’; an approach that is nowadays mostly abandoned (Martinson *et al.* 2005).

On an institutional level, counteracting incentive structures within organizations could facilitate misconduct, as Sovacool (2008) notes. While universities risk reputational damage, potential whistleblowers often believe that they face adverse consequences and, hence, do not report incidences of misconduct (Frankel 2000). Yet, most of the commonly discussed antecedents of scientific misbehaviour occur at the systemic level: overemphasis on publications and citations by universities, alienation of subordinates, and a replacement of cooperation by competition (Sovacool 2008). Government policy to deal with these problems mostly addresses the issue at the systemic level (Kumar 2010). The organizational science literature, however, shows the need for the integration of both personal and system level antecedents on general organizational misbehaviour (Vardi and Wiener 1996).

Performance indicators constitute a large part of what is often seen as a systemic incentive structure for academic misbehaviour. Performance indicators were introduced into the public sector as part of a broader shift towards output control (Van Dooren and Van de Walle 2008). Under the guise of New Public Management (NPM), various techniques from business environments were imported into the public sector – even though worker efficiency measurements are as old as public administration itself (Hood 1991; Pollitt and Bouckaert 2011, p. 106). Performance measurement can take the form of explicit output indicators – for example, the number of articles published, the number of citations, or minimum standards on students’ course evaluations. Performance indicators are often used as steering instruments. A specific sanction, or reward, is tied to pre-set evaluation criteria (Pollitt 2006). Examples are bonuses or a salary raise after attaining a specific target, such as the publication of an article in an A-tier journal. These performance indicators have become increasingly apparent in higher education institutions in the Netherlands and elsewhere (Teelken 2015).

A diverse set of factors have been found to influence employee behaviours in relation to professional norms. Below, we develop two alternative theoretical explanations for how the use of performance indicators influences employees’ integrity norms. The first mechanism is closely connected to incentives induced by the rat race that results from the use of performance indicators. These systemic incentives lead to a crowding out of personal intrinsic motivation within academics for academic work and academic ethics. The second mechanism develops an inverse argument. The presence of clear and explicit performance indicators facilitates the individual ownership of specific work results by academics, leading to a stronger identification with their work, and hence more incentives to perform better, including adherence to integrity norms.

## CROWDING OUT INTRINSIC MOTIVATION

Following the commonly discussed discourse, the theoretical fundamentals of academic misconduct lie in a social-psychological phenomenon called ‘the hidden cost of rewards’. These hidden costs are observed when controlled motivation replaces autonomous motivation for behaviour. Autonomous motivation reflects the satisfaction from the work itself (Deci and Ryan 1985; Frey 1994; Frey and Oberholzer-Gee 1997; Vandenabeele 2007; Jacobsen 2014). When individuals perceive a reward system introduced by external actors (such as a supervisor) to be controlling, they incur costs. Individuals perceive a reward for their autonomously motivated performance as a shift in the locus of control: from self-control to

external control. In addition, external rewards for performance could be perceived as a lack of acknowledgement of individuals' intrinsic motivation. That lack of acknowledgement would reduce their self-esteem (Frey 1994, p. 344; Frey 1998, p. 443). In effect, intrinsic motivation is reduced. Hence, the introduction of external motivators would lead to reduced performance instead of an increase in performance.

The mechanism applies to performance rewards in higher education as well. At the level of higher education institutions, an obvious example is the link between funding and diplomas. Incentives that arise from a funding system based on the number of diplomas clearly threatens the intrinsic motivation of higher education institutions for high-quality training of students. Intrinsic motivation is replaced by the motivation to deliver great quantities of diplomas regardless of the quality of the programme.

At the level of individual employees in higher education, the mechanism works in a similar fashion, even in the absence of explicit financial rewards linked to performance. The careers and research funding of academics, for example, are increasingly dependent on tangible performance indicators, such as the number of (top) publications, professor ratings in course evaluations, and citation indices. Hence, future rewards (promotion, tenure, reputation) become more dependent on specific performance measures. Since most performance measures explicitly reward quantity of output (in terms of productivity or reputation), these measures crowd out the autonomous motivation of academics to engage in high-quality scientific research and teaching.

We argue that one of the crucial dimensions of quality in academic work is scientific integrity, which we previously conceptualized as 'behaving as a "good academic" (Macfarlane *et al.* 2014). Examples of established norms of academic integrity are the proper acknowledgement of the work of others or adherence to transparent scientific (methodological) standards. Hence, if the crowding out mechanism is triggered by the introduction of performance measurement in higher education, we would expect a negative relation between the use of performance measurements and academic integrity. On the basis of the mechanism proposed by the crowding out theory, we formulate the following hypothesis.

*Hypothesis 1:* In a work context in which more explicit performance targets are formulated, employees experience less integrity in their work environment.

## PSYCHOLOGICAL OWNERSHIP AND SELF-DETERMINATION

It remains an open question whether the introduction of performance targets really leads to a crowding out of intrinsic motivation – or whether performance targets merely serve as an easy justification for bad behaviour. Other mechanisms may be at work as well and may lead to different behaviours. One possible alternative mechanism would be that performance agreements render the merit from academic work more tangible, and help academics to be more in control of their academic performance. Once academic work becomes better measurable/observable, the merits from academic achievement likewise become more tangible and comparable, allowing academics to be more in control of their own work. We call this social-psychological mechanism the 'ownership of results'.

Psychological ownership of results extends the idea of psychological ownership in organizations (Pierce *et al.* 2001). Psychological ownership is the experience of possession and the feeling of being tied to an object, which can be tangible or intangible (Pierce *et al.* 2001, p. 299). While Pierce *et al.* gear psychological ownership toward (parts of) organizations, we extend the concept to individual academic achievements or the achievements of academic teams. The object of the possession is, in this case, the results of the work conducted.



Pierce *et al.* describe that one of the routes to psychological ownership is controlling the target (2001, p. 301). In the context of the work of academic professionals, the target can be specified in terms of performance indicators. The presence of explicit standards set by performance indicators about specific dimensions of academic work (such as quantity, quality and reputation) facilitates academics' control over their achievement of results. Consequently, the presence of performance indicators enhances the sense of ownership of and identification with academic results.

The presence of extrinsic rewards does not necessarily preclude that employees' work attitudes may have an autonomous motivational basis. In contrast to the mechanism of crowding out, self-determination theory suggests that 'extrinsically motivated actions can also become self-determined as individuals identify with and fully assimilate their regulation' (Ryan and Deci 2000, p. 74). Employees may embrace the rules (Tummers 2013) and, thus, have no incentive to deviate from the rules (Torenvlied 1996). In such circumstances, performance targets may become rewarding. The reason is that adhering to targets positively affects the external reputation of an employee – as they make visible the extent to which the employee has complied with the regulation. In addition, performance targets provide employees with the opportunity to enjoy the rewards that performance targets entail (Vandenabeele 2007). Because we explicitly assume that academic integrity is a dimension of the quality of performance, the mechanism of 'ownership of results' and 'self-determination theory' would predict that employees have a lower tendency to commit fraud, or be dishonest, when working with clear performance targets. Hence, we formulate an alternative hypothesis, 2, which contrasts with our first hypothesis.

*Hypothesis 2:* In a work context in which more explicit performance targets are formulated, employees experience *more* integrity in their work environment.

## DATA AND MEASUREMENT

We use data collected in the 'POMO' personnel and mobility study by the Netherlands Ministry of the Interior and Kingdom Relations.<sup>2</sup> A web-based questionnaire was sent to a sample of 80,000+ employees within the Dutch public sector, in two waves (2010 and 2012).<sup>3</sup> This is a very large sample, given the total number of public servants in the Netherlands, amounting to approximately 850,000. The response rate varies between 34 to 40 per cent (Den Besten *et al.* 2010). For the present study we only include employees in the sector of higher education for analysis. The Dutch sector of higher education encompasses academic universities and universities of applied science, which offer higher vocational education. The total N amounts to 4,775 for the two years combined.

### Measurement

#### *Integrity*

Measuring the prevalence of academic misconduct is commonly carried out by survey research, asking scientists or higher education professionals to evaluate themselves or their environment (Greenberg and Goldberg 1994; Martinson *et al.* 2005; Fanelli 2009). In our study, integrity was measured using a five-item questionnaire probing employees' experience of their colleagues' integrity. The question for colleague integrity was purposefully and carefully designed, as direct questions to respondents about their own integrity are likely to prompt (more) social desirability bias. The integrity scale is composed of a number of items that are scored on a 5-point Likert scale, ranging from 'totally disagree' to

TABLE 1 *Integrity (reversed)*

| <b>'My colleagues ...'</b>                           | <b>Factor loading</b> |
|--|-----------------------|
| 'Do not take responsibility for their work results.' | .71                   |
| 'Are not honest about their work results.'           | .88                   |
| 'Do not do their best.'                              | .84                   |
| 'Modify the truth for their managers.'               | .88                   |
| 'Try to hide mistakes.'                              | .84                   |

Note: Principal axis factoring. Unrotated factor solution. Proportion explained variance = .69. Based on POMO survey waves 2010–12, N = 4,775.

TABLE 2 *Performance targets*

| <b>'Are the following issues discussed and did you formulate concrete agreements about them?'</b> | <b>Factor 1 loading</b> | <b>Factor 2 loading</b> |
|---|-------------------------|-------------------------|
| 'Your work results.'  | .74                     | .03                     |
| 'Work results for the coming period.'   | .85                     | -.04                    |
| 'Division of tasks within your team.'   | .45                     | .20                     |
| 'Your remuneration.'  | -.01                    | .45                     |
| 'Your personal circumstances.'  | .01                     | .54                     |
| 'Your personal development.'  | .08                     | .65                     |
| 'Your career.'  | -.05                    | .71                     |

Note: Principal axis factoring. Oblique rotated two-factor solution. Cumulative proportion explained variance: .43. Based on POMO survey waves 2010–12, N = 4,775.

'totally agree'. The questionnaire was designed to prompt answers about last year's situation in the workplace. The specific items we used for the present study, therefore, pertain to the situation just before the time of measurement. An exploratory factor analysis reveals the existence of a single-factor solution, explaining 69 per cent of the variance between the items. We saved the factor scores for each respondent using the regression mode. Table 1 presents the items and factor loadings for each of them. For the analyses, we reversed the scale so that higher scores reflect higher levels of perceived integrity.

### *Specificity of performance contracts*

The presence of performance targets and their level of detail was measured with seven items, with three response categories. The questions asked respondents to what extent specific dimensions of performance are formally discussed with their manager for performance appraisal. The response categories range from '1' (not discussed), via '2' (discussed without explicit agreement), to '3' (discussed and explicit agreement). Table 2 presents the results of an exploratory factor analysis. Clearly, two distinct factors emerge from the analysis. The pattern of factor loadings shows that a fundamental difference exists between performance targets that relate to the work content (results) and performance targets that relate to the rewards associated with results.

### **Control variables**

We control our analyses for public service motivation (PSM) (Perry 1996), which represents the individual's predisposition to serve the public good (Perry and Wise 1990). Motivational bases of employees differ, and some are naturally more focused on results and

TABLE 3 *Public service motivation*

| Item  | Factor 1 loading | Factor 2 loading | Factor 3 loading |
|---|------------------|------------------|------------------|
| Politics is a dirty word. (reversed)  | -.02             | -.06             | .75              |
| I don't care much for politicians. (reversed)   | -.13             | -.11             | .69              |
| I unselfishly contribute to my community.   | .53              | .14              | -.08             |
| Meaningful public service is very important to me.  | .69              | .19              | -.18             |
| I find it more important to make a significant contribution to society than to have personal success.               | .57              | .20              | .04              |
| Serving the public interest is an important driver for me in my daily life (in my job or elsewhere).                | .67              | .25              | -.07             |
| It is difficult for me to contain my feelings when I see people in distress.  | .17              | .67              | -.03             |
| I have little compassion for people in need who are unwilling to take the first step to help themselves. (reversed) | -.12             | -.55             | .22              |
| I consider the well-being of my fellow citizens to be very important.   | .25              | .64              | -.08             |
| If we do not show more solidarity, our society is doomed.   | .25              | .43              | .02              |

Note: Principal axis factoring. Oblique rotated 3-factor solution. Cumulative proportion explained variance: .44. Based on POMO survey waves 2010–12, N = 4,775.

motivated by measurable outcomes, while others have a disposition to work for the public interest. While PSM is not considered to be a motivation itself, individuals with high levels of PSM feel more rewarded by delivering service to the public in order to do something good for others or society (Jacobsen *et al.* 2014). It is somewhat related to intrinsic motivation, in the sense that there is no direct (future pecuniary) reward needed for motivation, even though the locus of control remains outside the self (Deci and Ryan 1985; Vandenberghe 2007). Individuals with high levels of PSM also tend to be more perceptive of the societal impact of their work (Stritch and Christensen 2014).

PSM consists of three dimensions: attraction to public policy making, commitment to the public interest, and compassion (Perry 1996; Vandenberghe and Van de Walle 2008). Some authors also add self-sacrifice as a separate dimension of PSM (Perry and Wise 1990). All three dimensions are assumed to have a positive effect on a person's integrity. The first dimension, attraction to public policy making, is a rational motive as Perry and Wise (1990, p. 368) explain. This might include the urge of self-importance, or the advocacy of an issue. Therefore, this dimension might have a less strong influence on integrity than the other two, which are norm-based and affective motives, respectively.

*Public service motivation* is measured using the shortened scale developed by Leisink and Stein (2009), as an adaptation of Perry's (1996) PSM scale for the Dutch context. Table 3 shows the results of a factor analysis exploring the dimensionality of this scale. Indeed, the three original dimensions emerge in the response patterns of the respondents in the survey. Factor 1 represents 'commitment to the public interest', factor 2 is a combination of 'compassion' and 'self-sacrifice', while factor 3 represents 'attraction to public policy making'. We saved the factor scores for all respondents for each of the factors using the regression option.

Finally, we use a number of additional predictors in our model to control for potentially confounding effects of individual traits and characteristics of employees. Gender is asked



TABLE 4 *Descriptive statistics*

|  | Min   | Max  | Mean  | SD    |
|--|-------|------|-------|-------|
| Integrity                                | -3.51 | 1.39 | 0.00  | 0.96  |
| Performance targets                      | -2.55 | 0.95 | 0.00  | 0.90  |
| Reward targets                           | -1.70 | 1.75 | 0.00  | 0.96  |
| PSM: Commitment to the public interest   | -4.17 | 2.12 | 0.00  | 0.81  |
| PSM: Attraction to public policy making  | -3.78 | 2.49 | 0.00  | 0.84  |
| PSM: Compassion                          | -1.69 | 3.30 | 0.00  | 0.83  |
| Salary (×1000 €)                         | 1.25  | 8.75 | 3.35  | 1.49  |
| Age                                      | 19    | 68   | 46.39 | 11.12 |
| Education                                |       |      |       |       |
| • Secondary or lower                     | 0     | 1    | .17   | –     |
| • Bachelor                               | 0     | 1    | .31   | –     |
| • Master                                 | 0     | 1    | .33   | –     |
| • Doctorate                              | 0     | 1    | .19   | –     |
| • Other                                  | 0     | 1    | .01   | –     |
| Staff category (Universities) *          |       |      |       |       |
| • Full professor                         | 0     | 1    | .03   | –     |
| • Associate professor                    | 0     | 1    | .03   | –     |
| • Assistant professor                    | 0     | 1    | .07   | –     |
| • PhD candidate                          | 0     | 1    | .07   | –     |
| • Post doc                               | 0     | 1    | .02   | –     |
| • Other scientific staff                 | 0     | 1    | .02   | –     |
| • Scientific support staff               | 0     | 1    | .11   | –     |
| • Support staff                          | 0     | 1    | .11   | –     |
| • Management (teaching)                  | 0     | 1    | .00   | –     |
| • Management (non-teaching)              | 0     | 1    | .01   | –     |
| Staff category (Higher vocational educ.) |       |      |       |       |
| • Management (non-teaching)              | 0     | 1    | .04   | –     |
| • Management (teaching)                  | 0     | 1    | .04   | –     |
| • Instructor                             | 0     | 1    | .28   | –     |
| • Teaching support                       | 0     | 1    | .11   | –     |
| • Support                                | 0     | 1    | .07   | –     |
| • Lector                                 | 0     | 1    | .00   | –     |
| Female                                   | 0     | 1    | .49   | –     |
| 2012**                                   | 0     | 1    | .40   | –     |

Note: Based on POMO survey waves 2010–12.

\*Dummy variables; university and higher vocational education staff combined.

\*\*Reference: 2010.

N = 4,775.

directly. Age is calculated from the year of birth, reported by the respondents. A 13-point scale was used to measure salary; each step represents a €500 increment. An 11-point scale was applied to measure the highest level of education attained, which we recoded to four levels, based on the ISCED standards (UNESCO Institute for Statistics 2012). We included a dummy variable for academic universities (1) versus universities of applied science (the reference category), as well as a dummy for survey year. Table 4 presents the descriptive statistics for these variables.

TABLE 5 *Regression analysis*<sup>4</sup>

|   | Estimate       | SE   |
|---|----------------|------|
| Constant                                  | -0.00          | 0.14 |
| Performance targets                       | <b>0.11 *</b>  | 0.02 |
| Reward targets                            | <b>0.09 *</b>  | 0.02 |
| PSM: Attraction to public policy making   | -0.01          | 0.02 |
| PSM: Compassion                           | <b>-0.12 *</b> | 0.02 |
| PSM: Commitment to the public interest    | <b>0.13 *</b>  | 0.02 |
| Age                                       | <b>0.00 *</b>  | 0.00 |
| Female                                    | 0.06           | 0.03 |
| Salary                                    | -0.00          | 0.00 |
| Education (ref.: secondary or lower)      |                |      |
| - Bachelor                                | <b>0.22 *</b>  | 0.05 |
| - Master                                  | <b>0.28 *</b>  | 0.05 |
| - Doctorate                               | <b>0.29 *</b>  | 0.07 |
| - Other                                   | 0.06           | 0.18 |
| Wave: 2012                                | 0.03           | 0.03 |
| Staff category ( $\Delta F$ ; $df = 15$ ) | <b>3.63 *</b>  |      |
| R <sup>2</sup>                            | .08            |      |
| F ( $df = 28$ ; 4,746)                    | <b>14.86 *</b> |      |

Note:

\* $p < .05$ . Dependent variable: perceived integrity in the work environment. Dummies for staff category omitted in table. Based on POMO survey waves 2010–12.  $N = 4,775$ .

## RESULTS

Table 5 presents the results of the OLS regression analysis. We observe a significant and positive effect of performance agreements on perceived integrity. Clear agreements on both output of work targets and the rewards related to those targets appear to be positively related to integrity in the direct work environment. These findings support the psychological ownership hypothesis (hypothesis 2) and refute the crowding out hypothesis (hypothesis 1). We further find that higher educated employees judge the integrity of their work environment more positively than their colleagues who completed secondary education or lower.

The levels of integrity as assessed by the respondents differ among the various academic occupations. Figure 1 presents these differences, based on quasi standard errors (Firth 2003). This method facilitates an interpretation of contrasts among groups represented by dummies in regression analyses. The most remarkable result is that full professors assess their work environment more positively with regard to integrity than other university employees, such as associate or assistant professors, or PhD candidates. The large group of teachers (28 per cent of the sample) and their support staff (11 per cent) at universities of applied sciences rate their environments as having lower integrity levels than other job categories. There is no clear tendency toward an extreme along the line of a career path in science, contrasting the findings by Martinson and colleagues (2006, p. 59), who found misbehaviour less commonly reported by early-career researchers.

The effects of the individual predisposition, measured as dimensions of PSM – one of the control variables – deserve some discussion. Attraction to public policy making has no statistically significant effects on the assessment of academic integrity. Both of the other

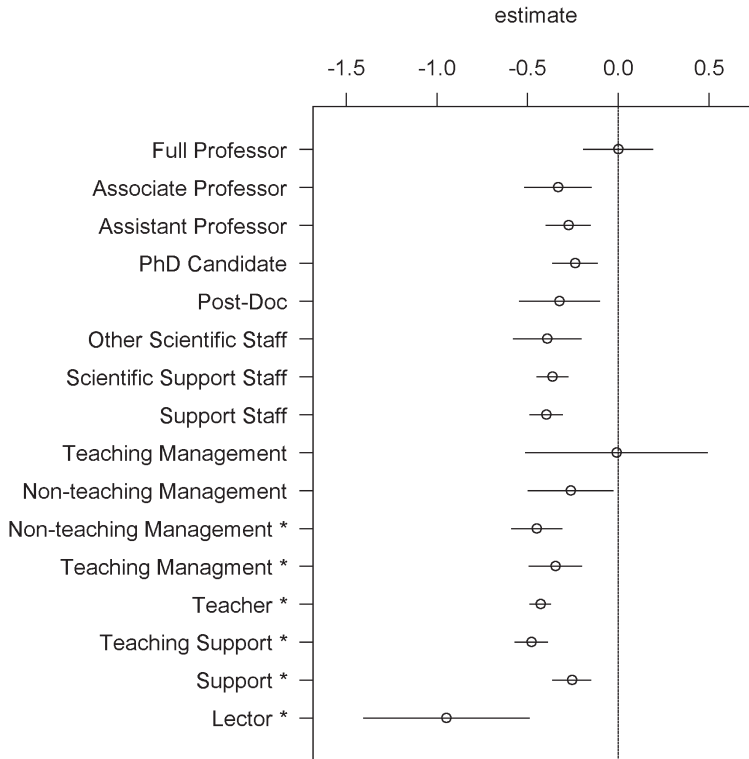


FIGURE 1 *Effect size contrasts between staff categories*

\* Universities of applied sciences. Point estimates as dots, lines represent  $\pm 2$  quasi-standard errors (Firth 2003).

dimensions show paradoxical effects. Commitment to the public interest is positively related to the perception of integrity in the work environment. This finding is in line with the findings by Brewer and Selden (1998), who found a relationship between high PSM and whistleblowing behaviour. At the same time, our findings along the compassion dimension of PSM further refine these insights. The dimension of compassion shows, contrastingly, an equally large but negative effect. This suggests that employees who have higher levels of compassion and self-sacrifice assess their academic environment as having less integrity.

## DISCUSSION AND CONCLUSIONS

In this study we aimed to establish a theoretical and empirical link between the use of performance targets in higher education and academic integrity. The recent public debate has revolved around the perverse effects of production-related reward systems in higher education. The links between output measurement, pecuniary rewards, and career perspectives of academics are assumed to have created perverse incentives in individuals. In response, policy instruments promoting performance assessment swiftly changed their focus from quantity of production to quality of production – such as meeting high-level methodological standards. While an inverse relation between the use of performance

measurement and academic integrity may be plausible intuitively, and may reflect the day-to-day sentiments of scholars, the relation is hardly based on systematic empirical research.

In this article we tested two competitive hypotheses. One hypothesis is based on the 'crowding out' mechanism which reflects the implicit mechanism dominant in the public debate. This hypothesis predicts that performance measurement acts as an extrinsic motivator and will replace intrinsic motivators for working in academics, resulting in lower levels of academic integrity. The other hypothesis, conversely, is based on the premise that clear performance agreements increase academics' feeling of 'psychological ownership' of rules. Hence, performance agreements will promote both the quantity of production and feelings of honesty.

Some limitations of this study should be addressed. First, even though we employed data from two time points, the current data are two cross-sections that do not allow us to test causal effects of changes in performance targets. Follow-up studies using panel data should further analyse our current findings. Second, we measured employees' assessment of their colleagues' integrity. Therefore, we cannot draw conclusions on actual behaviour. However, questioning respondents about their environment is less susceptible to socially desirable answering patterns and, hence, potentially less biased than asking them about their own behaviour. Future studies that use observations and behavioural experiments could overcome this limitation. Third, we do not know whether the introduction of performance targets might change the way employees reflect on results. Explicit performance rules could make clear whether certain types of behaviour are in line with integrity expectations. As such, the concept of integrity could become more salient, which could influence employees' perception. With these limitations in mind, let us turn to discussing our conclusions.

We tested the hypotheses on a dataset of employees in higher education in the Netherlands ( $N = 4,775$ ) collected in two waves in 2010 and 2012. Our results showed a clear positive effect of performance measurement and perception of integrity in employees' work environment. This supports the hypothesis based on the mechanism of psychological ownership and refutes the crowding out hypothesis. We find support for the notion that performance targets could live alongside other forms of motivation found in academic researchers and teachers.

This is not to say that extrinsic motivators cannot have overruled intrinsic motivation in the cases of fraud referred to in the introduction. Our findings show that, in general, there is not a statistical negative but a positive relation between performance management and integrity. Therefore, these results contest the justification or the explanation that is commonly used in the description of recent fraud cases that have occurred in the Dutch academic environment. The focus on output and clear targets could facilitate the acceptance of results. Employees might well be motivated in multiple ways. Performance targets do give clear benchmarks of what is meant by a good result and might facilitate the psychological ownership of these results. We should be very careful, therefore, about blaming performance targets for unethical behaviour. Governments might have been too quick in their reactions (Claxton 2005). While several policies address the reduced significance of output measures, such incentives might, actually, improve the integrity of the academic system. This empirical evidence and the explanatory analysis is valuable information that can further the, hitherto mostly theoretical or descriptive, debate on academic misconduct.

The current findings also nuance the debate on the undesirable outcomes of performance management and goal-setting. The public administration literature tends to emphasize the absence of effects or the negative effects of performance management instruments in the public sector (Bevan and Hood 2006; Hvidman and Andersen 2014). The results of the current study also show potential positive effects of performance indicators, appraisal, and rewards. Performance targets and rewards can, indeed, have a positive effect on public service employees and their work (see also Boyne and Chen 2007; Walker *et al.* 2011; Nielsen 2014). This study shows that positive effects can be found beyond effectiveness and efficiency. A traditional public value such as integrity also has a positive association with performance rewards.

Governments around the world have tried to address the recent spike in retractions of journal articles and high-profile fraud cases with measures and recommendations to lower the pressure on academics at a systemic level (Claxton 2005; Sovacool 2005). This study analysed how system-level developments, such as NPM-inspired performance targets, are associated with integrity at the personal level. We showed that both such system-level arrangements (performance targets) and individual predispositions (PSM) are associated with different levels of academic integrity. Moreover, these effects were in contrast to the ideas that are commonly employed in theory (Sovacool 2008), as well as in societal debates (Marcus and Oransky 2015). This has some important implications. First, theoretically, future studies into academic integrity should analyse antecedents at multiple levels, rather than the system level only. Second, the idea of the hidden cost of reward in universities should be critically re-evaluated. The third implication of this study is practical. While increased pressure and output orientation might have a negative effect on scientific work, it does not necessarily affect researchers' integrity in a negative way. Instead, integrity might benefit from such an orientation. Governments should be cautious and critical in developing and implementing new policies that address academic misconduct.

## ACKNOWLEDGEMENT

Data for this study were collected and graciously provided by the BZK Personeels- en Mobiliteitsonderzoek 2010, 2012 (Netherlands Ministry of the Interior and Kingdom Relations, personnel and mobility survey). We thank the Netherlands Ministry of the Interior and Kingdom Relations for providing these data.

## NOTES

<sup>1</sup> See [www.commissielevelt.nl](http://www.commissielevelt.nl).

<sup>2</sup> BZK Personeels- en Mobiliteitsonderzoek 2010, 2012. More information can be retrieved from: [www.arbeidenoverheid.nl](http://www.arbeidenoverheid.nl).

<sup>3</sup> Samples 2010: 22,446; 2012: 26,876.

<sup>4</sup> No data were available on the institutions of respondents employed in higher vocational education. Additional tests were performed on a sample including only respondents from universities to test the difference in effect sizes among universities, by using dummy variables. These analyses yielded similar results.

## REFERENCES

- Bevan, G. and C. Hood. 2006. 'What's Measured is What Matters: Targets and Gaming in the English Public Health Care System', *Public Administration*, 84, 3, 517–38.
- Boyne, G. and A.A. Chen. 2007. 'Performance Targets and Public Service Improvement', *Journal of Public Administration Research and Theory*, 17, 3, 455–77.
- Brewer, G.A. and S.C. Selden. 1998. 'Whistle Blowers in the Federal Civil Service: New Evidence of the Public Service Ethic', *Journal of Public Administration Research and Theory*, 8, 3, 413–39.

- Broockman, D., J. Kalla and P. Aronow. 2015. 'Irregularities in LaCour (2014).' [http://web.stanford.edu/~dbroock/broockman\\_kalla\\_aronow\\_lg\\_irregularities.pdf](http://web.stanford.edu/~dbroock/broockman_kalla_aronow_lg_irregularities.pdf).
- Budd, J. 2013. 'The Stapel Case: An Object Lesson in Research Integrity and its Lapses', *Synesis: A Journal of Science, Technology, Ethics, and Policy*, 4, 1, G47–G53.
- Butler, J.K. and R.S. Cantrell. 1984. 'A Behavioral Decision Theory Approach to Modeling Dyadic Trust in Superiors and Subordinates', *Psychological Reports*, 55, 1, 19–28.
- Carey, B. 2015. 'Study on Attitudes Toward Same-Sex Marriage Is Retracted by a Scientific Journal', *The New York Times*, 28 May. <http://www.nytimes.com/2015/05/29/science/journal-science-retracts-study-on-gay-canvassers-and-same-sex-marriage.html>.
- Claxton, L.D. 2005. 'Scientific Authorship: Part 1. A Window into Scientific Fraud?', *Mutation Research/Reviews in Mutation Research*, 589, 1, 17–30.
- D'Angelo, J. 2012. *Ethics in Science: Ethical Misconduct in Scientific Research*. New York: CRC Press.
- Deci, E.L. and R.M. Ryan. 1985. *Intrinsic Motivation and Self-Determination in Human Behavior*. New York: Plenum.
- De George, R.T. 1993. *Competing with Integrity in International Business*. New York: Oxford University Press.
- De Morgen. 2013. <http://www.demorgen.be/binnenland/brief-van-een-ontspoorde-wetenschapper-a1700383/>
- Den Besten, K., A. Postma and A. Sital. 2010. 'Quickscan Personeels- En Mobiliteitsonderzoek 2010.' MWM2 – bureau voor online onderzoek. <http://arbeidenoverheid.nl/media/31821/quickscan%20pomo%202010.pdf>.
- Dutch Inspectorate for Education. 2011. *Alternatieve afstudeerrichtingen en de bewaking van het eindniveau bij Hogeschool InHolland*. [Alternative graduation and the safeguarding of quality at Hogeschool in Holland]. Utrecht: Inspectie voor het Onderwijs.
- Fanelli, D. 2009. 'How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data', *PLoS ONE*, 4, 5, e5738.
- Fanelli D., R. Costas and V. Larivière. 2015. 'Misconduct Policies, Academic Culture and Career Stage, Not Gender or Pressures to Publish, Affect Scientific Integrity', *PLoS ONE*, 10, 6, e0127556.
- Firth, D. 2003. 'Overcoming the Reference Category Problem in the Presentation of Statistical Models', *Sociological Methodology*, 33, 1, 1–18.
- FoliaWeb. 2014. <http://www.foliaweb.nl/wetenschap/wetenschapsfraude-%E2%80%98het-probleem-ligt-bij-publicatieruk%E2%80%99/>.
- Frankel, M.S. 2000. 'Scientific Societies as Sentinels of Responsible Research Conduct', *Experimental Biology and Medicine*, 224, 4, 216–19.
- Frey, B.S. 1994. 'How Intrinsic Motivation is Crowded Out and In', *Rationality and Society*, 6, 3, 334–52.
- Frey, B.S. 1998. 'Institutions and Morale: The Crowding-Out Effect', in A. Ben-Ner and L. Putterman (eds), *Economics, Values, and Organization*. Cambridge: Cambridge University Press, pp. 437–60.
- Frey, B.S. and F. Oberholzer-Gee. 1997. 'The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding-Out', *American Economic Review*, 87, 4, 746–55.
- Greenberg, M. and L. Goldberg. 1994. 'Ethical Challenges to Risk Scientists: An Exploratory Analysis of Survey Data', *Science, Technology & Human Values*, 19, 2, 223–41.
- Hood, C. 1991. 'A Public Management for All Seasons?', *Public Administration*, 69, 1, 3–19.
- Hvidman, U. and S.C. Andersen. 2014. 'Impact of Performance Management in Public and Private Organizations', *Journal of Public Administration Research and Theory*, 24, 1, 35–58.
- Institute of Medicine (US). Committee on Assessing Integrity in Research Environments, National Research Council (US), et al. (2002). *Integrity in Scientific Research: Creating an Environment that Promotes Responsible Conduct*. Washington, DC: The National Academies Press.
- Jacobsen, C.B., J. Hvitved and L.B. Andersen. 2014. 'Command and Motivation: How the Perception of External Interventions Relates to Intrinsic Motivation and Public Service Motivation', *Public Administration*, 92, 4, 790–806.
- Jones, J. and G. Spraakman. 2011. 'A Case of Academic Misconduct: Does Self-Interest Rule?', *Accounting Perspectives*, 10, 1, 1–22.
- Kernaghan, K. 2003. 'Integrating Values into Public Service: The Values Statement as Centerpiece', *Public Administration Review*, 63, 6, 711–19.
- KNAW (Netherlands Academy of Arts and Sciences). 2013. *Vertouwen in Wetenschap* [Trust in Science]. For an English summary: <http://www.knaw.nl/nl/actueel/publicaties/vertrouwen-in-wetenschap>
- KNAW, VSNU and NOW. 2014. *Standard Evaluation Protocol 2015–2021. Protocol for Research Assessments in the Netherlands*. <http://www.vsnunl.nl/files/documenten/Domeinen/Onderzoek/SEP2015-2021.pdf>
- Kumar, M.N. 2010. 'A Theoretical Comparison of the Models of Prevention of Research Misconduct', *Accountability in Research*, 17, 2, 51–66.
- LaFollette, M.C. 2000. 'The Evolution of the "Scientific Misconduct" Issue: An Historical Overview', *Experimental Biology and Medicine*, 224, 4, 211–15.
- Leisink, P. and B. Steijn. 2009. 'Public Service Motivation and Job Performance of Public Sector Employees in the Netherlands', *International Review of Administrative Sciences*, 75, 1, 35–52.



- Macfarlen, B., J. Zhang and A. Pun. 2014. 'Academic Integrity: A Review of the Literature', *Studies in Higher Education*, 39, 2, 339–58.
- Marcus, A. and I. Oransky. 2015. 'What's Behind Big Science Frauds?', *The New York Times*, 22 May. <http://www.nytimes.com/2015/05/23/opinion/whats-behind-big-science-frauds.html>.
- Marshall, E. 2000. 'How Prevalent is Fraud? That's a Million-Dollar Question', *Science*, 290, 5497, 1662–63.
- Martinson, B.C., M.S. Anderson and R. de Vries. 2005. 'Scientists Behaving Badly', *Nature*, 435, 7043, 737–38.
- Martinson, B.C., M.S. Anderson, A.L. Crain and R. De Vries. 2006. 'Scientists' Perceptions of Organizational Justice and Self-Reported Misbehaviors', *Journal of Empirical Research on Human Research Ethics*, 1, 1, 51–66.
- Nielsen, P.A. 2014. 'Performance Management, Managerial Authority, and Public Service Performance', *Journal of Public Administration Research and Theory*, 24, 2, 431–58.
- NRC.next. 2011. 17 November, 13.
- Nylenna, M. and S. Simonsen. 2006. 'Scientific Misconduct: A New Approach to Prevention', *The Lancet*, 367, 9526, 1882–84.
- Perry, J.L. 1996. 'Measuring Public Service Motivation: An Assessment of Construct Reliability and Validity', *Journal of Public Administration Research and Theory*, 6, 1, 5–22.
- Perry, J.L. and L.R. Wise. 1990. 'The Motivational Bases of Public Service', *Public Administration Review*, 50, 3, 367–73.
- Pierce, J.L., T. Kostova and K.T. Dirks. 2001. 'Toward a Theory of Psychological Ownership in Organizations', *Academy of Management Review*, 26, 2, 298–310.
- Pollitt, C. 2006. 'Performance Management in Practice: A Comparative Study of Executive Agencies', *Journal of Public Administration Research and Theory*, 16, 1, 25–44.
- Pollitt, C. and G. Bouckaert. 2011. *Public Management Reform: A Comparative Analysis: New Public Management, Governance, and the Neo-Weberian State*, 3rd edn. Oxford and New York: Oxford University Press.
- Ryan, R.M. and E.L. Deci. 2000. 'Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being', *American Psychologist*, 55, 1, 68–78.
- Sovacool, B.K. 2005. 'Using Criminalization and Due Process to Reduce Scientific Misconduct', *American Journal of Bioethics*, 5, 5, W1–W7.
- Sovacool, B.K. 2008. 'Exploring Scientific Misconduct: Isolated Individuals, Impure Institutions, or an Inevitable Idiom of Modern Science?', *Journal of Bioethical Inquiry*, 5, 4, 271–82.
- Steen, R.G., A. Casadevall and F.C. Fang. 2013. 'Why Has the Number of Scientific Retractions Increased?', *PLoS ONE*, 8, 7, e68397.
- Steneck, N.H. 2003. *ORI Introduction to the Responsible Conduct of Research*. Washington, DC: Government Printing Office.
- Steneck, N.H. 2006. 'Fostering Integrity in Research: Definitions, Current Knowledge, and Future Directions', *Science and Engineering Ethics*, 12, 1, 53–74.
- Stone, T.H., I.M. Jawahar and J.L. Kisamore. 2010. 'Predicting Academic Misconduct Intentions and Behavior Using the Theory of Planned Behavior and Personality', *Basic and Applied Social Psychology*, 32, 1, 35–45.
- Stritch, J.M. and R.K. Christensen. 2014. 'Looking at a Job's Social Impact through PSM-Tinted Lenses: Probing the Motivation-Perception Relationship', *Public Administration*, 92, 4, 826–42.
- Teelken, C. 2015. 'Hybridity, Coping Mechanisms, and Academic Performance Management: Comparing Three Countries', *Public Administration*, 93, 2, 307–23.
- Torenvlied, R. 1996. 'Political Control of Implementation Agencies: Effects of Political Consensus on Agency Compliance', *Rationality and Society*, 8, 1, 25–56.
- Treviño, L.K., G.R. Weaver and S.J. Reynolds. 2006. 'Behavioral Ethics in Organizations: A Review', *Journal of Management*, 32, 6, 951–90.
- Tummers, L. 2013. *Policy Alienation and the Power of Professionals Confronting New Policies*. Cheltenham: Edward Elgar.
- UNESCO Institute for Statistics. 2012. *International Standard Classification of Education: ISCED 2011*. Montreal, Quebec: UNESCO Institute for Statistics.
- Vandenabeele, W. 2007. 'Toward a Public Administration Theory of Public Service Motivation', *Public Management Review*, 9, 4, 545–56.
- Vandenabeele, W. and S. van de Walle. 2008. 'International Differences in Public Service Motivation: Comparing Regions across the World', in J.L. Perry and A. Hondeghem (eds), *Motivation in Public Management: The Call of Public Service*. Oxford and New York: Oxford University Press, pp. 223–44.
- Van Dooren, W. and S. van de Walle. 2008. *Performance Information in the Public Sector: How It Is Used*. Basingstoke: Palgrave Macmillan.
- Vardi, Y. and Y. Wiener. 1996. 'Misbehavior in Organizations: A Motivational Framework', *Organization Science*, 7, 2, 151–65.
- Walker R.M., F. Damanpour and C.A. Devece. 2011. 'Management Innovation and Organizational Performance: The Mediating Effect of Performance Management', *Journal of Public Administration Research and Theory*, 21, 2, 367–86.