

# Behaviour of Finance Professionals Under the Balanced Scorecard

## Elizabeth Sheedy

Associate Professor Elizabeth Sheedy is a specialist in financial risk management based at the Macquarie Applied Finance Centre.

Email: [esheedy@mafc.mq.edu.au](mailto:esheedy@mafc.mq.edu.au)

## Lyla Zhang

Dr Le (Lyla) Zhang is an experimental economist at the Macquarie Graduate School of Management.

Email: [lyla.zhang@mgs.m.edu.au](mailto:lyla.zhang@mgs.m.edu.au)

## Dominik Steffan

Dominik Steffan is a visiting scholar from Bavaria, enrolled in the elite Finance and Information Management program.



## Contents

<b>Executive Summary</b> .....	<b>3</b>
<b>Introduction/Background</b> .....	<b>4</b>
<b>Experiment (1)</b> .....	<b>6</b>
<b>Experiment (2)</b> .....	<b>15</b>
<b>References</b> .....	<b>18</b>

### List of Tables and Figures

<b>Table 1: Experiment (1) Groups</b> .....	<b>6</b>
<b>Table 2: Balanced Scorecard (1)</b> .....	<b>7</b>
<b>Table 3: Payment Schedule with Behaviour Gateway (1)</b> .....	<b>7</b>
<b>Table 4: Fixed Remuneration (No Upside) Payment Schedule</b> .....	<b>7</b>
<b>Table 5: Sample Demographics for Finance Professionals</b> .....	<b>9</b>
<b>Table 5: Sample Demographics for Finance Professionals (Cont'd)</b> .....	<b>10</b>
<b>Table 6: Attitudes to Remuneration</b> .....	<b>11</b>
<b>Table 7: Treatment Analysis (1)</b> .....	<b>12</b>
<b>Table 8: Other influences on Behaviour</b> .....	<b>13</b>
<b>Figure 1: Compliance Rates Through the Experiment</b> .....	<b>14</b>
<b>Table 9: Sample Demographics for Students</b> .....	<b>16</b>
<b>Table 10: Treatment Analysis (2)</b> .....	<b>17</b>

This paper provides a summary of research conducted in 2018. A more detailed paper with full statistical analyses and reference list can be obtained by contacting Elizabeth Sheedy. Be sure to 'Link In' with the authors to receive ongoing research updates.

This version: November 2018  
© Macquarie University

# Executive Summary

---

The Royal Commission into Misconduct in Banking, Insurance and Financial Services has refocused attention on how remuneration practices influence behaviour. In particular the ‘Balanced Scorecard’, seen by some as a cure for misconduct, has come under scrutiny. In 2018 experimental researchers based at Macquarie University have investigated how various remuneration structures affect compliance with company policy.

The study concluded that the highest rates of compliance were achieved under a fixed remuneration structure; 75% of participants were completely compliant with policy across all transactions. Under a simplified Balanced Scorecard, compliance rates fell significantly to 62%. Under the compliance gateway, compliance rates fell further to 51% of participants. This suggests that the Balanced Scorecard is significantly less effective than fixed remuneration with regard to compliance outcomes.

Interestingly, the Balanced Scorecard we tested did not lead to any significant uplift in productivity relative to fixed remuneration. This may be because the decision to

breach, or not breach policy slows participants down as they weigh up the chance of being caught and possible consequences. They must also consider issues such as moral identity and social standing, further slowing down the speed of mental processing. Under fixed remuneration, there is less need to consider such issues and so the often-claimed loss in productivity is not observed.

With these findings, it seems appropriate for the financial services industry to reconsider the use of the Balanced Scorecard for remuneration purposes.



# Introduction/ Background

---

In 2018, researchers at Macquarie conducted a new experimental study to investigate the effectiveness of the popular Balanced Scorecard in financial institutions. The aim of this study was to better understand the usefulness of the Balanced Scorecard for ensuring compliance with company policies.

The study was supported by four industry partners. Our funding partners were Deloitte Australia and the Insurance Council of Australia; in addition, these partners provided advice and input to the experimental design to help ensure industry relevance. Our partners for participant recruitment were The Australian and New Zealand Institute of Insurance and Finance (ANZIIF) and Financial Services Institute of Australasia (FINSIA). We greatly appreciate the role of all four partners in generously supporting industry-relevant academic research which is conducted to the most rigorous standards.

Why is investigation of the Balanced Scorecard of interest? The Balanced Scorecard is a system of performance measurement used when employees are expected to complete multiple, competing tasks. In a financial institution, employees are often expected to generate sales/profits while also complying with company policy. Policies could be designed to limit credit risk, underwriting risk, market risk and operational risks for the firm. This last category includes the possibility of legal/reputational costs flowing from, for example, misconduct toward customers. At times these aims of profit generation and compliance are in conflict, at least in the short-term; it is sometimes necessary to reject profitable opportunities in order to comply with company policy.

The concept of the Balanced Scorecard was introduced by Kaplan and Norton (1992). The original version expanded on mere financial performance measures and incorporated customer satisfaction, internal business processes and innovation and learning. Subsequently Kaplan and Norton emphasised that scorecard measures should depend on the organisation's strategy. In other words, the criteria chosen to measure performance should vary depending on what the organisation is trying to achieve. In an Australian financial institution, the typical Balanced Scorecard has multiple performance criteria of which some are financial (e.g. sales/profits) but others are non-financial (e.g. customer satisfaction, compliance with policy, behaviour consistent with company values), as shown in APRA (2018).





In 2016, the Australian Banking Association (ABA) commissioned Stephen Sedgwick to conduct a review of remuneration practices in retail banking. The final report, released in 2017, recommended that variable remuneration should not be based solely on financial measures, but on a broader set of measures (Sedgwick, 2017, see recommendation 3). This recommendation appears to be based primarily on guidance from the UK conduct regulator (FSA, 2013, at p. 20). The same FSA guidance, however, highlights that the success of the Balanced Scorecard is likely to depend on the importance given to non-sales criteria and whether those criteria reflect the fair treatment of customers. Commonly used and simplistic measures of customer satisfaction, such as the net promoter score<sup>1</sup>, may not achieve this. Further, a subsequent report from the UK conduct regulator documented that the increasing use of the Balanced Scorecard had not been as successful as had been hoped; this was attributed to the role of managerial discretion and biased ratings for some performance criteria (FCA, 2014, at p. 13).

When the Balanced Scorecard is implemented in a financial institution, it sometimes includes criteria that are difficult to judge objectively. An example of such a criterion would be 'Behaviour is consistent with organizational values' where the manager applies a rating. There is often doubt as to whether these ratings are credible; previous researchers have documented a range of problems with subjective management ratings. Centrality bias is the tendency to give all employees a similar rating, despite variation in performance, while leniency bias is the tendency to give higher performance ratings than are warranted. Bol (2011) shows that managers' own incentives and preferences can influence their ratings. Due to the large amount of managerial discretion in the ratings, and the desire to retain top performers in sales/profits, it is likely that managers may give a high rating to such staff despite poor behaviour. Even more worrying, subjective performance ratings can be prone to favouritism, collusion and extortion (Delfgaauw and Souverijn, 2016).

Not only is there a potential problem due to subjectivity, but bad behaviour often goes undetected (especially in the short-term) thus further limiting the usefulness of the rating. The difficulty of perfectly monitoring compliance behaviour, and hence the tendency for compliance ratings to be overstated, is a major challenge for the Balanced Scorecard.

If we have multiple performance criteria in a Balanced Scorecard, some objective and some subjective or measured imperfectly, what is the likely outcome? The economics literature suggests that in such a case staff will tend to focus on the objective criteria which are usually things like sales and profits (Holmstrom and Milgrom, 1991). In other words, the balance in the Balanced Scorecard goes out the window.

A recent study by Danish researchers Albertson and Lueg reviewed 117 empirical studies of the Balanced Scorecard in leading academic journals that were published between 1992 and 2012. Overall, the authors claim that the "relevance" of the Balanced Scorecard system for remuneration remains unproven.

The Royal Commission into Misconduct in the Banking, Superannuation and Financial Services Industry (Royal Commission) has identified many examples of misconduct resulting in bad outcomes for customers. According to the Interim Report (Royal Commission, 2018, p. 301), 'the culture and conduct of the banks was driven by, and was reflected in, their remuneration practices and policies'. Chapter 9 contains extensive discussion of the Balanced Scorecard, detailing the design of these systems at ANZ and Westpac. The main finding here is that in practice, and despite appearances to the contrary, Balanced Scorecards continue to be dominated by sales criteria with minimal importance accorded to other criteria.



<sup>1</sup> The Net Promoter Score (NPS) is calculated based on responses to a single question: *How likely is it that you would recommend our company/product/service to a friend or colleague?* The scoring for this answer is most often based on a 0 to 10 scale. [4] Those who respond with a score of 9 to 10 are called Promoters; those who respond with a score of 0 to 6 are labeled Detractors. The Net Promoter Score is calculated by subtracting the percentage of customers who are Detractors from the percentage of customers who are Promoters. In the field of financial services, misconduct toward customers may not be detected for months, years or ever. In this context NPS may not be a reliable indicator for fair treatment of customers.

# Experiment (1)

The study focused specifically on how the Balanced Scorecard works to promote behaviour that is compliant with policy, in an environment where many policy breaches are not always identified in the short term. In other words, compliance is imperfectly measured. In the simulated work environment of the laboratory, 20% of transactions were audited for policy compliance, meaning that many violations of policy went unnoticed by the hypothetical ‘audit office’. As researchers, however, we were able to observe all violations of policy, something that is not achievable in the field.

The experiment was designed to mimic the typical financial services setting, where employees are under pressure to complete transactions for the benefit of shareholders. In a session lasting 20 minutes, participants could complete up to 60 transactions. The participants had to do some simple analysis (with a calculator) and then decide whether to complete the transaction or reject it. Just like any financial services workplace, we provided participants with rules to be followed during the experiment. These rules are designed to prevent transactions considered to be too risky by the employer. Of the 60 transactions, many would result in a breach of policy if completed. Consistent application of the policy would mean rejecting many transactions, potentially reducing performance rankings on the dimension of transaction volumes.

We tested five different treatments as set out in [Table 1](#). In treatment 1 (Balanced Scorecard, No peer information provided) participants received a cash payment at the

end of the session that depended on performance during the session, subject to a minimum payment of \$50. We tested a simplified ‘Balanced Scorecard’ which was partly determined by success in completing transactions, and partly based on behaviour i.e. whether or not it was consistent with risk policy. Compliance had a weighting of around 20%, slightly above typical industry practice (see APRA, 2018, p. 17). To achieve the maximum payment of \$300, participants had to be among the 10% most productive participants and achieve a perfect behaviour rating with zero observed policy breaches. As noted previously, the relatively high weighting to transactions is consistent with industry practices noted by the Royal Commission.

In all treatments participants received feedback during the experiment about their own performance. After evaluating 15, 30, 45 and 60 transactions, they were informed if they had been observed breaching policy by the hypothetical ‘audit office’. They were also informed of the number of transactions they had completed.

**Table 1: Experiment (1) Groups**

	Balanced Scorecard	Behaviour Gateway	Fixed (No Upside) Remuneration
No peer information provided	Treatment 1	Treatment 2	Treatment 3
Peer information provided	Treatment 4	-	Treatment 5

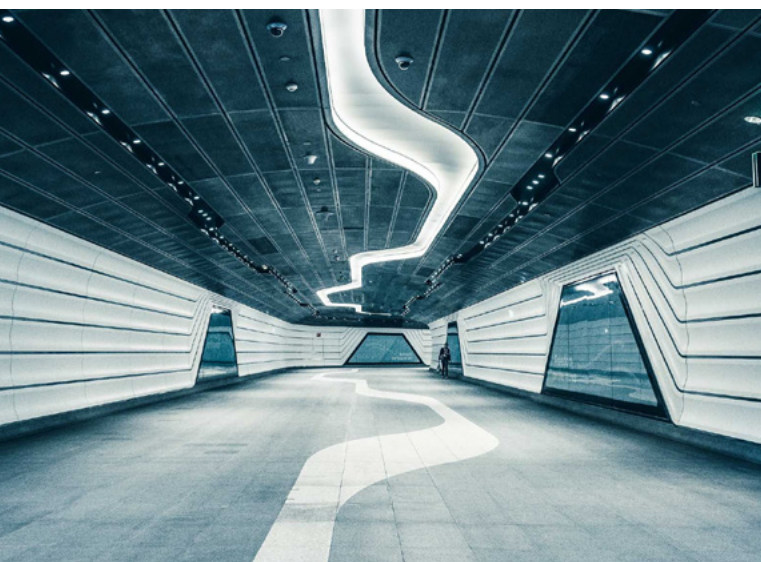
In treatment 2, we tested a behaviour ‘gateway’, whereby participants were excluded from receiving more than the minimum payment (\$50) if they were observed violating policy on three or more occasions.

These practices were compared to a ‘fixed’ remuneration structure (treatment 3). Payment was fixed in the sense that there was no financial benefit associated with violating policy, but there were penalties for observed breaches.

Treatments 4 and 5 were similar to treatments 1 and 3 respectively, with one important difference: after the first ten minutes, participants received information about the performance of their peers. They were told their transaction ranking i.e. performance relative to the other participants in the same session. They were also informed about the proportion of peers that have been found, by the ‘audit office’, to have breached policy in similar sessions.

We tested these treatments (4 and 5) with information about peer performance because recent experimental research has highlighted that ranking effects can be important determinants of behaviour. Kirchler, Linder and Weitzel, 2018, examined the behaviour of finance professionals with regard to rank incentives where there was no direct monetary reward for good performance. That study found that simply displaying a non-incentivised ranking activates concerns about relative performance and leads to high levels of risk-taking. This could be driven by the desire either for a positive self-image or enhanced reputation among peers and/or managers.

Understanding these ranking effects is important, particularly in the context of possible new initiatives and/or regulations to replace variable remuneration with fixed remuneration. There is a risk that eliminating variable remuneration may not improve behaviour if ranking effects prove to be significant.



**Table 2: Balanced Scorecard (1)**

Transactions	Behaviour
Number Transacted Maximum \$240	Consistent with Risk Policy Maximum \$60
Outstanding: \$240 (top 10% of participants)	Perfect i.e. zero observed policy breaches: \$60
Exceeds Requirements: \$80 (30% of participants)	Near Perfect i.e. one or two observed policy breaches: \$10
Meets Requirements: \$40 (40% of participants)	
Needs Improvement: zero (lowest 20% of participants)	Needs Improvement: Three or more observed policy breaches: zero

**Table 3: Payment Schedule with Behaviour Gateway (1)**

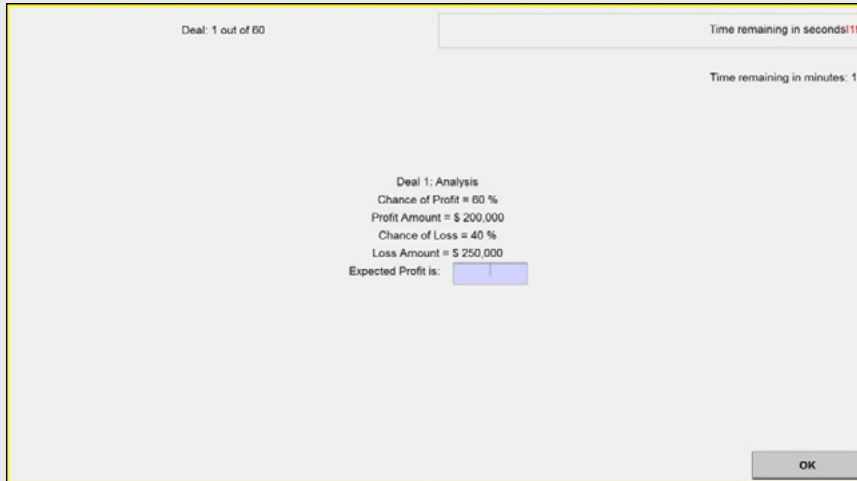
Transactions	Behaviour Gateway
Number Transacted	Consistent with Risk Policy
Outstanding: \$300 (top 10% of participants)	If fewer than 3 observed policy breaches
Exceeds Requirements: \$120 (30% of participants)	
Meets Requirements: \$80 (40% of participants)	
Needs Improvement: \$50 (lowest 20% of participants)	If 3 or more observed policy breaches,
\$50 regardless of number of transactions	

**Table 4: Fixed Remuneration (No Upside) Payment Schedule**

Policy Breaches Observed by Audit Office	Behaviour Rating	Payment
Zero	Perfect	\$110
One or Two	Near Perfect	\$60
Three or More	Needs Improvement	\$50

**Example**

This example illustrates the analysis that participants in the experiment were expected to perform. The risk limit (for the Loss Amount) is \$200,000.



The transaction has 60% chance to gain \$200,000 and 40% chance to lose \$250,000. So the expected profits can be easily calculated:  $60\% \times 200,000 - 40\% \times 250,000 = 20,000$

This transaction violates risk policy since the loss amount of \$250,000 is more than the specified limit of \$200,000.

At the start of the experiment we asked participants to provide their perceptions of the likely compliance behaviour in the laboratory. This was done after participants had read the instructions pertinent to their treatment and had correctly answered a set of comprehension questions to confirm their understanding.

**Perception Question:**

In the experiment, some transactions will violate the risk policy. Some participants may choose to transact despite this. In your opinion, what percentage of your fellow participants in the experiment would ALWAYS follow risk policy (i.e. not transact if outside the risk policy)? (Enter X%).

We consider that the answer to this question is a good measure of perceived behavioural norms for the experiment i.e. a measure of ‘culture’. We expected that this measure would predict how individuals behaved during the experiment. In other words, people, who crave social acceptance, will tend to behave in a manner consistent with perceived behavioural norms.

At the end of the experiment we asked participants to complete a short survey so we could understand and control for demographics, attitudes etc. The survey included measures of:

- **Individual risk tolerance.** We expected that people who are more personally risk tolerant would be more likely to breach policy, taking the risk of being observed and receiving a lower behaviour rating, in order to boost transactions.
- **Preferences for relative performance.** We expected that people who are very concerned about their performance relative to others would be more likely to be influenced by the information about peer performance.
- **Moral disengagement.** Moral disengagement refers to a set of cognitive mechanisms that some people use to enable them to behave unethically without feeling distress (Moore et. al, 2012). Research has suggested that moral disengagement is a good predictor of unethical behaviour in the workplace. Some of these mechanisms are: moral justification (e.g. Playing dirty is sometimes necessary in order to achieve noble ends); advantageous comparison (e.g. Considering the ways people grossly misrepresent themselves, it’s hardly a sin to inflate your own credentials a bit); displacement of responsibility (e.g. People shouldn’t be held accountable for doing questionable things when they were just doing what an authority figure told them to do); dehumanisation (e.g. Some people have to be treated roughly because they lack feelings that can be hurt) etc. We expected that people who are more prone to moral disengagement would be more likely to violate policy in order to boost transactions.



## Recruitment of Participants

With assistance from two professional associations, we recruited 318 finance professionals to participate in this study. Members of ANZIIF (The Australian and New Zealand Institute of Insurance and Finance) made up 59% of the sample while 21% were members of FINSIA (Financial Services Institute of Australasia). Those who attended were invited to refer their work

colleagues which provided some participants who were not members of professional associations. The sole criterion for participation was employment in the financial services industry for a minimum of six months. Participants were required to spend approximately one hour in a temporary experimental laboratory located at Macquarie University's city campus at 123 Pitt St, Sydney. Sample demographics are presented in [Table 5](#).

**Table 5: Sample Demographics for Finance Professionals**

	Percentage of Responses (Number of valid answers)
<b>Gender</b>	<b>(318)</b>
Male	60.4%
Female	39.6%
<b>Age</b>	<b>(318)</b>
Less than 25 years	6.0%
25-34 years	34.6%
35-44 years	25.2%
45-54 years	22.3%
55 years and above	12.0%
<b>Industry Tenure</b>	<b>(318)</b>
6 months to less than 1 year	1.6%
1 year to less than 3 years	9.4%
3 years to less than 5 years	11.3%
5 years to less than 10 years	20.8%
10 years to less than 15 years	12.9%
15 years to less than 25 years	26.7%
More than 25 years	17.3%
<b>Membership in a Professional Association</b>	<b>(318)</b>
ANZIIF	59.1%
FINSIA	20.8%
Other	4.7%
No	15.4%
<b>Finance Industry Segment</b>	<b>(318)</b>
Banking and Finance	16.0%
Insurance	(177) 55.7%
Of which:	
a) General	76.8%
b) Life	6.2%
c) Other	17.0%
Superannuation (Pension Funds)	2.2%
Broking	8.8%

**Table 5: Sample Demographics for Finance Professionals (Cont'd)**

	<b>Percentage of Responses (Number of valid answers)</b>
Financial Planning/Wealth Management	4.1%
Funds Management	4.4%
Consulting	2.5%
Professional Services	2.8%
Others	3.5%
<b>Seniority</b>	<b>(318)</b>
Team member / front-line employee	13.2%
Professional employee (but not a manager)	43.7%
Team Leader	7.2%
Middle Management	11.0%
Report to Senior Management	10.7%
Senior Management	14.2%
<b>What best describes your role (also known as the three Line of Defence model)?</b>	<b>(318)</b>
Line 1 (Business)	58.5%
Line 2 (Independent/Specialist Risk Manager, including Compliance)	21.4%
Line 3 (Internal Audit/Assurance)	1.6%
Do not know	18.6%

## Results

The results of the survey on attitudes to remuneration structures are presented in [Table 6](#). It is clear that more than half of the participants in this study are eligible for variable remuneration in their usual workplace. Of these, around two-thirds have their performance assessed with a balanced scorecard, highlighting the importance of research on this topic. Of the people working under a balanced scorecard, the majority have a compliance gateway of some sort. Finally, the balanced scorecard enjoys popularity with most participants preferring this system of performance measurement linked to variable remuneration.

**Table 6: Attitudes to Remuneration**

<b>In your everyday workplace, are you eligible for variable remuneration?</b>	<b>(318)</b>
No	47.5%
Yes	(167) 52.5%
<b>I) The following proportions of your remuneration are at risk (i.e. not fixed)</b>	
a) 0-10%	50.9%
b) 10-30%	32.3%
c) 30-50%	10.2%
d) More than 50%	6.6%
<b>II) Is your variable remuneration determined with reference to a balanced scorecard</b>	
Unsure	10.8%
No	21.6%
Yes	(113) 67.7%
<b>I) In the balanced scorecard that applies to you, what percentage is comprised of risk/compliance/behavioural matters?</b>	
Less than or equal to 10%	15.0%
10-20%	19.5%
20-30%	16.8%
30-50%	21.2%
Unsure	27.4%
<b>II) In the balanced scorecard that applies to you, is there a risk or compliance or behaviour gateway? That is, would a poor rating for risk/compliance/behaviour make you ineligible for variable remuneration?</b>	
No	8.9%
Yes	69.9%
Unsure	21.2%
<b>Which of the following remuneration systems would you prefer to work under?</b>	<b>(318)</b>
Base salary plus variable remuneration based on a balanced scorecard	64.5%
Base salary plus variable remuneration based on financials only	11.0%
Fixed salary	22.6%
Other	1.9%

**Table 7** shows how behaviour varied across the different remuneration and peer information treatments.

Unsurprisingly, compliance rates were consistently higher under fixed remuneration where there was no financial benefit for breaching policy. It is interesting to note, however, that compliance rates fell well short of 100% even when there was no upside. One possible explanation relates to errors i.e. participants sometimes breach policy in error. Another possible explanation is that some participants feel intrinsic pressure to boost transaction volumes despite the fact that there is no financial reward for doing so.

We observed significantly lower rates of compliance in the treatment with a behaviour gateway. This is the treatment having the lowest perceptions of compliance (see row g) i.e. participants predicted at the start of the experiment that compliance would be lower than for the other treatments.

The effect of information about relative performance appears not to be a significant determinant of behaviour. This is an important finding for those people who may be concerned about the use of non-monetary rank incentives and how these might encourage misconduct even in the absence of variable remuneration

Productivity across treatments was remarkably similar. Although the number of compliant transactions was a little lower under fixed remuneration, the difference was not enough to be statistically significant. This finding may be surprising to some, but it is not entirely at odds with other research in this area. The relationship between incentives and performance is complex, depending on the kind of task, the individual and the work environment. Variable remuneration is most effective for enhancing performance in very simple situations e.g. one performance criterion only that can be objectively and easily measured.

In this situation, participants in the Balanced Scorecard treatment have to make a mental calculation as to whether it is beneficial to breach the policy in order to obtain a financial benefit. At the same time, they may also be considering how this choice might impact on their moral identity and social standing. All this additional mental processing is likely to slow participants down and may explain why the incentive scheme is not resulting in higher productivity.

**Table 7: Treatment Analysis (1)**

Payment	Fixed Remuneration (No Upside)		Balanced Scorecard		
	No peer information	With peer information	No peer information	With peer information	Gateway (no peer information)
Treatment	3	5	1	4	2
a. Number of Participants	56	53	105	47	57
b. Proportion of Participants with Perfect Compliance Throughout	75.0%	79.3%	62.9%	61.7%	50.9%
c. Proportion of Participants with Fewer than 3 Policy Breaches	87.5%	86.8%	76.2%	74.5%	52.6%
d. Proportion of Participants with 3 or More Policy Breaches	12.5%	13.2%	23.8%	25.5%	47.4%
e. Average number of completed, compliant transactions (with standard deviation)	18.6 (8.1)	20.5 (6.3)	22.4 (10.7)	24.4 (11.4)	23.5 (8.5)
f. Average number of Transactions Evaluated per Participant (with standard deviation)	38.2 (12.7)	40.5 (12.8)	40.8 (13.3)	43.3 (13.6)	42.1 (11.9)
g. Average Perception i.e. forecast proportion of participants who will always comply (with standard deviation)	64.3% (29.6%)	67.0% (29.2%)	64.6% (27.6%)	70.4% (27.1%)	56.7% (31.0%)



## Regression Analysis

Regression analysis allowed us to control for other factors that might influence behaviour such as demographics and personal attitudes. [Table 8](#) below details which of the variables we tested had a statistically

significant impact on compliance behaviour. Of these variables, Perceived Compliance Norms has the most significant influence, highlighting the importance of perceptions (i.e. culture) for determining compliance behaviour.

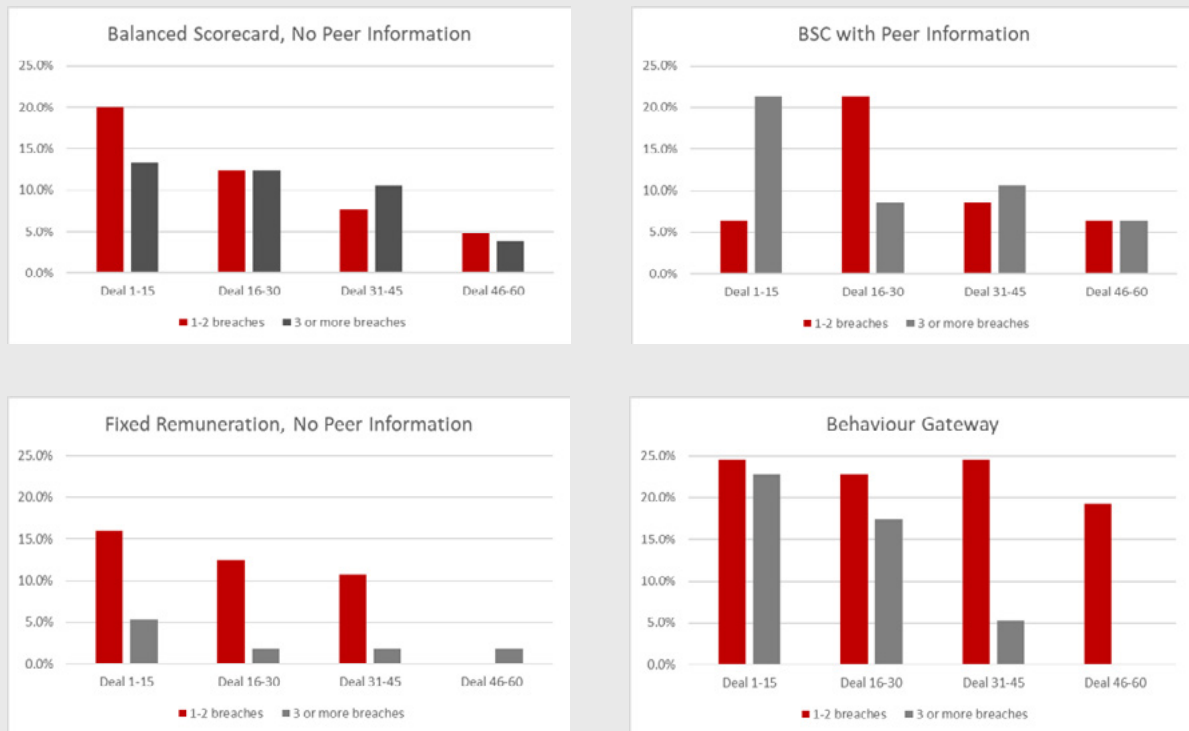
**Table 8: Other influences on Behaviour**

Variable	Detail	Impact on Compliance Behaviour
<b>Perceived Compliance Norms</b>	'In your opinion, what percentage of participants in the experiment will ALWAYS follow risk policy (i.e. not invest if outside the risk policy)? (Enter X%)'	Those who perceive higher compliance norms were more likely to themselves comply and any non-compliance was to a lesser extent.
<b>Individual Age</b>	How old are you? Less than 25yrs, 25-34, 35-44, 45-54, 55yrs or over	Older people were slightly less likely to comply with policy.
<b>Individual Gender</b>	Are you male/female?	No impact
<b>Seniority</b>	At what level are you currently working? Senior Management; Report to Senior Management; Middle Management; Team Leader; Professional Employee (but not a Manager); Team Member/ Front-Line Employee	No impact
<b>Lines of Defence</b>	What best describes your role (also known as the Lines Of Defence model)? 1) Business (Line 1); 2) Independent/Specialist Risk Manager, including Compliance (Line 2); 3) Internal Audit/Assurance (Line 3); 4) Don't know	No impact
<b>Segment</b>	In what segment of financial services do you work? Banking and Finance; Insurance; Other (includes Superannuation; Broking; Financial Planning/ Wealth Management; Funds Management; Consulting; Professional Services).	Those working in banking were less likely than other participants to be fully compliant (by 16%).
<b>Individual Risk Tolerance (Financial)</b>	A set of 2 items that assess an individual's propensity to take financial risk in their life outside of work. E.g. 'To achieve something in life, one has to take risks (Strongly Agree to Strongly Disagree)'	No impact
<b>Individual Preferences for Relative Performance</b>	A single item: 'How important is it for you to be the best at what you do? (1: not important; 7: very important)'	Participants with high preferences for relative performance were more likely to breach policy on 3 or more occasions.
<b>Moral Disengagement</b>	A set of 24 items that measuring the tendency employ cognitive processes that allow one to behave unethically without feeling guilty e.g. 'People shouldn't be held accountable for doing questionable things when they were just doing what an authority figure told them to do. (Strongly Agree to Strongly Disagree)'	Participants with a tendency to morally disengage are less likely to be fully compliant.
<b>Daily Incentives</b>	Eligibility for variable remuneration in usual workplace. 'In your everyday workplace, are you eligible for variable remuneration? (Yes/No)'	No impact
<b>Member of Professional Association</b>	'Are you a member of a professional association? (No, ANZIIIF, FINSIA, Other)'	No impact on the decision to comply, but among those participants who sometimes breach policy, those who are members of FINSIA have significantly higher compliance rates (by 24%).

Figure 1 shows how behaviour evolved during the experiment. Bear in mind that the average number of transactions evaluated (per participant) was around 40; many participants did not complete all 60. After

15, 30 and 45 transactions, participants were given feedback about their own performance. They were told the number of transactions they had completed and the number of policy breaches observed by the ‘audit office’.

**Figure 1: Compliance Rates Through the Experiment**



# Experiment (2)

---

The second experiment was conducted using a sample of students. The main difference here is the fact that we adjusted the audit rate during the experiment to see how behaviour changes. The experiment consisted of three sections, each of which lasted 8 minutes. In each section students were asked to evaluate up to 30 transactions.

	Balanced Scorecard	Compliance Gateway
Audit rate increasing	20%, 50%, 80%	20%, 50%, 80%
Audit rate decreasing	80%, 50%, 20%	80%, 50%, 20%

[Table 9](#) provides sample demographics for the student cohort while [Table 10](#) presents a summary of findings. We expected that a higher audit rate would result in higher rates of compliance. This was true when the audit rate was decreasing but not when the audit rate was increasing. In other words, it seems much harder to increase compliance rates than to reduce. This poses a challenge for those organisations trying to improve compliance; it suggests that increasing the audit rate might not necessarily be effective.

#### Some other points that are worth noting:

- Perceptions of compliance rates were again one of the most important determinants of compliance rates. This highlights the importance of behavioural norms or culture.
- Gender was significant here with females being significantly more compliant (unlike the professional sample).

**Table 9: Sample Demographics for Students**

	Percentage of Responses (Number of valid answers)
<b>Gender</b>	<b>(133)</b>
Male	55.64%
Female	43.61%
Other	0.75%
<b>Age</b>	<b>(133)</b>
Average, (SD), Median	24.53, (5.29), 23
<b>Student's Origin</b>	<b>(133)</b>
International	65.41%
Domestic	34.59%
<b>Academic Progress</b>	<b>(133)</b>
1st year of undergraduate	12.78%
2nd year of undergraduate	27.82%
3rd year of undergraduate	10.53%
4th year of undergraduate	5.26%
Postgraduate	43.61%
<b>Completed University-Courses in Finance</b>	<b>(133)</b>
0	27.82%
1	27.07%
2	15.04%
3	3.76%
>= 4	26.32%
<b>Attitude to Working in the Financial Services Industry in the Future</b>	<b>(133)</b>
Working in financial services is my goal	22.56%
I would consider working in financial services	48.87%
It's unlikely I'll ever work in financial services	28.57%
<b>Working-Experience in Financial Services Industry</b>	<b>(133)</b>
No	81.95%
Yes	(24) 18.05%
<b>Which of the following remuneration systems would you prefer to work under?</b>	<b>(133)</b>
Base salary plus variable remuneration based on a balanced scorecard	60.15%
Base salary plus variable remuneration based on financials only	18.05%
Fixed salary	20.30%
Other	1.50%



Table 10: Treatment Analysis (2)

Payment	Balanced Scorecard Increasing Audit Rates			Balanced Scorecard Decreasing Audit Rates		
Treatment	3a			3b		
Audit Rate	20%	50%	80%	80%	50%	20%
a. Number of Participants	33	33	33	34	34	34
b. Proportion of Participants with Perfect Compliance Throughout	57.6%	69.7%	69.7%	73.5%	73.5%	44.1%
c. Proportion of Participants with Fewer than 3 Policy Breaches	75.8%	78.8%	78.8%	94.1%	88.2%	58.8%
d. Those with 3 or more breaches	24.24%	21.21%	21.21%	5.9%	11.8%	41.2%
e. Average number of completed transactions (with standard deviation)	6.9 (3.8)	11.1 (5.6)	11.7 (7.1)	6.5 (5.0)	10.3 (5.5)	11.9 (6.0)
f. Average number of transactions evaluated (with standard deviation)	13.1 (4.6)	19.7 (6.7)	20.4 (6.8)	13.1 (5.8)	18.8 (7.4)	19.1 (6.5)

# References

---

Alexandra Albertsen, O., & Lueg, R. (2014). The balanced scorecard's missing link to compensation: A literature review and an agenda for future research. *Journal of Accounting & Organizational Change*, 10(4), 431-465.

APRA (2018) Information Paper: Remuneration Practices in Financial Institutions. [www.apra.gov.au](http://www.apra.gov.au)

Bol, J.C., 2011. The determinants and performance effects of managers' performance evaluation biases. *The Accounting Review*, 86(5), pp.1549-1575.

Delfgaauw, J. and Souverijn, M., 2016. Biased supervision. *Journal of Economic Behavior & Organization*, 130, pp.107-125.

Kirchler, M., Lindner, F., & Weitzel, U. (2017). Rankings and Risk Taking in the Finance Industry. *The Journal of Finance*.

Moore, C., Detert, J. R., Klebe Treviño, L., Baker, V. L., & Mayer, D. M. (2012). Why employees do bad things: Moral disengagement and unethical organizational behavior. *Personnel Psychology*, 65(1), 1-48.

Royal Commission into Misconduct in the Banking, Superannuation and Financial Services Industry (2018) Interim Report [financialservices.royalcommission.gov.au/Pages/interim-report.aspx](http://financialservices.royalcommission.gov.au/Pages/interim-report.aspx)

---

**APPLIED FINANCE CENTRE**

*Faculty of Business  
and Economics*



**MACQUARIE**  
University

**APPLIED FINANCE CENTRE**

**SYDNEY CBD**

Level 24, 123 Pitt Street  
Sydney NSW 2000  
Australia  
T: +61 2 9234 1700

**MELBOURNE CBD**

Level 6, Room 6.1, 500 Collins Street  
Melbourne VIC 3000  
Australia  
T: +61 3 9614 0005

**BEIJING**

Tsinghua University  
School of Continuing Education  
7th Floor, Tower B  
Innovation Plaza  
Beijing China 100084  
T: +86 10 6279 7825

[mafc.mq.edu.au](http://mafc.mq.edu.au)

CRICOS Provider 00002J