UNITED NATIONS
UNIVERSITY
UNU-WIDER

This paper is a draft submission to the

WIDER Development Conference

Human capital and growth

6-7 June 2016 Helsinki, Finland

An experimental investigation of the effects of affirmative action in Australian high schools

This a preliminary version - Please do not cite

Juliana Silva-Goncalves*[1,2], Uwe Dulleck[2], Anita Lee Hong[2], and Markus Schaffner[2]

[1]WZB Berlin
[2]Queensland University of Technology

January 23, 2016

Abstract

In this study we experimentally investigate the effects of affirmative action on effort, in socioeconomically disadvantaged high schools with a large representation of indigenous Australians. We conduct experiments in the classroom, with students performing a real-effort task in a competitive setting with monetary incentives. We introduce affirmative action by offering a starting advantage to those in the bottom third of the performance distribution. Our findings show that affirmative action increases effort in the task of those that the rule aims to favour, without necessarily discouraging effort of those who are indirectly penalised by affirmative action. Our results are very significant from a public policy perspective given that our participant pool includes a large representation of individuals who are tipically targeted by affirmative action policies.

Keywords: Affirmative action, Education, Classroom experiment, Real-effort task, Public policy

JEL Classification: C91, I24, J15

*Corresponding author: juliana.silvagncalves@gmail.com

This research was funded by the United Nations University World Institute for Development Economics Research and the Queensland University of Technology.

1 Introduction

In most societies there is persistent inequality in educational achievement and labour market outcomes between different social groups. Gender disparity in career choices and employment conditions, as well as unequal education and labour market outcomes between individuals from different socioeconomic backgrounds are recurrently observed. A very relevant case in the context of our study is the worldwide persistent disadvantage of native populations.[1] In Australia in particular, indigenous people systematically have less favorable educational and labour market outcomes compared to non-indigenous people (Bath and Biddle, 2011; Altman, 2000).

Affirmative action is often used as a policy tool to reduce such patterns of inequality. It gives preferential treatment to specific groups of people to compensate for their disadvantaged trajectory, influenced by family background, school environment or stereotypes. Affirmative action has been widely used in developed and developing countries. In the United States many programs promote minorities in the labour market and higher education institutions. In India, quota systems are in place to favour representation of women and people from lower castes in government and higher education institutions. In Australia, many universities and employers try to achieve participation of indigenous people, at least equivalent to their representation in the Australian population. Despite its popularity, to date there is little evidence on whether affirmative action increases incentives to invest in education or whether by inducing lower standards, reinforces the gap in educational attainment and stereotypes (Coate and Loury, 1993).

In this study, we conduct a laboratory experiment in socioeconomically disadvantaged high schools with a large representation of indigenous Australians, and test whether affirmative action influences effort of low performing students. In our experiment, participants perform a simple real-effort task in a competitive setting with monetary incentives. Those whose performance is within the top third of the distribution receive a high piece-rate payment, whereas the remaining participants receive a low piece-rate payment. In this setting, we test whether a starting advantage given to the bottom third performers influences effort in a subsequent stage. Our results indicate that affirmative action increases effort of those that the rule aims to favour. Moreover, the discouraging effect on those who are not benefited (but indirectly penalised) by affirmative action is at best small.

An important contribution of our study is the specificity of our participant pool, students in socioeconomically disadvantaged schools, with a large representation of indigenous. Additionaly, the schools use positive discrimination by offering and partnering

[1] See, for example, Patrinos (1992) for Canada, Bradley et al. (2007) for Australia, Ladson-Billings (2006) for the US, Patrinos (2004) for Bolivia, Ecuador, Guatemala, Mexico and Peru.

with non-governmental organisations that offer special programs and learning support to indigenous students. The characteristics of our participant pool enhance the external validity and policy significance of our findings and differentiate our study from the existing experimental literature on affirmative action. Closely related to our study, Calsamiglia et al. (2013) test whether affirmative action affects performance in an effort-based task (sudoku) in a tournament setting. The participants were primary school children and half of them had received training at the task in a prior stage. In this setting, affirmative action compensating for unequal experience in the task increased participants' performance. Another related study is the one by Dulleck et al. (2015) which tests in cross-country lab experiment whether affirmative action influences effort in an math task in a tournament setting. The participant pool allows for a large disadvantage in performance for half of the participants and a real stereotype. They find that affirmative action does not affect effort when it compensates for a disadvantage, but decreases effort in an alternative task in which there is no disadvantage. Also related to our study is the work by Balafoutas and Sutter (2012) and Niederle et al. (2013). Both studies test whether affirmative action reduces the gender gap in tournament entry decisions. Using effort-based tasks, they find that affirmative action doubles tournament entry for women. Early laboratory experiments with a standard participant pool by Schotter and Weigelt (1992) use a tournament setting with an abstract framing, where effort consists in choosing a number which decreases the payoff according to a specified cost function, and the disadvantage is introduced by varying the parameters of the cost function between participants. In accordance with theoretical predictions (see also Lazear and Rosen, 1981), the authors show a positive effect of affirmative action on effort.

Laboratory experiments are a very valuable tool to study the incentive effects of affirmative action. A major reason is that such policies are often adopted endogenously, challenging the interpretation of its causal effects by means of observational data.[2] Moreover, ethical and political considerations prevent conducting field experiments on such policies, with a valid counterfactual. Nonetheless, standard participants in experiments, i.e. university students, often do not offer a good representation of the population targeted by affirmative action. As in this study, experiments which add the realism of field data, with a participant pool very similar to the one that would be part of a field experiment, are likely to provide more accurate predictions on how affirmative action influences individual behaviour (Falk and Fehr, 2003, see).

[2]Many studies have investigated the incentive effects of affirmative using existing data. For instance, some studies have reported the effects of affirmative action on admission to college of minority students using policy simulations (see Arcidiacono, 2005; Howell, 2010). Other studies explored policy shifts in the US on university admissions of minority students (see Card and Krueger, 2005; Long, 2004). In the labour market context, Leonard (1990) and Miller and Segal (2012) investigated how affirmative action influences employment of women and black people.

The remainder of the paper is organised as follows. Section 2 describes the experimental design and participant pool. Section 3 presents the results, Section 4 summarises and discusses the findings.

2 Description of the experiment

2.1 Participant pool

We conducted the experiments with a total of 263 high school students, enrolled in Year 8 and Year 9 (aged between 13 and 15), across four public schools in the Australian State of Queensland. Two of the participating schools are located in metropolitan areas in South East Queensland, and the other two are located in a metropolitan area in North East Queensland. The schools are characterised by similar and below national average socio-educational advantage indicators. The selection of the schools was based on two criteria. These schools have large numbers of indigenous students as well as special learning support programs for these students. A student is classified as indigenous if he or she self-identifies as Aboriginal and/or Torres Strait Islander.[3] Within each participating school, all students enrolled in Year 8 and Year 9 were invited to participate.

It is an important advantage to work with underprivileged high school students as opposed to a standard subject pool when conducting laboratory experiments on affirmative action. The former are more representative of the population that may be subject to positive discrimination than university students, who are expected to be the best achievers during compulsory education. Moreover, it gives direct insights on how affirmative action policies that benefit high school graduates historically disadvantaged in the labour market, university admissions and/or with low school achievement, influence their incentives to exert effort in school.

The study was approved by the Queensland Department of Education, Training and Employment, and by the Queensland University of Technology ethics committee. All participants received an information letter briefly describing what their participation would involve, and had to return the consent of their parents or caregivers for participation.

We present the socio-demographic characteristics of our sample in Table 1. It counts with a total of 55 students who identify as indigenous and 208 students who do not identify as indigenous. Fifty-four percent of the participants are enrolled in Year 8 and 46 percent are enrolled in Year 9. Our sample is balanced in terms of gender. There is some variation across treatments, which is due to the fact that we conducted few sessions

[3]In Queensland, 7.1 percent of primary and high school students identify as indigenous. In public schools, the proportion of self-identified indigenous students is 8.9 percent (Steering Committee Review of Government Service Provision, 2015). In the four participating schools, the percentage of indigenous students in the school population is 7, 11, 19 and 27 percent.

(16 in total) and one treatment per session. However, any potential sample selection issues are strongly attenuated as we contrast within-subject variation in behaviour across treatments. Additionally, we control for the socio-demographic variables in our regression analysis. We also have information with regards to participants' achievement in math and English subjects obtained from each school's principal teacher. We do not observe significant differences in the achievement distribution across treatments.[4] In addition, 9 percent of the students in our sample have been held back in school at least one year. We also observe that the large majority of our participants (92 percent) primarily speak English at home. With regards to participants' living situation, 54 percent live with both parents and 42 percent live with only one of their parents. With respect to parents' occupation, 26 percent of the participants indicated that their mother is unemployed or a housewife, whereas only 6 percent declared that their father is unemployed. The father's occupation of 51 percent of our participants is classified as low skill (e.g. construction worker, truck driver, miner) and this also is the case for the mother's occupation of 25 percent of the participants (typically, cashier at a supermarket, waitress or cleaner). Nineteen percent of the participants indicated that their mother is a nurse or caregiver, 8 percent indicated that their mother has an administrative job and 3 percent indicated that their mother is a teacher. With regards to the father's occupation of the remaining participants, 6 percent indicated that the father is a technician, 5 percent that he has an administrative job, works in the army or police and only 2 percent indicated that the father has a high skilled job (e.g. engineer, medical doctor).

[4]The p-value of the two-sided Wilcoxon rank-sum test is 0.39 for math and 0.25 for English scores.

Table 1: Descriptive statistics of the participant pool

	Baseline		Affirmative action		Total	
	N	%	N	%	N	%
Non-indigenous	112	87	96	72	208	79
Indigenous	17	13	38	28	55	21
Year 8	96	74	46	34	142	54
Year 9	33	26	88	66	121	46
Female	57	44	71	53	128	49
Male	72	56	63	47	135	51
Math grade *(1=worst; 5=best)*						
1	9	7	6	5	15	6
2	11	9	17	13	28	11
3	43	34	56	42	99	38
4	41	33	33	25	74	29
5	22	17	21	16	43	17
English grade						
1	1	1	0	0	1	0
2	9	7	14	11	23	9
3	62	48	55	43	117	46
4	42	33	58	45	100	39
5	14	11	1	1	15	6
Were held back in school	13	10	11	8	24	9
Speak English at home	121	94	120	90	241	92
Family living situation						
Both parents	76	59	65	49	143	54
Single-parent	47	36	63	47	110	42
Other	6	5	6	4	12	4
Mother occupation						
Unemployed/housewife	32	25	37	28	69	26
Low skill job	32	25	35	26	67	25
Administration	11	9	10	7	21	8
Army/police	1	1	1	1	2	1
Business owner	1	1	0	0	1	0
Nurse/care	25	19	24	18	49	19
Teacher	4	3	4	3	8	3
Other	23	18	23	17	46	17
Father occupation						
Unemployed	5	4	10	7	15	6
Low skill job	70	54	64	48	134	51
Admin/manager	8	6	5	4	13	5
Army/police	10	8	3	2	13	5
Business owner	1	1	2	1	3	1
Nurse/care	2	2	4	3	6	2
Technician	8	6	8	6	16	6
High skill job	1	1	4	3	5	2
Other	24	19	34	25	58	22
Total	129	100	134	100	263	100

2.2 Experimental setup

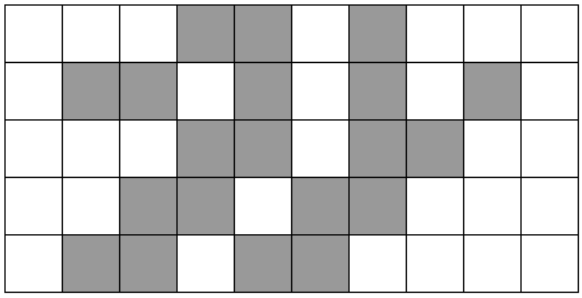
All experimental sessions were conducted by the same experimenter (author in this study) and a research assistant. The experiments were conducted during regular school hours and lasted around 60 minutes. A teacher was always present during the sessions, so that the participants would perceive the experiment as a formal activity. All participants received the exact same instructions. The instructions for all tasks were given aloud at the beginning of the session, following a written script (see Appendix A.2). Examples were used to illustrate each the tasks and participants were encouraged to ask questions to ensure a good understanding of the tasks and the associated payoff structure. Subsequently, participants performed the tasks individually on a website, programmed with the experimental software Coral (Schaffner, 2013) and specifically designed for our experiment.[5] Once participants started performing the tasks, all questions were answered privately. The experimental webpages displayed screens with a brief summary of the instructions before participants started to perform each stage of the task.

The experiment was incentivised with real monetary payments. At the start of the experiment, participants were aware that they would earn 7 AUD as a participation fee and that their final payoff would be determined by their performance. Participants could earn up to 14 AUD. The payoffs were distributed to each participant privately in cash, immediately at the end of the session.

Participants performed a real-effort task in two stages. The task consisted of counting the number of shaded squares in a grid with a total of fifty squares, as shown in Figure 1 below. Each time the participant entered an answer, a new grid appeared.

Figure 1: Experimental task

How many squares in the grid are shaded?



[5]The screenshots of the experimental website are provided in Appendix A.3.

The first stage of the task is identical in both treatments. Participants perform the task for eight minutes. Those ranked among the top third of the performance distribution receive 15 cents per correct answer, and all the remaining participants receive 5 cents per correct answer. Ties are broken randomly by the experimental software. Subsequently, participants are asked to guess their rank. After guessing their rank, each participant is informed about his or her total number of correct answers in the first stage and whether he or she was ranked among the top third.

In the second stage of the task, participants perform the exact same task. However, we now introduce two different treatments, the baseline and affirmative action treatment. In the baseline treatment, the payoffs are calculated exactly as in the first stage. In the affirmative action treatment, participants who are ranked in the bottom third of the performance distribution in the first stage of the task receive 15 extra points, that are added to their number of correct answers in stage 2 to determine their score.[6] For all other participants, that is, those ranked above the bottom third, the score is determined by their number of correct answers in stage 2 only. Participants whose score is within the top third receive 15 cents per correct answer. All the other participants receive 5 cents per correct answer. The starting advantage increases the chances for the lowest performers to be ranked among the top performers in the second stage, and therefore, receiving the high piece-rate payment. However, only their number of correct answers enters in the payoff calculation. As in the first stage, participants are asked to guess their rank after performing the second stage of the task.

In both the baseline and affirmative action treatment, before performing the second stage, participants are informed about whether one-third of the participants in the room is receiving the 15 extra points in the second stage. If that is the case, they are also told whether they are one of the participants receiving the extra points.

Only one treatment was conducted in each session and each participant took part in one session only. With our experimental design, we test whether affirmative action influences effort of the lowest performers by comparing the within subject variation in performance between the two stages, between the baseline and affirmative action condition. We also test whether affirmative action impacts the effort of participants who are not targeted by affirmative action, but are indirectly penalised when the lowest performers are given a starting advantage.

[6]The size of the starting advantage was defined following the observation of students' performance in the task in a pilot session, with students in the same age group as our participants but who did not participate in the experiment. By giving 15 extra points to all participants in the bottom third, half of them were lifted to the top-third.

Table 2: Description of the treatments

	Baseline	**Affirmative action**
Stage 1	Participants whose number of correct answers is within the top 3rd of the distribution get the high piece-rate payment. All the others get the low piece-rate payment.	
Stage 2	As in stage 1.	Participants whose number of correct answers is below the bottom 3rd in stage 1 receive 15 extra points in stage 2. Participants whose score (number of correct answers + extra points) is within the top 3rd, receive the high piece-rate payment for each correct answer. All the others receive the low piece-rate payment.

After performing the two stages of the task, participants are asked to answer a short questionnaire, asking basic demographic information (gender, date of birth, year level, indigenous status). They are also asked whether English is the primary language spoken at home, who they live with and their parents' or caregivers' occupation. Participants are also asked questions in relation to their educational and professional aspirations. Specifically, they are asked whether they intend to complete high school, what is the highest level of education they intend to reach and which occupation they would like to have once they leave school.

Additionally, participants are asked questions about the enjoyable and difficulty aspects of the task and the clarity of the instructions.[7] Fifty-three percent of the participants considered the task entertaining, 42 percent found it a bit long and 5 percent found it boring. Forty-four percent found the task easy, 55 percent not too hard and only 2 percent considered it very hard. With regards to the clarity of the instructions, 80 percent found them easy to follow, 18 percent a bit confusing and only 2 percent found them very confusing. Finally, we asked participants whether they received extra points in the second stage of the task. The answers allow us to control in our analysis for participants' understanding of a fundamental element of our experimental design.

At the end of the experiment, we asked the teacher present in the room to draw a ball from a bag, containing two balls of different colours, to determine which of the stages of the task was paid. Participants were informed at the start of the session that although

[7] At the end of the questionnaire, participants are asked the following questions: "How did you like the task?" (answer options: "Entertaining", "A bit long", "Boring"), "How did you find the task?" (answer options: "Easy", "Not too hard", "Hard"), "How did you find the instructions?" (answer options: "Easy too follow", "A bit confusing", "Very confusing").

they had the possibility to earn money in both stages of the task, only one randomly determined at the end of the session would be considered for the final payoff.

3 Results

3.1 Descriptive analysis

We start with a descriptive analysis of the data in Table 3 and Figure 2, Figure 3 and Figure 4. We describe the data for the whole sample, and for two subsamples. We consider the subsample of participants ranked in the bottom third of the performance distribution in stage 1 and the subsample of participants ranked above the bottom third threshold in stage 1.

In Table 3, we report the performance mean (average number of correct answers) and standard deviation in stage 1, stage 2 as well as the average difference in performance between the two stages (correct answers in stage 2 - correct answers in stage 1), separately by treatment. On average, participants answered 25 questions correctly in the first stage, in both the baseline and affirmative action treatment (p-value=0.98). Likewise, within each subsample, none of the differences in average performance in stage 1 across treatments is statistically significant at conventional levels. When restricting the sample to the bottom third of the performance distribution in stage 1, the average number of correct answers is 18.33 and 17.09 in the baseline and affirmative action treatment respectively (p-value=0.23). For the subsample of participants ranked above the bottom third threshold, the average number of correct answers in stage 1 is 28.7 and 29.3 in the baseline and affirmative action treatment respectively (p-value=0.41). We also observe that standard deviations are relatively large, pointing out heterogeneous performance between participants.

The progression in performance between stage 1 and stage 2 is positive and statistically significant in both treatments. This is likely to be explained by a learning effect, as participants may become more efficient at the task in the second stage. For the whole sample, the performance in stage 2 increased by 2.42 and 2.93 correct answers in the baseline and affirmative action treatment, respectively (p-values<0.00). The increase in performance is the largest for the subsample of participants in the bottom third of the distribution in the affirmative action treatment (those who received the 15 extra points in stage 2). While participants in the bottom third of the distribution in the baseline treatment obtained, on average, 1.67 more correct answers in stage 2, those in the affirmative action treatment obtained 4.17 more correct answers (p-values are 0.12 and 0.00, respectively). The difference between the two treatments is weakly statistically significant at the 10 percent level (p-value=0.10). The difference in performance

progression between the baseline and affirmative action treatment for the participants who were ranked above the bottom third in stage 1 (therefore, not eligible for affirmative action), is negative but not statistically significant at conventional levels.

Table 3: Description of participants' performance

	Stage 1		Stage 2		Difference	
	mean	sd	mean	sd	mean	sd
All						
Baseline	25.09	6.84	27.50	8.89	2.42***	5.71
Affirmative action	25.10	7.57	28.03	8.39	2.93***	5.94
Difference	0.01		0.53		0.51	
Bottom 3rd in Stage 1						
Baseline	18.33	4.75	20.00	7.43	1.67	7.12
Affirmative action	17.09	5.12	21.26	7.80	4.17***	7.09
Difference	-1.24		1.26		2.5*	
Above the bottom 3rd in Stage 1						
Baseline	28.70	4.71	31.52	6.75	2.82***	4.80
Affirmative action	29.30	4.72	31.57	6.26	2.27***	5.17
Difference	0.60		0.05		-0.55	

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

We also show the performance distribution by treatment, for the whole sample in Figure 2, for the subsample of participants in the bottom third of the performance distribution in stage 1 in Figure 3, and for the subsample of participants above the bottom third threshold in Figure 4. A considerable heterogeneity in individual performance is noticeable in all figures. We contrast the performance distribution in stage 1, stage 2 and the distribution of the variation in performance between stages, across the baseline and affirmative action treatment, using two-sided Wilcoxon rank-sum tests. The results with regards to performance in stage 1 (subfigures A and D) do not yield any significant differences between treatments (p-value is 0.791 for the whole sample, 0.251 for the subsample of participants in the bottom third and 0.347 for the subsample of participants above the bottom third). Similarly, the performance distribution in stage 2 (subfigures B and E) is not significantly different between the two treatments for each of the samples considered (p-values are 0.608, 0.530 and 0.958). Turning now to the progression in performance between the two stages (subfigures C and F), we also do not observe any statistically significant differences at conventional level between treatments (p-values are 0.454, 0.119 and 0.700). Nonetheless, the systematic non-significant differences do not rule out differences in behaviour across the baseline and affirmative action treatment.

The absence of statistically significant results in our descriptive analysis may be explained by a substantial heterogeneity in performance across participants combined with our small sample size and potential heterogeneous effects of affirmative action for participants with different characteristics.

In the following section, we describe the results obtained with the regression analysis, which allows us to control for other factors that may affect the magnitude and statistical significance of the treatment effect. These are individual level factors, specifically, gender, indigenous status, Year level, math scores and participants' correct understanding of the instructions, as well as session effects.

Figure 2: Performance distribution: Whole sample

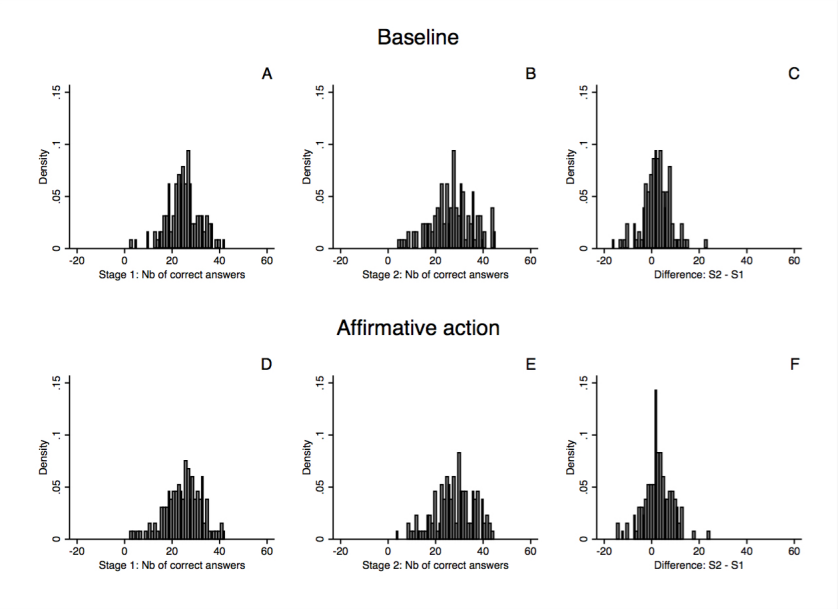


Figure 3: Performance distribution: Bottom 3rd in stage 1

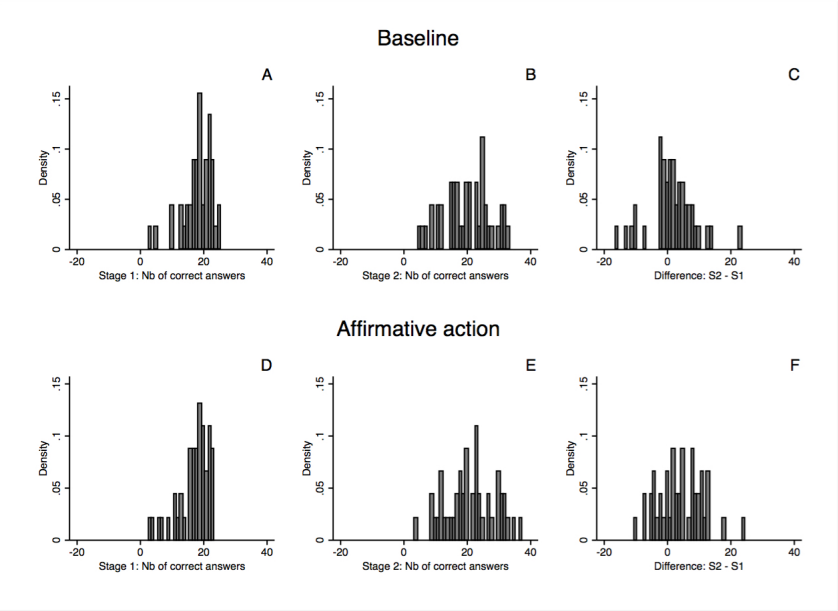
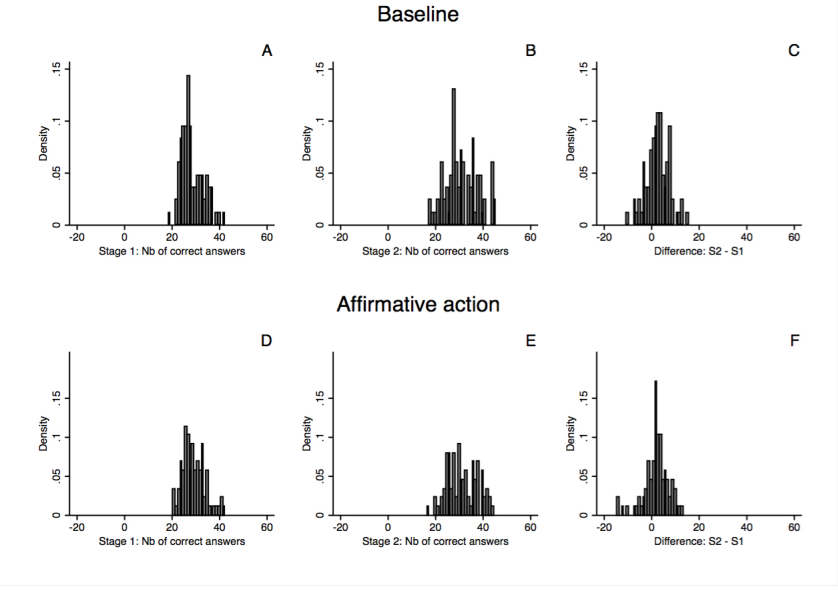


Figure 4: Performance distribution: Above the bottom 3rd in stage 1



3.2 Regression results

We present the OLS regression results for the effect of affirmative action on participants' performance and beliefs for their relative performance. We also examine the consequences of affirmative action in terms of efficiency. We discuss the results focussing on the subsample of participants who are ranked in the bottom third of the performance distribution in stage 1. We also analyse the effect on the subsample of participants who are not targeted but indirectly impacted by affirmative action. We systematically report the results obtained with two specifications, a basic one and an extended specification with additional covariates. In both specifications, the reported standard errors are robust and clustered by session, to account for the fact that the error term is unlikely to be independent across observations within session. The basic model includes an indicator variable for whether the participant is in the baseline or affirmative action treatment (**AA**), an indicator variable for whether the participant was ranked in the bottom third of the performance distribution in stage 1 (**Bottom 3rd**), and the interaction between both variables. The extended model controls for gender, Year level, indigenous status, math scores, and whether participants gave the correct answer when asked if they received 15 extra points in stage 2 (Checkbonus). As mentioned in section 2.2, at the end of the experiment all participants were asked to indicate whether they received 15 extra points in the second stage. Eighty-eight percent of the participants gave the correct answer. Among the 12 percent of participants who gave an incorrect answer, 3 percent were in the baseline treatment and 9 percent in the affirmative action treatment. We include a control variable, which takes the value 1 if the participant gave a correct answer and 0 otherwise. Arguably participants' incorrect understanding of whether they received the extra points will attenuate the treatment effect. For comparative purposes, we report in the appendix the results obtained when excluding the 35 participants who gave an incorrect answer. We also briefly comment on the results obtained by their exclusion throughout the description of the results in the following sections. Overall, excluding these participants increases the magnitude and statistical significance of the treatment effect. Finally, there are small variations in sample sizes between the basic and extended specification, which are due to missing information on the math scores of 4 participants.[8]

3.2.1 The effect of affirmative action on effort

In Table 4, we report the effect of affirmative action on participants' effort measured by performance in stage 1, stage 2 and by the progression in performance between the two stages. Our main estimates of interest are those for the variables **AA x Bottom**

[8]The principal teacher did not provide us information on these students' math scores as they had been enrolled in other schools in the school term prior to our experiment.

3rd and **AA**. The former estimate corresponds to the effect of affirmative action on the performance gap between participants ranked below the bottom third (who are eligible for affirmative action) and those ranked above the bottom third.[9] The estimate for **AA** corresponds to the effect of affirmative action on performance of participants who were ranked above the bottom third of the performance distribution in stage 1, and are therefore indirectly penalised by affirmative action.[10]

We first look at the regression results with performance in stage 1 as the dependent variable (columns 1a and 1b in Table 4). As expected, none of our main variables of interest are statistically significant. Given that the experimental instructions and the first stage of the task are identical for both treatments, the absence of statistically significant differences in the first stage further supports the validity of our experimental design. By construction, participants ranked in the bottom third have a lower performance than those ranked above the bottom third. The estimate for the lower performance of participants ranked in the bottom third relative to those ranked above, is slightly larger than 10 in both specifications (significant at 1 percent level). We also observe that participants with better math scores perform slightly better in the task. An increase in one unit on the math scores scale (1 to 5), on average, increases performance by 0.68. Moreover, performance does not seem to be influenced by the gender or indigenous status of the participants.[11] However, participants in Year 9 perform better, by 1.61 units (significant at 10 percent level), than those in Year 8.

The regression results for performance in stage 2 (columns 2a and 2b in Table 4) suggest that affirmative action does not affect performance in the task, as the coefficient estimate for **AA x Bottom 3rd**, although positive, is not statistically significant at conventional levels. However, looking at the variation in performance between the two stages provides a better estimate of the treatment effect, by allowing to neutralise individual effects, which may be important given our relatively small sample. A similar approach is to control for performance in stage 1 in the regression with performance in stage 2 as the dependent variable.[12] As reported in Table 4, we find that affirmative action leads to an increase in performance of participants in the bottom third of the distribution relative to those who are not eligible for affirmative action, by 3.1 units, sig-

[9] The estimate for the interaction term **AA x Bottom 3rd** corresponds to $[Y_{AA}^{B3rd} - Y_{Base}^{B3rd}] - [Y_{AA}^{>B3rd} - Y_{Base}^{>B3rd}]$, where Y denotes performance, $B3rd$ the group in the bottom third of the performance distribution in stage 1 and $> B3rd$ the group above the bottom third threshold.

[10] The estimate for the variable **AA** corresponds to $[Y_{AA}^{>B3rd} - Y_{Base}^{>B3rd}]$.

[11] The absence of gender differences in performance is line with the findings in the existing literature (see Niederle and Vesterlund, 2007; Niederle et al., 2013; Balafoutas and Sutter, 2012). These studies systematically observe differences in competitive preferences between men and women, but not in average performance in simple real-effort tasks analogous to the one used in our study.

[12] The results obtained are qualitatively and quantitatively very similar to the ones reported in this section, and reported in Table A.1 in Appendix A.1.

nificant at the 5 percent level (columns 3a and 3b). The estimate confounds the effect of affirmative action on those in the bottom third of the performance distribution who are benefitted by affirmative action and any potential effect on the remaining participants who are indirectly penalised by affirmative action. The latter is given by the estimate for the variable Treatment **A**A, which is not statistically significant at conventional levels. The participants who are benefited by affirmative action have a progression in performance between the two stages that exceeds the one for their counterparts in the baseline treatment by 2.5 units in the basic specification (column 3a) and 2.3 units when adding additional control variables (column 3b).[13] The magnitude of the treatment effect is large, representing an increase in performance by approximately 50 percent relative to the baseline.

Finally, we observe that participants who reported correctly whether they received the 15 extra points in the final questionnaire have a larger improvement in performance by 1 unit between the two stages, significant at 5 percent level. By excluding the 35 participants who gave an incorrect answer, the magnitude and statistical significance of the estimate for **AA x Bottom 3rd** increases slightly. The estimates are 3.57 in the basic model and 3.59 in the extended one, and affirmative action increases the performance gap between the two stages for participants ranked in the bottom third by 3.15 and 2.91 correct answers, respectively (see Table A.2 in Appendix A.1).[14]

[13]These values are obtained by adding the coefficient estimates for **AA x Bottom 3rd** and **Treatment** AA.

[14]Our results are qualitatively and quantitatively very similar to the findings in Calsamiglia et al. (2013). In their experiment, controlling for participants' ability in the task (with pre-test scores), affirmative action increases performance of participants non-experienced at the task by about 4 units and their average performance without affirmative action is 23 units.

Table 4: OLS regressions on performance

	Stage 1		Stage 2		Difference	
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)
Treatment **AA**	0.593	-0.033	0.044	-0.837	-0.549	-0.804
	(0.901)	(0.706)	(1.130)	(1.044)	(0.530)	(0.534)
Bottom 3rd	-10.369***	-10.180***	-11.524***	-10.925***	-1.155	-0.745
	(0.898)	(0.918)	(1.328)	(1.291)	(0.913)	(0.950)
AA x Bottom 3rd	-1.839	-1.650	1.216	1.433	3.056**	3.083**
	(1.279)	(1.336)	(1.643)	(1.568)	(1.389)	(1.289)
Male		0.497		-0.450		-0.947*
		(0.729)		(0.864)		(0.540)
Year 9		1.611*		2.177***		0.566
		(0.763)		(0.732)		(0.771)
Indigenous		-0.452		-0.226		0.225
		(0.756)		(0.963)		(0.625)
Math		0.679*		0.566		-0.113
		(0.347)		(0.478)		(0.288)
Checkbonus		-0.305		0.719		1.024**
		(0.838)		(0.740)		(0.478)
Constant	28.702***	13.091*	31.524***	11.117	2.821***	-1.975
	(0.797)	(6.988)	(0.941)	(6.618)	(0.378)	(6.295)
N	263	259	263	259	263	259
Adj. R^2	0.557	0.568	0.358	0.356	0.007	-0.001

Robust standard errors in parentheses, clustered by session. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Our results suggest that affirmative action does not influence performance of participants ranked above the bottom third in stage 1, although they are indirectly penalised by affirmative action. Nevertheless, the coefficient estimates obtained in Table 4 for the variable Treatment **AA** are informative about the average effect on participants whose performance exceeds the affirmative action eligibility threshold. We now test for potential heterogeneous effects of affirmative action on that subsample, according to participants' performance in the task. Our total number of participants who are not eligible for affirmative action in stage 2 allows us to restrict our analysis to this subsample. We report in Table 5 our regression results, which include an interaction term between the treatment variable and performance in stage 1.[15] The coefficient estimate for Treatment **AA** gives us the effect of affirmative action for participants with a very low performance in stage 1 (just above the bottom third threshold). The positive estimates suggest that affirmative action enhances performance of participants with a very low performance in

[15]Looking at performance in stage 2 and at the difference in performance across the two stages yields exactly the same results, given that we interact the treatment variable with the performance in stage 1 and control for performance in stage 1.

stage 1. However, the coefficient is only weakly statistically significant (at the 10 percent level) in the basic specification (columns 1a) and not statistically significant at conventional levels in the extended model (column 1b). In addition, we observe that affirmative action has a discouraging effect for participants with a higher performance in stage 1. The estimate for **AA x Stage 1** is negative (approximately -0.2) and weakly significant in both specifications.[16]

We also present in Table A.3 in Appendix A.1 the regression results obtained when excluding the participants who incorrectly reported whether the received the 15 extra points (accounting for 20 observations). As previously, the results do not qualitatively change, but their magnitude and statistical significance slightly increases (the estimate for **AA x Stage 1** is approximately -0.3 in both specifications).

Table 5: OLS regressions on performance

	Difference (S2 - S1)	
	(1a)	(1b)
Treatment **AA**	6.150*	5.809
	(3.338)	(3.616)
Stage 1	0.009	0.019
	(0.049)	(0.059)
AA x Stage 1	-0.229*	-0.211*
	(0.116)	(0.120)
Male		-1.637**
		(0.750)
Year 9		-0.541
		(0.749)
Indigenous		-0.032
		(0.722)
Math		-0.245
		(0.304)
Checkbonus		0.965
		(0.724)
Constant	2.564	7.539
	(1.687)	(6.888)
N	172	172
Adj. R^2	0.008	0.009

Robust standard errors in parentheses, clustered by session. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

[16]These results are in line with Calsamiglia et al. (2013), who find that affirmative action does not affect the average performance of participants who do not benefit from the policy, but that it positively affects those with a low performance level (at a decreasing rate) and has a negative impact on participants with the highest performance in the task.

3.2.2 The effect of affirmative action on beliefs for relative performance

We now investigate whether affirmative action influences beliefs for relative performance. We present the results in Table 6, where the dependent variables correspond to participants' percentile rank guesses within session.[17] The estimates for the variables **AA x Bottom 3rd** and Treatment **AA** on the variation in expectations between the two stages are not statistically different from zero (columns 3a and 3b), suggesting that affirmative action does not influence participants' expectations for their rank. This result is also in line with the findings in Calsamiglia et al. (2013).

We also observe that participants with a higher actual rank in stage 1 expect to be ranked higher in stage 1 by 14 percentile points, significant at the 10 percent level (columns 1a and 1b). Similarly, participants with a higher actual rank in stage 2 expect to be ranked higher in stage 2 by 21 percentile points significant at the 1 percent level (columns 2a and 2b). Moreover, indigenous participants systematically expect to be ranked lower than their non-indigenous counterparts, by approximately 8 percentile points, significant at the 1 percent level. This is an interesting observation given the absence of differences in performance in the task between indigenous and non-indigenous participants. Another interesting result is that we do not have strong evidence that male participants are more confident with respect to their relative performance than females. The estimate for the variable male is systematically negative in our regressions, but not statistically significant at conventional levels.[18] When excluding the subsample of participants who gave an incorrect answer to whether they received the extra points, the results do not change qualitatively or quantitatively (see Table A.4 in Appendix A.1).

As for performance in the previous section, we test whether the impact of affirmative action on beliefs for relative performance is dependent on the actual rank, for participants with a performance level above the bottom third threshold. There is no evidence that the non-statistically significant average effect for this subsample of participants (given by the estimate for Treatment **AA**) reported in Table 6 countervails heterogeneous effects (see Table A.5 in Appendix A.1).

[17] For simplicity, in the experiment participants were asked to guess their rank between 1 and the total number of participants in the session. We converted their answers in percentile rank guesses by dividing their expected rank by the total number of participants in the session.

[18] The results is also observed in Calsamiglia et al. (2013).

Table 6: OLS regressions on expected percentile ranks

	Stage 1		Stage 2		Difference	
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)
Treatment **AA**	-0.031	-0.028	-0.010	-0.015	0.023	0.016
	(0.030)	(0.036)	(0.031)	(0.032)	(0.018)	(0.020)
Bottom 3rd	-0.010	-0.017	0.023	0.012	0.039	0.036
	(0.044)	(0.048)	(0.038)	(0.041)	(0.033)	(0.040)
AA x Bottom 3rd	-0.019	0.000	0.026	0.045	0.035	0.034
	(0.056)	(0.060)	(0.041)	(0.045)	(0.045)	(0.048)
Male		-0.020		-0.034		-0.011
		(0.022)		(0.020)		(0.024)
Year 9		-0.019		-0.007		0.012
		(0.032)		(0.022)		(0.025)
Math		0.020		0.009		-0.013
		(0.016)		(0.011)		(0.012)
Indigenous		0.084***		0.087***		-0.006
		(0.025)		(0.020)		(0.020)
Checkbonus		0.057		0.048		-0.013
		(0.043)		(0.038)		(0.023)
Rank in stage 1	0.145*	0.143*				
	(0.070)	(0.070)				
Rank in stage 2			0.205***	0.213***		
			(0.068)	(0.067)		
Constant	0.279***	0.309	0.279***	0.262	0.032**	-0.001
	(0.036)	(0.304)	(0.038)	(0.217)	(0.014)	(0.238)
N	263	259	263	259	263	259
Adj. R^2	0.027	0.049	0.088	0.096	0.031	0.016

Robust standard errors in parentheses, clustered by session. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.2.3 The effect of affirmative action on efficiency

In this section, we describe how affirmative action affects the pool of participants in the top third of the performance distribution and who receive the high piece-rate payoff. A recurrent argument against affirmative action is the constraint it imposes on matching candidates who are expected to be the most productive at the task, in favour of individuals expected to be less productive who would not be selected in the absence of affirmative action. Our task does not require a participant to acquire specific skills to increase his or her performance beyong higher effort. Hence, it allows to investigate whether affirmative action is likely to raise individuals' effort enough, so that it does not induce a lower average performance of those who receive the high piece-rate payment in the second stage

and even increases overall performance.

In what follows, we look at how affirmative action influences the chances of a participant with a low performance to be ranked among the top third in the second stage and how it affects the overall performance of the participants receiving the high piece-rate payoff. In Table 7, we report the number of participants who were ranked in the bottom third in stage 1 and in the top third in stage 2, as well as the number of participants who were ranked among the top third in stage 1 and below the top third in stage 2, separately by treatment. Overall, affirmative action leads to substantial changes in the composition of the group of participants receiving the high and the low piece-rate payments. In the baseline treatment, only 2 of the participants who were ranked in the bottom third in the first stage, were ranked in the top third in the second stage of the task. In contrast, in the affirmative action treatment, 22 participants who were ranked in the bottom third in stage 1 were ranked among the top third in stage 2. Among those who were ranked in the top third in stage 1, 11 were ranked below the top third in the second stage in the baseline treatment, whereas 24 were ranked below the top third in the affirmative action treatment.

The advantage given to participants in the bottom third of the performance distribution in the affirmative action treatment was very large, which strongly accounts for the large promotion of the low performers in stage 2. It is also interesting to look at how many participants in the bottom third in stage 1 would have been promoted to the top third based on their *real* rank in stage 2, that is, based on their number of correct answers only. When considering participants' real relative performance, we observe that only 5 would be promoted from the bottom third to the top third in stage 2. Although the number is very small, it represents an increase by 150 percent relative to the baseline. Likewise, a smaller number of participants, 14, would be moved from the top third to below the top third.

Table 7: Variation in the composition of the top 3rd

	Baseline	Affirmative action
	N	N
Rank in S1: bottom 3rd; Rank in S2: top 3rd	2	22
Rank in S1: top 3rd; Rank in S2: below top 3rd	11	24
Rank in S1: bottom 3rd; Real rank in S2: top 3rd	-	5
Rank in S1: top 3rd; Real rank in S2: below top 3rd	-	14

In Table 8, we report the OLS regression results on participants' chances of being

ranked among the top third in stage 2. The results are consistent with our observations above, as affirmative increases the chances for an individual ranked in the bottom third in stage 1 of getting the high piece-rate payment in stage 2 by 0.655, statistically significant at 1 percent level.

Table 8: OLS regressions on the prob. of being in the top 3rd in stage 2

	(1a)	(1b)
Treatment **AA**	-0.190***	-0.237***
	(0.030)	(0.038)
Bottom 3rd	-0.396***	-0.376***
	(0.040)	(0.045)
AA x Bottom 3rd	0.624***	0.655***
	(0.076)	(0.068)
Male		-0.013
		(0.054)
Year 9		0.119**
		(0.054)
Indigenous		-0.056
		(0.047)
Math		0.039
		(0.032)
Checkbonus		-0.009
		(0.053)
Constant	0.440***	-0.653
	(0.017)	(0.486)
N	263	259
Adj. R^2	0.099	0.109

Robust standard errors in parentheses, clustered by session. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In Table 9, we report the regression results on the probability of being ranked among the top third in stage 2, based on the real rank in stage 2. Even though the estimate for the interaction term **AA x Bottom 3rd** is positive, it is not statistically different from zero.

Table 9: OLS regressions on the prob. of being in the top 3rd in stage 2, based on real ranks

	(1a)	(1b)
Treatment **AA**	0.003	-0.025
	(0.027)	(0.042)
Bottom 3rd	-0.396***	-0.373***
	(0.040)	(0.046)
AA x Bottom 3rd	0.062	0.072
	(0.071)	(0.068)
Male		-0.066
		(0.063)
Year 9		0.048
		(0.058)
Indigenous		0.027
		(0.050)
Math		0.021
		(0.035)
Checkbonus		0.021
		(0.058)
Constant	0.440***	-0.018
	(0.017)	(0.476)
N	263	259
Adj. R^2	0.131	0.119

Robust standard errors in parentheses, clustered by session. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Lastly, we describe how affirmative action impacted overall performance in Table 10. When looking at the whole sample, we do not observe a significant difference in average performance between treatments. The average number of correct answers in the baseline treatment in stage 2 is 27.5 and 28.03 in the affirmative action treatment (p-value=0.622). However, there are heterogeneous distributional effects. For participants ranked in the top third in stage 2, their average performance in the baseline treatment exceeds the one of their counterparts in the affirmative action treatment by 3.48 units (p-value=0.009). Similarly, for participants ranked below the top third in stage 2, their average performance in the baseline treatment is lower than in the affirmative action treatment by 1.96 units (p-value=0.072). These observations suggest that, by promoting individuals with a relatively low performance, affirmative action does not affect overall performance. Nonetheless, it decreases average performance of participants receiving the high piece-rate payment and increases the one of those receiving the low piece-rate payment. That is not unexpected given that, as described in Table 7, a very large number

of participants ranked in the bottom third in stage 1 are promoted to the top third in the second stage by affirmative action.

Table 10: Average performance in stage 2

	Baseline	Affirmative action	Diff.
Whole sample	27.50	28.03	0.53
Top 3rd	37.15	33.66	- 3.49***
Below top 3rd	23.32	25.28	1.96*

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4 Concluding remarks

In this paper, we investigated how affirmative action impacts performance in a simple real-effort task, with a competitive setting and monetary incentives. The experiments were conducted in the classroom with 263 students aged 13 to 15, in Australian schools with a large representation of indigenous students and similar low socioeconomic indicators. In the experiment, participants performed the same effort-based task twice. In each stage participants in the top third of the performance distribution received a high piece-rate payment, while the remaining participants received a low piece-rate payment. In the second stage, in half of the experimental sessions, our affirmative action treatment, we offered a starting advantage to all participants in the bottom third of the performance distribution in the first stage. The starting advantage increased the chances for the low performers to be ranked among the top third in the second stage of the task and thereby, receiving the high piece-rate payment. In the other half of the experimental sessions, our baseline treatment, the second stage was identical to the first stage. Our results indicate that participants who benefit from affirmative action increase their effort in the task. On average, affirmative action increases the gap in performance between the two stages, by approximately 50 percent relative to the baseline. Moreover, our results suggest that any potential discouraging effects on those who are indirectly penalised by affirmative action are small. With regards to efficiency, we do not find that affirmative reduces overall performance. However, it decreases average performance of participants in the top third who receive a high piece-rate payment, and increases performance of the remaining participants, receiving a low piece-rate payment. This is explained by the fact that our affirmative action policy led to a large promotion of low performing participants into the top third.

The novelty of our study lies on the specificity of the participant pool, high school

students with a large representation of indigenous Australians. Indigenous populations tend to be systematically disadvantaged in many relevant outcomes and are often the target of affirmative action policies. In Australia, many special programs are implemented in schools and universities to encourage and support indigenous students to exert effort in school and raise their educational aspirations. Also, both governmental and private employers, have non-binding affirmative action rules in place to favour participation of indigenous people. The specificity of our participant pool increases the significance and external validity of our findings. Another aspect that differentiates our study from the existing literature is that we apply affirmative action based on actual relative performance.

In sum, our study shows that affirmative action is an effective policy tool to encourage effort of low achieving students and thereby reduce patterns of inequality in education. Importantly, even a very salient affirmative action policy in favour of low achieving students at best only weakly discourages those above the eligibility threshold. A limitation of our study is that, due to sample size considerations, we are unable to test for heterogeneity in response to affirmative action of different social groups (for instance, indigenous *versus* non-indigenous, male *versus* female). This is an interesting question for future research, given the substantial evidence that individual characteristics are relevant predictors of behaviour in many contexts.

References

Altman, J. (2000). The economic status of indigenous australians. CAEPR Discussion Paper No. 193. Canberra: Australian National University.

Arcidiacono, P. (2005). Affirmative action in higher education: How do admission and financial aid rules affect future earnings? *Econometrica*, 73(5):1477–1524.

Balafoutas, L. and Sutter, M. (2012). Affirmative action policies promote women and do not harm efficiency in the laboratory. *Science*, 335(6068):579–582.

Bath, M. J. and Biddle, N. (2011). Measures of indigenous wellbeing and their determinants across the lifecourse. *CAEPR Lecture Series, The Australian National University.*

Bradley, S., Draca, M., Green, C., and Leeves, G. (2007). The magnitude of educational disadvantage of indigenous minority groups in australia. *Journal of Population Economics*, 20(3):547–569.

Calsamiglia, C., Franke, J., and Rey-Biel, P. (2013). The incentive effects of affirmative action in a real-effort tournament. *Journal of Public Economics*, 98:15–31.

Card, D. E. and Krueger, A. B. (2005). Would the elimination of affirmative action affect highly qualified minority applicants? evidence from california and texas. *Industrial and Labor Relations Review*, 58(3):416–434.

Coate, S. and Loury, G. C. (1993). Will affirmative-action policies eliminate negative stereotypes? *The American Economic Review*, 83(5):1220–1240.

Dulleck, U., He, Y., Kidd, M., and Silva-Goncalves, J. (2015). The impact of affirmative action: Evidence from a cross-country laboratory experiment. QuBE working paper.

Falk, A. and Fehr, E. (2003). Why labour market experiments? *Labour Economics*, 10(4):399–406.

Howell, J. S. (2010). Assessing the impact of eliminating affirmative action in higher education. *Journal of Labor Economics*, 28(1):113–166.

Ladson-Billings, G. (2006). From the achievement gap to the education debt: Understanding achievement in us schools. *Educational researcher*, 35(7):3–12.

Lazear, E. P. and Rosen, S. (1981). Rank-order tournaments as optimum labor contracts. *The Journal of Political Economy*, 89(5):841–864.

Leonard, J. S. (1990). The impact of affirmative action regulation and equal employment law on black employment. *The Journal of Economic Perspectives*, 4(4):47–63.

Long, M. C. (2004). College applications and the effect of affirmative action. *Journal of Econometrics*, 121(1):319–342.

Miller, A. R. and Segal, C. (2012). Does temporary affirmative action produce persistent effects? a study of black and female employment in law enforcement. *Review of Economics and Statistics*, 94(4):1107–1125.

Niederle, M., Segal, C., and Vesterlund, L. (2013). How costly is diversity? Affirmative action in light of gender differences in competitiveness. *Management Science*, 59(1):1–16.

Niederle, M. and Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *The Quarterly Journal of Economics*, 122(3):1067–1101.

Patrinos, Harry Anthony, H. G. (2004). *Indigenous Peoples, Poverty and Human Development in Latin America*. The World Bank.

Patrinos, Harry Anthony, S. C. N. (1992). North american indians in the canadian labour market: A decomposition of wage differentials. *Economics of Education Review*, 11(3):257–266.

Schaffner, M. (2013). Programming for experimental economics: Introducing CORAL-a lightweight framework for experimental economic experiments.

Schotter, A. and Weigelt, K. (1992). Asymmetric tournaments, equal opportunity laws, and affirmative action: Some experimental results. *The Quarterly Journal of Economics*, 107(2):511–539.

Steering Committee Review of Government Service Provision (2015). Report on Australian Government Productivity Commission Services.

A Appendix

A.1 Additional regression results

Table A.1: OLS regressions on performance

	Stage 2	
	(1a)	(1b)
Treatment **AA**	-0.434	-0.811
	(0.577)	(0.609)
Bottom 3rd	-3.166**	-2.810*
	(1.315)	(1.392)
AA x Bottom 3rd	2.699*	2.749**
	(1.310)	(1.192)
Stage 1: Nb of correct answers	0.806***	0.797***
	(0.066)	(0.071)
Male		-0.846
		(0.536)
Year 9		0.893
		(0.780)
Indigenous		0.134
		(0.645)
Math		0.025
		(0.283)
Checkbonus		0.962**
		(0.427)
Constant	8.388***	0.680
	(2.174)	(5.302)
N	263	259
Adj. R^2	0.557	0.548

Robust standard errors in parentheses, clustered by session.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.2: OLS regressions on performance

	Stage 1		Stage 2		Difference	
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)
Treatment **AA**	0.348	-0.254	-0.074	-0.936	-0.421	-0.682
	(0.904)	(0.720)	(1.085)	(1.052)	(0.540)	(0.658)
Bottom 3rd	-10.613***	-10.460***	-11.986***	-11.522***	-1.373	-1.062
	(0.939)	(0.922)	(1.375)	(1.308)	(0.910)	(0.956)
AA x Bottom 3rd	-1.957	-2.133	1.611	1.462	3.568**	3.595**
	(1.484)	(1.531)	(1.630)	(1.615)	(1.439)	(1.370)
Male		0.909		-0.214		-1.123*
		(0.744)		(0.912)		(0.603)
Year 9		1.915**		2.006**		0.091
		(0.870)		(0.807)		(0.796)
Indigenous		-0.190		0.176		0.365
		(0.801)		(1.174)		(0.774)
Math		0.585		0.231		-0.354
		(0.348)		(0.441)		(0.312)
Constant	28.899***	10.529	31.772***	14.439*	2.873***	3.910
	(0.801)	(7.948)	(0.904)	(7.350)	(0.368)	(6.465)
N	228	224	228	224	228	224
Adj. R^2	0.553	0.564	0.383	0.369	0.013	0.006

Robust standard errors in parentheses, clustered by session. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.3: OLS regressions on performance

	Difference (S2-S1)	
	(1a)	(1b)
Treatment **AA**	7.764*	7.278*
	(3.686)	(4.090)
Stage 1	0.002	0.022
	(0.052)	(0.062)
AA x Stage 1	-0.280**	-0.264*
	(0.123)	(0.132)
Male		-1.320*
		(0.730)
Year 9		-0.531
		(0.826)
Indigenous		0.288
		(0.838)
Math		-0.240
		(0.312)
Checkbonus		0.000
		(.)
Constant	2.807	8.137
	(1.753)	(7.439)
N	152	152
Adj. R^2	0.021	0.015

Robust standard errors in parentheses, clustered by session. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.4: OLS regressions on expected percentile ranks

	Stage 1		Stage 2		Difference	
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)
Treatment **AA**	-0.029	-0.045	-0.000	-0.024	0.029	0.022
	(0.032)	(0.036)	(0.034)	(0.032)	(0.020)	(0.021)
Bottom 3rd	0.009	-0.001	0.048	0.032	0.042	0.038
	(0.038)	(0.045)	(0.040)	(0.041)	(0.038)	(0.041)
AA x Bottom 3rd	-0.055	-0.030	-0.036	-0.012	0.011	0.008
	(0.046)	(0.048)	(0.043)	(0.051)	(0.051)	(0.056)
Male		-0.015		-0.021		-0.001
		(0.022)		(0.023)		(0.024)
Year 9		-0.009		0.004		0.016
		(0.035)		(0.027)		(0.027)
Math		0.019		0.007		-0.012
		(0.017)		(0.011)		(0.013)
Indigenous		0.100***		0.107***		-0.001
		(0.026)		(0.022)		(0.014)
Rank in stage 1	0.134*	0.133*				
	(0.066)	(0.073)				
Rank in stage 2			0.178**	0.187**		
			(0.064)	(0.067)		
Constant	0.288***	0.297	0.294***	0.228	0.030*	-0.057
	(0.035)	(0.322)	(0.036)	(0.227)	(0.017)	(0.247)
N	228.000	224.000	228.000	224.000	228.000	224.000
Adj. R^2	0.033	0.053	0.067	0.075	0.017	0.006

Robust standard errors in parentheses, clustered by session. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.5: OLS regressions on expected ranks

	Difference in expected ranks (S2-S1)			
	(1a)	(1b)	(2a)	(2b)
Treatment **AA**	0.009	-0.009	-0.001	-0.019
	(0.033)	(0.036)	(0.033)	(0.036)
Rank in stage 1	0.154***	0.122***	0.155**	0.124**
	(0.043)	(0.038)	(0.053)	(0.056)
AA x Rank in stage 1	0.041	0.067	0.085	0.106
	(0.082)	(0.074)	(0.083)	(0.090)
Male		-0.017		-0.021
		(0.024)		(0.027)
Year 9		0.019		0.016
		(0.025)		(0.023)
Math		-0.027**		-0.028**
		(0.013)		(0.013)
Indigenous		-0.009		0.001
		(0.029)		(0.030)
Checkbonus		0.010		
		(0.032)		
Constant	-0.023	-0.077	-0.025	-0.040
	(0.023)	(0.233)	(0.024)	(0.223)
N	172	172	152	152
Adj. R^2	0.038	0.043	0.057	0.070

Columns (1a) and (1b) give the coefficient estimates obtianed for the whole sample of participants ranked above the bottom third in stage 1; columns (2a) and (2b) give the coefficient estimates for the sample of participants ranked above the bottom third in stage 1, excluding those who incorrectly reported whether they received the extra points. Robust standard errors in parentheses, clustered by session. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A.2 Experimental instructions

Welcome

Thank you for taking part in this session! It will be fun and you will earn real money. You will get the money in cash at the end of the session.

What you will do:

You will perform the same task twice and then answer a short questionnaire. How much effort you put in the task will determine how much money you get. At the end of the session, one of the two stages of the task will be randomly picked for payment. Your final earnings will consist of $7 for your participation, plus up to another $7, depending on your performance in the stage picked for payment. We will now explain in detail what you will be doing in the task and go through examples. While we go through the instructions, please only proceed to the next screen when we ask you to do so. It is very important that you listen carefully to the explanations and you are encouraged to ask questions. Once you start performing the task, you can still raise your hand if you have any question and one of us will come and answer it privately. You must perform the task and answer all questions individually, so please do not talk to your neighbours during the session.

Wait screen

Stage 1: The task

We will now describe in detail stage 1.

You will have 8 minutes to answer questions similar to the one that you see now on your screen.

In each question, your task is to count the number of shaded squares in a grid with a total of 50 squares (the grid has 5 rows and 10 columns), and enter this number in the empty box as indicated on your screen. In the example, the number of shaded squares in the grid is 20. You would write the number 20 in the empty box.

Once you enter your answer, you need to press the button Continue on the bottom right of the screen. Then, a new grid will appear and you will be asked the exact same question.

You will have 8 minutes to answer as many questions as you can.

Do you have any questions? Is there something which you would like me to explain better?

OK, please press the button Continue.

Stage 1: Your earnings

Your earnings will depend on two things: your number of correct answers and whether you are one of the students in the room who had the highest number of correct answers in the task.

This is how your earnings will be calculated:

If your number of correct answers places you within the top third of students, you get 15 cents per correct answer. If you are not among the top third of students with the highest score, you get 5 cents per correct answer.

In this room you are X students. Those ranked A or above will receive 15 cents per correct answer. All the others will get 5 cents per correct answer.

Let's consider a first example.

A student answered 40 questions correctly and is ranked A (within the top third). He earns 40*$0.15.

Let's now turn to a second example.

A student answered 15 questions correctly and is ranked B (below the top third). He earns 15*$0.05.

What if two or more students get exactly the same score, how are the ranks allocated?

Ties will be broken randomly. For instance, if the 3rd ranked student answered 32 questions correctly, and two students answered 31 questions correctly, then one of them will be ranked 4th and the other 5th.

Do you have any questions? Is there something which you would like me to explain better?

OK, please press the button Continue.

Stage 1: Guess your rank

Once the 8 minutes are over, you will be asked to guess how well you did in the task compared to the other X students in the room. You can choose any number between 1 and X+1 for your rank.

For instance, if you think that you had the highest number of correct answers, your guess for your rank will be 1.

If you think that you had the second highest number of correct answers, your guess for your rank will be 2.

If you think that you had the lowest number of correct answers, your guess for your rank will be X+1.

You will need to enter the number corresponding to your guess for your rank in the box, as you can now see on your screen.

Do you have any questions? Is there something which you would like me to explain better?

OK, please press the button Continue.

Stage 1: Your score

At the end of Stage 1, you will be informed about your number of correct answers in Stage 1 and whether your score in Stage 1 was in the top third or not.

Stage 2: The task

We will now describe stage 2.

In stage 2, the task is exactly the same as in stage 1. Again, you will be given 8 minutes to solve as many questions as you can.

Do you have any questions? Is there something which you would like me to explain better?

OK, please press the button Continue.

Stage 2: Your earnings

The difference with stage 1 is that, in stage 2, your earnings may be determined differently than in stage 1. There are two possible cases:

- 1/3 of the students in the room receive 15 extra points added to their number of correct answers to calculate their score in stage 2.
- Nobody receives any extra points.

Let's consider the case where 1/3 of the students in the room receive the 15 extra points.

Scores are determined as follows:

If you receive the 15 extra points, your score will be determined by your number of correct answers, plus the 15 extra points.

If you do not receive the extra points, your score will be determined by your number of correct answers only.

If your score is within the top third, you get 15 cents per correct answer. If it is not, you get 5 cents per correct answer. Note that if you receive the extra points, it is easier for you to be in the top third than if you do not receive the extra points. However, you only receive money for your correct answers.

Let's consider an example:

A student received the extra points and answered 18 questions correctly in stage 2. His or her score will be equal to 18 + 15

If the score 33 is within the top third, he or she will receive 18 x 15 cents = \$2.7

If the score 18 + 15 is not within the top third, he or she will receive 18 x 5 cents = \$0.9

Let's consider another example:

A student did not receive any extra points and got 40 correct answers in stage 2.

His or her score will be equal to 40 + 0.

Again, if the score 40 is within the top third, he or she will receive 40 x15 cents, and 40 x 5 cents otherwise.

Let's now consider the case where nobody receives any extra points:

In that case, your earnings are determined exactly as in Stage 1.

Remember that in Stage 2 not everybody will receive the extra points. If extra points are given out, only 1/3 of the students will receive extra points. If no extra points are given out, nobody receives any extra points.

Before performing Stage 2, you will be informed whether any students received the extra points, and if so, whether you are one of these students.

Do you have any questions? Is there something which you would like me to explain better? OK, please press the button Continue.

Stage 2: Guess your rank

As in stage 1, you will be asked to guess how well you did in in Stage 2 compared to the other X students in the room. You can choose any number between 1 and X+1 for your

guess for the rank of your score.

You will need to enter the number corresponding to your guess for the rank of your score in the box, as you can now see on your screen.

Do you have any questions? Is there something which you would like me to explain better?

OK, please press the button Continue.

Final questionnaire

Once you have completed Stage 1 and Stage 2, you will be asked to answer a short questionnaire.

Last screen before starting

You will now start. Before performing each stage, you will have a short explanation screen reminding you the instructions for the stage you are about to perform.

You will also have screens indicating you to wait before continuing. When you are on a wait screen, please do not click continue before we invite you to do so. Also, it is very important that you stay quiet and do not talk to your neighbours. You will only have to wait a few seconds.

Remember that although you will have the possibility to earn money in both stages of the task, only one stage, randomly picked, will be paid in the end. If you have any question while performing the task, please raise your hand and one of us will come and answer your question privately.

Once you have finished, please remain seated until and quiet until everybody has completed all the tasks. Once everybody has finished, you will receive your earnings privately and then get back to your normal school activities.

Do you have any questions before starting the task?

If you are ready to start, please press the button READY.

Wait screen

Wait screen after stage 1:

You will now be told about your score in stage 1 and whether you were in the top third. You will then be told that:

AA:

1/3 of the students in the room will receive the extra points in stage 2. Please read the information displayed on the screens carefully as you will be informed whether you are one of the students who received the extra points

Baseline:

nobody will receive extra points in stage 2.

A.3 Screenshots

Welcome

Thank you for taking part in this session! It will be fun and you will earn real money. You will get the money (in cash) at the end of the session.

What you will do:

You will perform the same task twice and then answer a short questionnaire. **How much effort you put in the task will determine how much money you get.** At the end of the session, one of the two stages of the task will be randomly picked for payment. Your final earnings will consist of $7 for your participation, plus up to another $7, depending on your performance in the stage picked for payment.

We will now explain in detail what you will be doing in the task and go through examples. While we go through the instructions, **please only proceed to the next screen when we ask you to do so. It is very important that you listen carefully to the explanations** and you are encouraged to ask questions. Once you start performing the task, you can still raise your hand if you have any question and one of us will come and answer it privately.

You must perform the task and answer all questions individually, so please do not talk to your neighbours during the session.

Enter your participant number here:

Continue

WAIT

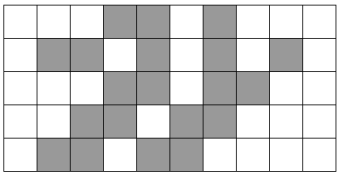
Please wait until instructed to continue.

Continue

Stage 1: The task

You will have **8 minutes** to answer questions similar to the one below.

How many squares in the grid are shaded?

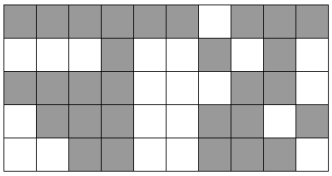


The number of shaded squares is:

Continue

Stage 1: The task

How many squares in the grid are shaded?



The number of shaded squares is:

You have **8 minutes** to answer as many questions as you can.

Continue

Stage 1: Your earnings

Your **earnings** will be determined as follows:

	Rank		Score
	1	→	Highest
	2	→	Second Highest
	3	→	Third Highest
$0.15 per correct answer	...	→	...
↑	0		
	1		
↓			
$0.05 per correct answer	...	→	...
	...	→	...
	...	→	...
	-1	→	...
	0	→	...
	1	→	Lowest

Continue

Stage 1: Guess your rank

You will be asked to **guess how well you did in the task compared to the other students**, as shown below.

What is your guess for your rank in Stage 1 of the task?

Rank		Score
1	→	Highest
2	→	Second Highest
3	→	Third Highest
...	→	...
0	→	...
1	→	...
...	→	...
...	→	...
...	→	...
...	→	...
...	→	...
1	→	Lowest

Continue

Stage 1: Your score

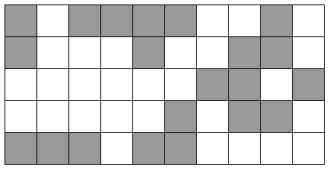
At the **end of Stage 1**, you will be informed about **your number of correct answers** in Stage 1 and **whether your score was in the top third or not.**

Continue

Stage 2: The task

Exact same task **as in Stage 1:**

How many squares in the grid are shaded?



The number of shaded squares is:

You have **8 minutes** to answer as many questions as you can.

Continue

Stage 2: Your earnings

How your earnings are determined in Stage 2 **may be different than in Stage 1**, as explained below.

Two possible cases

1/3 of the students in the room receive 15 extra points added to their number of correct answers to calculate their score in Stage 2	Nobody receives any extra points

Continue

Stage 2: Your earnings

How your earnings are determined in Stage 2 **may be different than in Stage 1**, as explained below.

Two possible cases

1/3 of the students in the room receive 15 extra points added to their number of correct answers to calculate their score in Stage 2	Nobody receives any extra points

Your Score:

If you **receive the 15 extra points** ⟶ **Score = Number of correct answers in Stage 2 + 15**

If you **do not receive any extra points** ⟶ **Score = Number of correct answers in Stage 2**

Your Earnings:

If your **score is within the top third** ⟶ **Earnings = $0.15 per correct answer**

If your **score is not within the top third** ⟶ **Earnings = $0.05 per correct answer**

Continue

Stage 2: Your earnings

How your earnings are determined in Stage 2 **may be different than in Stage 1**, as explained below.

Two possible cases

1/3 of the students in the room receive 15 extra points added to their number of correct answers to calculate their score in Stage 2	Nobody receives any extra points

Your Earnings:

Exactly as in Stage 1

Continue

Stage 2: Guess your rank

As in Stage 1, you will be asked to **guess how well you did in the task compared to the other students**, as shown below.

What is your guess for your rank in Stage 2 of the task?

Rank		Score
1	→	Highest
2	→	Second Highest
3	→	Third Highest
...	→	...
0	→	...
1	→	...
...	→	...
...	→	...
...	→	...
...	→	...
...	→	...
1	→	Lowest

Continue

Last screen before starting

You will now start. Before performing each stage, you will have a short explanation screen reminding you the instructions for the stage you are about to perform.

Remember that although you will have the possibility to earn money in both stages of the task, **only one stage, randomly picked, will be paid in the end.**

If you have any question while performing the task, please raise your hand and one of us will come and answer your question privately.

Once you have finished, please remain seated and quiet until everybody has completed all the tasks. **Once everybody has finished, you will receive your earnings privately** and then get back to your normal school activities.

Do you have any questions before starting the task?

If you are ready to start, please press READY.

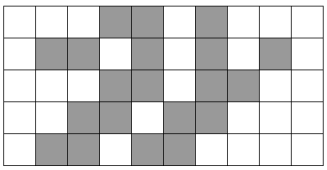
READY

Stage 1

Task

You have **8 minutes** to answer as many questions as you can, similar to the one below.

How many squares in the grid are shaded?



The number of shaded squares is:

Earnings

If your score is **within the top third**, you get **$0.15 for each correct answer**.

If your score is **not within the top third**, you get **$0.05 for each correct answer.**

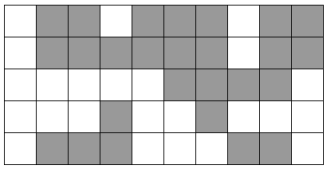
If you have understood what you are asked to do and are ready to start, please click on the START button below.

START

Stage 1

Time left: 1:54

How many squares in the grid are shaded?



The number of shaded squares is:

Continue

Stage 1: Guess your rank

Guess how well you did in the task compared to the other students.

What is your guess for your rank in Stage 1 of the task?

Rank		Score
1	→	Highest
2	→	Second Highest
3	→	Third Highest
...	→	...
0	→	...
1	→	...
...	→	...
...	→	...
...	→	...
...	→	...
...	→	...
1	→	Lowest

Continue

WAIT

Please wait until instructed to continue.

Continue

Stage 1: Your score

In **Stage 1**, you answered **0 questions correctly.**

Given your score, you are **not** ranked in the top third.

Continue

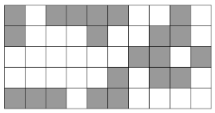
Stage 2

You will now perform **Stage 2**.

Task

As in Stage 1, you have **8 minutes** to answer as many questions as you can.

How many squares in the grid are shaded?



The number of shaded squares is:

Earnings

Nobody receives any extra points

Your score will correspond exactly to your number of correct answers.

As in Stage 1, if your score is within the top third, you get **$0.15 for each correct answer**. If your score is **not within the top third**, you get **$0.05 for each correct answer.**

If you have understood what you are asked to do and are ready to start, please click on the START button below.

START

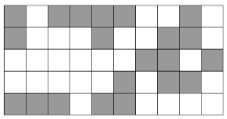
Stage 2

You will now perform **Stage 2**.

Task

As in Stage 1, you have **8 minutes** to answer as many questions as you can.

How many squares in the grid are shaded?



The number of shaded squares is:

Earnings

1/3 of the students in the room receive 15 extra points added to their number of correct answers to calculate their score in Stage 2

Continue

Stage 2

You are one of the students **receiving 15 extra points**

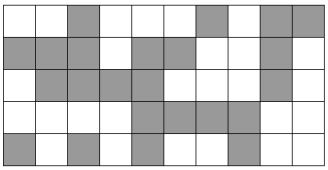
If you have understood the instructions please click on the START button below and you will start Stage 2. Otherwise please go back and read the instructions again.

Go Back START

Stage 2

Time left: 7:59

How many squares in the grid are shaded?



The number of shaded squares is:

Continue

Stage 2: Guess your rank

Guess how well you did in the task compared to the other students.

What is your guess for your rank in Stage 2 of the task?

Rank		Score
1	→	Highest
2	→	Second Highest
3	→	Third Highest
...	→	...
0	→	...
1	→	...
...	→	...
...	→	...
...	→	...
...	→	...
...	→	...
1	→	Lowest

Continue

WAIT

Please wait until instructed to continue.

Continue

We ask you to answer the following questions about yourself:

Are you: -

What is your date of birth? day: - month: - year: -

In which school year are you enrolled? -

Have you repeated any school year? -

Which grade did you get in your last report card for math? -

Which grade did you get in your last report card for English? -

Do you intend to complete Year 12? -

If you answered yes to the previous question, how far do you want to go with your education? -

Which job would you like to do once you leave school?

Do you primarily speak English at home? -

If you do not speak English at home, which language do you primarily speak?

Who do you live at home with? (Tick all that apply)

- Mother
- Father
- Stepmother
- Stepfather
- Brother(s)/sister(s)
- Grandparents
- Aunt
- Uncle

What is your mother's job?

What is your father's job?

Do you identify as Aboriginal and/or Torres Strait Islander? -

We now ask you what you thought of this session:

	Entertaining	A bit long	Boring
How did you like the task?			

	Easy	Not too hard	Hard
How did you find the task?			

	Easy to follow	A bit confusing	Very confusing
How did you find the initial instructions?			

Did you receive extra 15 points in Stage 2? -

Continue