Jobseekers' Beliefs about Comparative Advantage and (Mis)Directed Search*

Andrea Kiss ⓡ Robert Garlick ⓡ Kate Orkin ⓡ Lukas Hensel

August 12, 2024

Abstract

Worker sorting into tasks and occupations has long been recognized as an important feature of labor markets. But this sorting may be inefficient if jobseekers have imperfect information about their skills and therefore apply to jobs that poorly match their skills. To test this idea, we run two field experiments that give young South African jobseekers information on their results from standardized assessments of job-relevant skills. Half of untreated jobseekers believe they are better at skills in which their scores are low, relative to other jobseekers, and apply to jobs that poorly match their skills. The information treatment redirects jobseekers' search toward jobs that value skills where they score relatively highly – using measures from administrative, incentivized task, and survey data – without raising their search effort. Treatment also substantially raises earnings. These patterns are consistent with inefficient sorting due to imperfect information, which can be reduced by simple interventions.

*Andrea Kiss, Carnegie Mellon University, akiss@andrew.cmu.edu; Robert Garlick, Duke University, robert.garlick@duke.edu; Kate Orkin, Blavatnik School of Government, University of Oxford, kate.orkin@bsg.ox.ac.uk; Lukas Hensel, Guanghua School of Management, Peking University, lukas.hensel@gsm.pku.edu.cn. Order of authors is randomized and all authors contributed equally. For helpful comments, we thank Johannes Abeler, Peter Arcidiacono, Linda Babcock, Vittorio Bassi, Michele Belot, Matteo Bobba, Stefano Caria, Rebecca Dizon-Ross, Taryn Dinkelman, Simon Franklin, Clement Imbert, Laura Gee, Selim Gulesci, David Huffman, Jeremy Magruder, Barbara Petrongolo, Chris Roth, Jeffrey Smith, Duncan Thomas, and Basit Zafar, as well as seminar and conference participants at AFE, AFES, Bocconi, CEPR, Carnegie Mellon, Columbia, Duke, ESA, Essex, FIU, IFLAME, NBER development, NBER education, Notre Dame, Oxford, Peking, PHBS, Renmin, Pittsburgh, Royal Holloway, SEA, SITE labor, TCD and UC San Diego. We thank Alice Cahill, Raquel Caldeira, Aliya Chikte, Sabhya Gupta, Jenn Kades, Brynde Kreft, and Wim Louw for excellent research assistance. For collaboration on the broader research program, we thank Eliana Carranza and Neil Rankin and staff at J-PAL Africa and the Harambee Youth Employment Accelerator. This study has been approved by the ethics review boards at the University of Cape Town (REC 2020/02/001), Duke University (# D0368), University of Oxford (# ECONCIA15-055), and Carnegie Mellon University (# STUDY2022_00000166). The experiments used in this study are preregistered on the AEA's trial registry at https://www.socialscienceregistry.org/trials/10000/ and https://www.socialscienceregistry.org/trials/1631/. This paper has been produced through a research program financially supported by the World Bank Jobs Group, World Bank Africa Gender Innovation Lab, National Science Foundation (# 1824413), Private Enterprise Development in Low-Income Countries program (# 3024 and # 4728), UKRI GCRF Accelerating Achievement for Africa's Adolescents Hub, IPA Research Methods Initiative, and Upjohn Institute for Employment Research. Lukas Hensel acknowledges support by the National Natural Science Foundation of China (Grant No. 72394393).

1 Introduction

Worker sorting into tasks and occupations has long been recognized as an important feature of labor markets (Roy, 1951). Efficiently matching workers with the tasks where their skills are most productive offers the prospect of large output gains (Lise & Postel-Vinay, 2020). The efficiency of matching depends crucially on job search, particularly how jobseekers direct their search effort across different types of jobs.

In this paper, we use two field experiments to show that jobseekers can have imperfect information about their skills and that getting new information can update their beliefs, redirect their job search toward jobs that better match their skills, and substantially raise their earnings. Existing research, reviewed by Mueller & Spinnewijn (2023), has documented imperfect information about many aspects of job search but has not studied jobseekers' beliefs about their skill match with different job types. The idea that key labor market decisions may depend on imperfect information about skill match, shown in this paper, has broad applications that range from job search to education to migration.

We begin by proposing a simple model in which jobseekers with multidimensional skills have beliefs about their **skill comparative advantage** (CA): their ranks in different skill dimensions, relative to other jobseekers from similar backgrounds. These beliefs influence their **skill-directed job search**: how they allocate effort across searching for jobs with different skill demands. In this model, imperfect information about skill CA can distort skill-directed job search and worsen labor market outcomes.

To empirically study the consequences of limited information about skill CA, we recruit young jobseekers from disadvantaged backgrounds in Johannesburg. We invite them to day-long workshops run with a South African government job search assistance agency. In the workshops, we measure their communication and numeracy skills using established psychometric assessments. We verify that these skills are valued by firms hiring for non-specialist, entry-level jobs – the same types of jobs that jobseekers in our sample are searching for. We define each jobseeker's CA as the skill dimension in which they rank highest relative to other jobseekers from similar backgrounds who take the same assessments at this agency.[1] We also measure jobseekers' beliefs about their levels of these skills and their CA over these skills. We find that jobseekers' beliefs about their CA (1) are only slightly closer to their assessment results than random guesses, (2) persist over time, and (3) predict the skill demand of jobs where they apply. These descriptive patterns are consistent with the model and help to motivate our experimental analysis

In two field experiments, we randomize half the workshops to give participants their skill assessment results, shown as rankings relative to the assessment results of a large

[1] This definition follows work in both labor (e.g. Guvenen et al. 2020; Papageorgiou 2014) and education (Altonji et al., 2016). We discuss alternative definitions and the link to absolute advantage in Section 2.

group of jobseekers from similar backgrounds. Participants in control workshops take the same assessments but do not learn their assessment results. The two experiments have the same treatments, comparable groups of jobseekers recruited in the same way but at different times, and complementary but not identical outcome measures.[2]

Our first experiment shows that receiving information about one's skills facilitates skill-directed job search, using a unique combination of belief and novel search measures for 278 jobseekers. No existing data sources measure all variables necessary to study how jobseekers' beliefs influence how they direct search effort over jobs with different skill demands. To achieve this, we measure jobseekers' multidimensional skills, beliefs about these skills, jobseeker $\times$ vacancy-level data on whether jobseekers apply to specific vacancies, and these vacancies' skill demands.[3] Relative to the control group, treated participants' skill beliefs move substantially closer to their measured skills, including beliefs about their CA. Treated participants are more likely to apply to jobs with skill demands that match their CA, using multiple prespecified measures of skill-directed job search: applications on a job search platform, survey data, and a novel incentivized task in which participants choose between applying to jobs with different skill requirements. These results are driven by a prespecified subset of jobseekers: those whose baseline CA beliefs differ from their assessment results. These are the jobseekers who get more information from treatment and, to the extent that our assessments are accurate, have inaccurate baseline beliefs. Treatment closes roughly 80% of the gap in skill-directed job search between jobseekers with accurate and inaccurate baseline beliefs.

Our second experiment shows that receiving information about one's skills improves labor market outcomes, using survey data for 4,389 jobseekers collected 3.5 months after the workshops. Treated jobseekers have weekly earnings and hourly wages roughly 25% higher than control group jobseekers, with suggestive evidence that they move onto different short-run job ladders. These effects are driven by the subset of jobseekers whose baseline CA beliefs differ from their assessment results, as in the first experiment. Treatment effects on jobseekers' CA beliefs and skill-directed job search are similar across the two experiments, although the second experiment relies on simpler survey measures.

We call the two experiments "tight" and "big" to emphasize their respective strengths: detailed, novel measures of beliefs and skill-directed search in the first experiment versus a larger sample and longer timeframe to study labor market outcomes in the second.[4] We

[2] We interpret the treatments as providing information only to jobseekers because it is difficult for participants to credibly share assessment results with firms and we show that they rarely share these results.

[3] Survey datasets typically record aggregated search measures such as total applications submitted, not jobseeker*vacancy-level measures. Job search platforms seldom measure skills, beliefs, or labor market outcomes. Government administrative data do not measure beliefs and seldom measure skills or search.

[4] We use these light-hearted, nonstandard terms because standard terms don't capture the differences between our experiments. Both are field experiments, not lab-in-the-field, because they study real jobseekers' real search. Both have elements of what Harrison & List (2004) call framed and natural field experiments.

ran the tight experiment after the big experiment to better measure the belief and search effects that might explain the strikingly large earnings effects in the big experiment.

The two experiments show a consistent picture of this labor market: getting more information about their skills moves jobseekers' CA beliefs toward their measured CA, redirects their search toward jobs aligned with their measured CA, and raises their earnings. In contrast, we find evidence against a plausible alternative mechanism: that treatment changes beliefs about absolute skill *levels* and therefore shifts search effort and labor market outcomes. We see the first part of this mechanism in our data: treatment lowers the average jobseeker's belief about her skill level, because untreated jobseekers are on average overconfident about their skill levels relative to other jobseekers. But we do not see the second part of this mechanism: treatment has negligible effects on multiple measures of job search effort in both experiments. Our model shows that this can occur because the sign of the effect of beliefs about skill levels on search effort is theoretically ambiguous. We find little evidence that treatment changes jobseekers' skill investment decisions.

We contribute to a broad and long-standing research program studying barriers to efficient matching between workers and jobs. Our contribution is to provide the first direct evidence that jobseekers' beliefs about their skill CA influence their job search, in particular how they direct search over different job types, and therefore influence their labor market outcomes. We collect for the first time the unique combination of data needed to observe this causal chain. This bridges and extends three existing literatures.

First, one literature studies the labor market *consequences* of (mis)match between workers' multidimensional skills and jobs' skill demands. This work uses longitudinal labor market data and dynamic matching models to show that mismatch can lead to large long-term earnings losses.[5] But this literature typically infers the *causes* of mismatch from model-based arguments or does not specify them, as it does not observe data on beliefs, search, or sometimes even skills. We complement this work by designing experiments and collecting unique measures to understand how skill mismatch can be caused by beliefs about skills and job search choices, allowing us to to directly observe behavior otherwise requiring assumptions.[6] And we highlight one way policy might address skill mismatch: our skill information intervention raises the average treated participant's earnings by 1.8 times the average variable cost of the intervention, under plausible assumptions about the time path of earnings. This raises the possibility that similar interventions

[5]See Sanders & Taber (2012) for a review of this literature. Böhm et al. (2023) use direct skills measures to show that returns to skills vary substantially between firms, consistent with an important role for mismatch. Baley et al. (2022), Fredriksson et al. (2018), Guvenen et al. (2020) and Lise & Postel-Vinay (2020) combine direct skill measures with dynamic models to understand life-cycle earnings losses from mismatch.

[6]Our tasks build on work using choices in controlled environments to study preferences over education and jobs (e.g. Adams-Prassl & Andrew 2023; Cortés et al. 2023; Mas & Pallais 2017; Wiswall & Zafar 2015.)

might be cost-effective additions to active labor market programs or job search platforms.

A second literature studies imperfect information in education and hiring as potential causes of (mis)match. Education economists have shown that imperfect information about students' skills and skill match with specific occupations can influence education investments.[7] Education research has focused on decisions during formal education rather than decisions in the labor market, which depend on potentially very different learning processes. Labor economists have studied the implications of firms' imperfect information about applicants' skills for hiring and wage-setting decisions.[8] Labor research has focused almost entirely on firms' information about jobseekers' *level* of a single skill, rather than the multidimensional skill *match* that we emphasize. And labor research has focused almost entirely on firm-side decisions, although a few recent papers have noted that job search decisions may also react when the market acquires new information about jobseekers' skills (Abebe et al., 2021b; Bassi & Nansamba, 2022; Carranza et al., 2022).

We use some of the same data as our companion paper, Carranza et al. (2022) hereafter C22, but the two papers have substantially different goals and methods. C22's goal is to *separately identify firm-side responses to new information about jobseekers' skills from jobseeker-side responses.* To achieve this, C22 uses multiple, mostly firm-facing experiments to show that employment and earnings rise when firms learn more about jobseekers' skills, holding constant jobseekers' information. C22 also reports average treatment effects from our big experiment on employment and earnings, in order to show that these outcomes change less when only jobseekers receive information than when firms also do. C22 discusses jobseekers' beliefs and search but never uses the idea of skill CA and relies on a single measure of beliefs about skill levels and a single, self-reported, categorical measure of search direction. It acknowledges that this evidence merely "suggests ... but does not prove" a link between jobseekers' beliefs, search, and labor market outcomes (p. 3549).

In contrast, this new paper's goal is to *show how jobseekers' limited information about their skill CA can distort their search direction and worsen their labor market outcomes*. To achieve this, we propose a simple model of job search with CA beliefs over multidimensional skills. We test the model predictions using the entirely new tight experiment, new outcome measures from the big experiment, and new model-motivated heterogeneous treatment effects on jobseekers with different baseline beliefs about their skill CA in both

[7]For example, beliefs about skills can influence education enrollment and expenditure (e.g. Arcidiacono et al. 2016; Berry et al. 2022; Bobba & Frisancho 2022; Dizon-Ross 2019; Franco 2019) and beliefs about skill CA can also influence subject and major choices (e.g. Altonji et al. 2012, 2016; Aucejo & James 2021; Delaney & Devereux 2021; Saltiel 2022; Stinebrickner & Stinebrickner 2014).

[8]Hiring and wage-setting can change when firms observe new information about workers' skills from job performance (Altonji & Pierret, 2001; Arcidiacono et al., 2010; Hardy & McCasland, 2023; Kahn & Lange, 2014), references and referrals (Abel et al., 2020; Heath, 2018; Ioannides & Loury, 2004; Pallais, 2014), education qualifications (Alfonsi et al., 2020; Clark & Martorell, 2014; Jepsen et al., 2016; MacLeod et al., 2017), and skill certification (Abebe et al., 2021b; Bassi & Nansamba, 2022; Groh et al., 2015).

experiments. We also test and reject a different special case of the model where beliefs influence search effort and hence labor market outcomes.

The third related literature studies relationships between beliefs and job search.[9] This literature focuses on jobseekers' beliefs about the *level* of their labor market prospects, captured by job offer arrival rates or wage offer distributions. It has not studied beliefs about skills or skill CA, despite the centrality of sorting on skill CA in labor economics (Roy, 1951). And it typically studies levels of search effort, rather than how search is directed across different job types and what this implies for job-worker matching. Recent work has shed important light on search direction by studying the labor market consequences of nudging jobseekers to consider alternative occupations.[10] Unlike this work, we directly measure jobseekers' skills and beliefs and use information about their skills to help them to direct search over different types of jobs.

Section 2 describes our model, context, and sample characteristics. In Section 3 we show the relationship between CA beliefs and skill-directed job search in the tight experiment. In Section 4 we show the relationship between CA beliefs, skill-directed job search, and labor market outcomes in the big experiment. We show that treatment has little effect on search effort in Section 5 and on other possible mechanisms in Section 6. Section 7 concludes with reflections – on general equilibrium, generalizability, and markets for information about skills – that might inform future research.

2 Economic Environment

We begin with a conceptual framework and then describe our context, sample, and skill assessments. We then report four key descriptive patterns that inform our conceptual framework: jobseekers' skills vary across multiple dimensions, different firms value different skill dimensions, jobseekers' beliefs about their skills persistently differ from results on assessments, and jobseekers' beliefs about their skills predict their search decisions. These four patterns motivate the structure of our conceptual framework, which shows how jobseekers' imperfect information about their CA over multiple skills can distort skill-directed job search and worsen their labor market outcomes.

[9] This includes research on the co-evolution of search and search-related beliefs in panel data (Adams-Prassl et al., 2023; Conlon et al., 2018; He & Kircher, 2023; Mueller et al., 2021; Spinnewijn, 2015); on experiments providing information about labor market conditions (Altmann et al., 2018; Jäger et al., 2023; Jones & Santos, 2022); on search subsidies, matching services and technologies, or mentoring programs that influence multiple outcomes including jobseekers' beliefs (Abebe et al., 2022; Alfonsi et al., 2022; Bandiera et al., 2023; Banerjee & Sequeira, 2023; Beam, 2016; Caria et al., 2024; Kelley et al., 2023; Kroft & Pope, 2014; Vyborny ⓡ al., 2023; Wheeler et al., 2022); and on jobseekers' beliefs about attributes of specific jobs (Bazzi et al. 2021; Boudreau et al. 2023; Chakravorty et al. 2023; Sockin & Sojourner 2023; Subramanian 2022). Beliefs about skills have also been studied in workplace decisions (e.g. Hoffman & Burks 2020; Huffman et al. 2022; Malmendier & Tate 2015) and in lab settings with tasks that mimic job search (e.g. Falk et al. 2006).

[10] See Altmann et al. (2022), Behaghel et al. (2022) Belot et al. (2019, 2022) and Le Barbanchon et al. (2023).

In the conceptual framework, descriptive analysis, and tight experiment, we study jobseekers' beliefs about CA over their two skills – communication and numeracy – and their search over jobs demanding these skills. Two skill dimensions allow simplicity and are the minimum needed to study CA and skill-directed search. These skills suit our research: they are non-specialist skills used in many jobs in this economy, they are weakly correlated with each other, and different firms value them differently. The big experiment shows that results generalize to a less stylized setting where jobseekers receive information about six skills and search over jobs requiring many skills.

2.1 Conceptual Framework

Our conceptual framework combines elements from recent models of "partially directed job search," where jobseekers try to direct search to higher-wage vacancies but face uncertainty about wages (Lentz et al., 2022; Wu, 2021) with models of subject choice in education, reviewed by Altonji et al. (2012, 2016). We use a static partial equilibrium framework focusing on jobseekers' search and beliefs, and motivate this choice below.

We assume that each jobseeker has communication and numeracy skill levels S_C and S_N. Each job demands primarily communication or primarily numeracy skills.

Search over jobs demanding different skills: Jobseekers split fixed total search effort $\bar{E}$ between search for communication jobs E_C and numeracy jobs E_N. We generalize the framework later to allow endogenous choice of total effort. Searching for a type j job yields outcome $V_j(S_C, S_N, E_j)$, which captures the expected present value of a job offer multiplied by the probability of an offer. We make three assumptions. First, V_j is increasing and concave in all three arguments and $\partial V_j/\partial S_j > \partial V_j/\partial S_i > 0$ for $j \neq i$. This assumption allows both types of jobs to value both skills, but each job type to value one skill more. Second, we assume that skill and search effort are technical complements and are 'more complementary' within than across dimensions. Intuitively, a jobseeker with high communication skills will get a higher return to directing marginal search effort to communication than numeracy jobs and vice versa. Formally,

$$\frac{\partial^2 V_j}{\partial S_j \partial E_j} > \frac{\partial^2 V_i}{\partial S_j \partial E_i} > 0 \tag{1}$$

for $j \neq i$. Third, we assume that gross utility from job search $U(V_C, V_N)$ is increasing and concave in both arguments. This allows jobseekers to value the outcomes of searching for both job types without fully specifying the offer acceptance decision or reservation wage.

Under the first and third assumptions, jobseekers direct search effort to equalize the marginal utility of searching for each job type:

$$\frac{\partial U}{\partial V_C} \times \frac{\partial V_C}{\partial E_C} = \frac{\partial U}{\partial V_N} \times \frac{\partial V_N}{\partial E_N}, \tag{2}$$

where $\frac{\partial U}{\partial V_j}$ captures the jobseeker's preferences over nonpecuniary aspects of job type j. Conditional on these preferences, marginal search effort will be directed based on the relative magnitudes of $\frac{\partial V_C}{\partial E_C}$ and $\frac{\partial V_N}{\partial E_N}$. Under the second assumption, $\frac{\partial V_C}{\partial E_C}$ is more steeply increasing in communication skill than $\frac{\partial V_N}{\partial E_N}$. This means that if a jobseeker's communication skill rises, the left-hand side of the optimality condition in (2) will rise more than the right-hand side. The jobseeker will then increase E_C and decrease E_N to restore equality in condition (2), because of the first assumption that V_j is a concave function of E_j.

Jobseekers' beliefs about their skills: We assume each jobseeker has beliefs about their skill levels $\tilde{S}_C$ and $\tilde{S}_N$. They allocate search effort based on these beliefs, not their actual skill levels. We do not model belief formation, including belief updating in response to search outcomes, because we show in Section 2.5 that control group jobseekers learn little about their skills from search outcomes over the timeframe of our study.

Testable predictions: Our framework predicts two effects of a jobseeker receiving accurate new information about her skill levels, for example, learning that her communication skill is higher than she previously thought. First, she will redirect search effort away from numeracy-heavy jobs and toward communication-heavy jobs, because condition (1) means that her expected relative return to search is now higher for communication-heavy than for numeracy-heavy jobs. Second, her labor market outcomes will improve, because more of her search effort is now directed to job types that will reward her skills more. Note that these outcomes are driven by learning about her communication skill *relative* to numeracy skill. Learning about her *average* skill level across the two dimensions is irrelevant in this framework because total search effort is fixed. In Section 5, we generalize the framework to allow her belief about her average skill level to endogenously influence total search effort but show empirically that this does not occur in our data.

Skill comparative advantage (CA): We define a jobseeker as having a CA in communication if she ranks higher in the distribution of communication than numeracy skills and vice versa. This definition aligns with the spirit of standard definitions of CA in trade. There, a country has a CA in the product it can produce at lowest opportunity cost. Here, a jobseeker has a CA in the skill where she ranks higher, because she can supply it to the market at a lower opportunity cost in terms of time spent supplying the other skill.

This definition aligns with our empirical measures, described in Section 2.2. We focus on relative ranks in each skill distribution rather than absolute scores on assessments because absolute scores are sensitive to the difficulty of each assessment (Nielsen, 2023). This means that "absolute advantage" is not a well-defined concept: *scoring* higher in a communication than numeracy assessment does not necessarily show higher communication than numeracy skill. But *ranking* higher in a communication than numeracy assessment does show higher communication than numeracy skill, relative to other test-takers.

This definition takes advantage of our multidimensional skill measures and follows research using similar data in education (e.g. Altonji et al. 2016) and labor economics (e.g. Guvenen et al. 2020).[11] Readers could rename this measure from skill CA to "relative skill rank" if they prefer it, without changing our theoretical predictions or empirical analysis.

Using this definition of CA, our framework's predictions become that a jobseeker who gets accurate new information about her skill CA will (1) redirect search effort toward jobs that demand skills aligned with her CA and (2) have better labor market outcomes.

Before proceeding, we reflect on the framework's goals and scope. The framework motivates our experimental design by providing a precise definition of skill levels, skill CA, and beliefs about them; and by generating testable predictions about how these influence search. We could add more structure and estimate the $U(\cdot)$ and $V(\cdot)$ functions to quantify the sorting process or the welfare costs of imperfect information. However, this would crowd out other parts of our analysis. We view this as best left for follow-up work.

2.2 Context, Sample, and Skill Assessments

Context: We work in Johannesburg, part of the 14-million-person metropolitan area of Gauteng. Wage labor is the primary income source, self-employment is low, entry-level employment is mainly in services and manufacturing, and most employment is in formal firms, not always with formal contracts (SSA, 2022). Unemployment in Johannesburg is high: 40.5% for ages 15–34 at the time of the tight experiment (SSA, 2022).[12]

Target population: Our target population is active jobseekers who are likely to face limited information about their skills because they have limited higher education or work experience that might provide this information. In Section 7.2, we informally discuss if and how our results might generalize outside this context and target population, noting that many countries face low-information, high-unemployment labor markets.

Sample recruitment: We recruited from participant registries at the Harambee Youth Employment Accelerator, a public-private partnership that provides job search assistance. Harambee maintains a database of active jobseekers aged 18-34 with legal permission to work in South Africa who graduated from schools in low-income areas. Firms receive free access to the database for recruiting. The database captures a sizeable propor-

[11] Alternative approaches estimate occupation-specific wages by education and use this to define CA in occupations based on education levels (e.g. Acemoglu & Autor 2011). These approaches can also price the value of different types of workers in different types of jobs. But their wage estimates are conditional on the way workers currently sort into occupations. These may be sensitive to a Lucas-style critique that wages might be different under a different type of sorting, which is precisely the mechanism that we study. We show in Section 2.4 that demands for proxies of communication and numeracy skills are roughly equal in this setting. This suggests that defining CA using relative skill ranks and using skill prices might not produce different classifications in this setting.

[12] We use Statistics South Africa's definition of employment: engaging in any income-generating activity during the reference week. Unemployment rates exclude those in full-time study or not in the labor force.

tion of the population of interest: in Gauteng in 2022, restricting to ages 18-34, there were 1,078,745 jobseekers in the database and 1,403,064 unemployed people (SSA, 2022). These groups do not perfectly overlap because some jobseekers in the database are employed and some unemployed people do not sign up on the database.

We recruited a sample of 4389 people for our big experiment in September 2016 – April 2017 and another sample of 278 people in July – October 2022 for our tight experiment, using the same sampling frame and method. We contacted people from Harambee's database who lived within commuting distance of our field office in downtown Johannesburg. We invited them to a day-long job search assistance workshop, stating that the workshops would include receiving job search advice and taking assessments that could be used to match them to suitable vacancies. Harambee often runs workshops like these.

Data collection: In both experiments, participants complete baseline surveys about their demographics, beliefs about their skills and labor market prospects, recent job search activities, current employment, and employment history. They also take skill assessments in person. We discuss post-treatment measurement in sections 3 and 4.

For the tight experiment, we also observe participants' search behavior on the online job search and matching platform SAYouth.mobi. Harambee used its database of participants to set up this platform in 2019. The platform aggregates job advertisements from all online job boards in South Africa, allows jobseekers access without incurring mobile phone data charges, and allows firms to post advertisements and set up interviews.

Sample characteristics: We focus here on descriptive statistics for the tight experiment sample, shown in Table 1. Descriptive statistics for the big experiment sample are similar (see Section 4.1). Jobseekers are young, with 90% aged 21–32. Two patterns suggest that these jobseekers are likely to face limited information about their skills. First, 60% have no post-secondary education. There is little specialization by subject in secondary schools, limiting their scope to learn about their skills from specialized training. Grades and grade progression are only weakly correlated with results on independent skill assessments (Lam et al., 2011) and career counselling is rare (Pillay, 2020). There is a national secondary school graduation exam but grades are not strongly informative: they correlate weakly with results in post-secondary education and firms report that the grades convey limited information about skills (Schoer et al., 2010). Second, most jobseekers have limited, mostly informal work experience: 33% were employed at baseline but only 13% had a formal written contract and only 25% had *ever* held a long-term wage job. This limits their scope to have learned about their skills from work experience, the main mechanism in models of jobseeker learning in other contexts (e.g. Baley et al. 2022).

Their job search effort was high but met with limited success. 96% were actively searching. In the week before baseline, the average jobseeker submitted 10 job applica-

Table 1: Summary Statistics – Tight Experiment

	Mean (1)	Median (2)	Min (3)	Max (4)	SD (5)	Obs. (6)
Panel A: Demographics						
Black African	1.00	1.00	1.00	1.00	0.00	278
Male	0.33	0.00	0.00	1.00	0.47	278
Age	26.41	26.00	18.00	36.00	4.04	278
Completed secondary education only	0.60	1.00	0.00	1.00	0.49	278
University degree / diploma	0.22	0.00	0.00	1.00	0.41	278
Any other post-secondary qualification	0.15	0.00	0.00	1.00	0.36	278
Panel B: Labor market background						
Any work in last 7 days	0.33	0.00	0.00	1.00	0.47	278
Has worked in permanent wage job before	0.25	0.00	0.00	1.00	0.43	278
Earnings in USD (last 7 days, winsorized)	44.90	0.00	0.00	697.57	101.71	277
Written contract	0.13	0.00	0.00	1.00	0.33	278
Panel C: Search behavior						
Any job search in last 30 days	0.96	1.00	0.00	1.00	0.20	278
# applications (last 7 days, winsorized)	10.00	5.00	0.00	100.00	14.93	278
Search expenditure in USD (last 7 days, winsorized)	22.72	14.00	0.00	126.00	23.72	278
Hours spent searching (last 7 days, winsorized)	13.82	9.00	0.00	72.00	15.00	278
# job offers (last 30 days, winsorized)	0.17	0.00	0.00	4.00	0.56	278
Panel D: Skills beliefs						
Aligned belief about CA	0.49	0.00	0.00	1.00	0.50	278
Fraction aligned belief domains	0.22	0.00	0.00	1.00	0.31	278

Notes: Table 1 shows baseline summary statistics for the tight experiment. 'CA' stands for comparative advantage. Winsorization is at the 99^{th} percentile. All monetary values are in 2021 USD PPP.

tions, spent 14 hours searching, and spent 22.72 USD on search online and offline, including transport to drop off CVs and attend interviews and mobile phone airtime and data.[13] This search cost is high: roughly 20% of what a full-time minimum wage job would pay. This matches other work on job search costs in South Africa (Banerjee & Sequeira, 2023; Kerr, 2017). High search costs limit scope for jobseekers to learn about their skills through search and raise costs of misdirected job search. Under 1% of job applications yield offers.

Skill assessments: The numeracy assessment captures practical arithmetic. It was developed by a large retail chain to assess potential cashiers. The communication assessment captures English-language listening, reading, and comprehension skills at a high school level. It was developed by an adult education provider (www.mediaworks.co.za). Candidates also complete a "concept formation" assessment that captures fluid intelligence: the ability to identify patterns across situations and to use logic in new situations (Raven & Raven, 2003; Taylor, 1994). These assessments are also used in the big experiment, along with three more assessments described in Section 4.2. Appendix A.1 describes the assessment process and psychometric properties of the assessments, including

[13] All monetary values throughout the paper are in 2021 USD in purchasing power parity terms.

showing that scores are roughly normally distributed without ceiling or floor effects.

All our information treatments and measures of skill beliefs use assessment results relative to a benchmark population: roughly 12,000 jobseekers from similar backgrounds who have also been assessed by Harambee. We place jobseekers' communication and numeracy skills in quintiles relative to the distribution of assessment scores in this population. We define each jobseeker's CA as the skill in which they score in a higher quintile. This approach is designed to approximate each jobseeker's CA relative to likely competitors for similar jobs, because Harambee's database consists of jobseekers from similar backgrounds, searching for similar types of jobs, and likely covers the majority of young jobseekers in this province.

2.3 Skills Vary Substantially across Dimensions within Jobseeker

These two assessments differentiate jobseekers horizontally more than vertically because communication and numeracy scores are only moderately correlated with each other ($\rho = 0.31$). Hence the assessments show which jobseekers are better suited for jobs where each skill is more important, rather than identifying jobseekers who are likely to be better at most jobs. Concept formation is equally correlated with both other skills (ρ = 0.33–0.38), suggesting neither assessment captures fluid intelligence better.

2.4 Firms Value Different Skills for Different Jobs

We have already shown that the assessments mainly horizontally differentiate jobseekers: they identify jobseekers' relative strengths across different skills. Here we summarize evidence that firms value the skills that we measure and that there is variation in which firms value which skills, with details provided in the Online Supplement. Jointly, these patterns suggest scope for jobseekers to improve their labor market outcomes by searching for jobs that value the skills in which they perform best.

Firms value communication and numeracy skills: In an incentivized resume-ranking experiment, 91% of firms preferred job applicants with higher communication or numeracy assessment results to job applicants with lower assessment results and an additional one-year post-secondary training qualification. Over 500 client firms have paid Harambee to screen roughly 1 million jobseekers using these assessments, which we interpret as revealing a preference for using these skills in hiring. And firms prefer job applications with certified communication and numeracy scores (Carranza et al., 2022). We do not argue that communication and numeracy are the most important skills or the only relevant skills in this labor market, just that they are important.

These patterns are consistent with the national high school graduation exam not giving firms sufficient information on applicants' skills. The communication and numeracy skill quintiles are positively but weakly associated with jobseekers' self-reported grades on their graduation exams in English and mathematics, respectively (Table A.3, columns 1–2). The positive association suggests the Harambee assessments capture meaningful variation in skills; the weak association highlights that jobseekers and firms both have scope to learn from information on jobseekers' ranks on Harambee's assessments.

Different firms value different skills: In the same incentivized ranking experiment, 58% of the firms in our sample ranked a resume with high numeracy skills ahead of a resume with high communication skills, and 42% of firms had the opposite ranking. This cross-firm variation, combined with the fact that our assessments horizontally differentiate jobseekers on communication versus numeracy skills, creates scope for jobseekers to improve their labor market outcomes by searching for different types of jobs.

There is roughly equal demand for numeracy and communication skills in online job postings. To show this, we compiled all 69,000 vacancies for entry-level jobs in Johannesburg posted during our tight experiment on any South African job search platform. We classified 13% as numeracy- and 14% as communication-heavy jobs, using a method we describe in Section 3.4. We cannot construct wage-based measures of skill demand because few postings specify wages, as in many economies (Batra et al., 2023).

Firms can at least partly observe skills. We embedded a measurement exercise in one firm's hiring process. The firm substantially prefers applicants for positions with skill requirements that match the applicants' assessed CA, relative to positions that do not match their CA, even though the firm does not directly observe the assessment results.[14]

Together, these patterns suggest that redirecting jobseekers' search towards jobs that match their CA in skills has the potential to improve their labor market outcomes.

2.5 Jobseekers' Perceived and Measured Comparative Advantage in Skills Differ

We have already shown scope for skill-directed job search to improve jobseekers' labor market outcomes. Here we show that jobseekers' beliefs about their skills do not match their skill assessment results, which might distort search direction.

We measure jobseekers' beliefs about their communication and numeracy skill quintiles before they take assessments. We first define the skills, then explain the concept of quintiles, define the reference group, and ask jobseekers which quintile they are in on each skill, relative to the reference group. See Appendix A.3 for measurement details and Figure 1 for the sequencing of belief measurement, assessments, and treatment. We do not refer to jobseekers' beliefs about skills as "accurate" or "inaccurate," as no assessment perfectly measures skills, even the well-established assessments we use. Instead, we refer to beliefs about skills as "aligned" or "misaligned" with skill assessment results.

The joint distribution of jobseekers' skills and beliefs is a multidimensional object that

[14] This shows firms can partly but not fully observe applicants' skills, in line with Carranza et al. (2022).

can be described in many different ways. We focus here on two simple, binary measures of belief (mis)alignment that are guided by our conceptual framework. We show robustness checks later using other, continuous measures of belief (mis)alignment.

Jobseekers have misaligned beliefs about their CA between skills: We define a jobseeker as having an **aligned CA belief** about her CA over communication and numeracy if she believes she is in a higher quintile for the same skill in which she actually scores in a higher quintile on our assessments. Using this definition, 49% of jobseekers have aligned CA beliefs at baseline (Table 1, panel D).[15]

Assessment results are misaligned with jobseekers' beliefs both about their *skills* and about their *assessment results*. The measures above capture jobseekers' beliefs about their communication and numeracy skills. We also ask jobseekers which quintile they fall in on our communication and numeracy assessments, after taking the assessments but before we tell any jobseekers their assessment results. We show in Appendix A.3 that the two measures are strongly positively related, suggesting that jobseekers view the assessments as capturing relevant parts of their general skills.

Jobseekers have misaligned beliefs about their average level over the two skills: For each of the two skills, we construct an indicator equal to one if the perceived and assessed skill quintiles are equal. We average these and call this measure the **fraction of aligned beliefs**, which can take values of 0, 0.5, and 1. Using this measure, 22% of beliefs about skills are aligned (Table 1, panel D).[16] Misalignment is explained more by overconfidence than underconfidence: 63% of jobseekers' beliefs about their skills are above their assessed skills and only 19% are below. Similar patterns of overconfidence, sometimes called "overplacement," have been documented in other work, reviewed by Santos-Pinto & de la Rosa (2020).

Jobseekers learn little about their skills while searching: To show this, we use the big experiment's baseline and follow-up surveys. The share of control group jobseekers whose believed and assessed CA align hardly changes over 3.5 months, even for the employed and those with above-median search effort (Table A.6, columns 1–3). The fraction of aligned beliefs is similarly persistent (columns 4–6). Slow learning might reflect limited scope for feedback during search: only 3% of jobseekers report ever receiving feedback about their skills during an unsuccessful job application. It may also reflect the well-documented difficulty of Bayesian learning about a multi-input function, such as the

[15] If jobseekers guessed randomly, 40% would have aligned CA beliefs. The benchmark is not 50% because tied CA beliefs are misaligned, as we explain in Section 3.1.

[16] CA beliefs and the fraction of aligned beliefs are positively correlated by construction but have substantial separate variation ($\rho = 0.21$). To see the separate variation, consider a candidate who has measured skills in quintile 2 in communication and 4 in numeracy, and has aligned beliefs. She scores 1 on both belief measures. Raising her believed numeracy quintile without changing her communication belief will decrease the fraction of aligned beliefs to 0.5 without changing aligned CA.

search outcome-skill relationship (Banerjee & Sequeira, 2023).

Jobseekers' skill beliefs seem to draw on both school results and other information: Even with good information from the schooling system, high school graduation exam scores should not perfectly predict young adults' beliefs about their skills: many jobseekers took school exams multiple years ago, the exams and assessments do not test identical skills, and jobseekers may not perfectly recall their exam scores. Indeed, jobseekers' beliefs about their communication and numeracy skills are positively but weakly correlated with their self-reported results on the graduation exams in English and mathematics (Table A.3, columns 4–5). Beliefs about CA are positively associated with the difference in scores between the two exam subjects (columns 6–7).

Skill beliefs do not substantially vary by gender: In both the tight and big experiments, we find limited gender differences in baseline skill beliefs, with or without controls for assessment results, demographics, and education. See the Online Supplement for details. This matches a recent metastudy showing limited gender differences in confidence (Bandiera et al., 2022). This motivates our gender-pooled analysis in this section.

2.6 (Misaligned) Beliefs Predict Skill-Directed Job Search

We have already shown that there is scope for skill-directed job search to improve jobseekers' labor market outcomes but that jobseekers' beliefs about their skills don't match their assessment results. Here we show that the gap between jobseekers' beliefs about their skills and assessment results might shift their skill-directed job search.

CA beliefs predict job search decisions: In the big experiment, we ask candidates what skill is most valuable for the types of jobs for which they are applying. Control group candidates' answers are strongly associated with their beliefs about their CA in skills. They are 8–10pp ($p < 0.01$) more likely to state that they are applying for jobs that value the *skill in which they believe they have a CA*, compared to jobs valuing other skills (Table A.4, rows 1–2). The correlations are robust across skill domains and to controlling for assessed CA and demographics. These results suggest jobseekers try to search for jobs for which they believe they have a skill CA.

CA on assessments weakly predicts job search decisions: Jobseekers are only 2–5 pp more likely to state that they are applying for jobs that value the *skill in which they have a CA on our assessments*, compared to jobs valuing other skills (Table A.4, rows 3–4). These results suggest that jobseekers' skill beliefs might direct their search away from jobs that match their assessed skills.

Furthermore, jobseekers with higher skills expect better labor market outcomes, using several different measures and econometrics approaches (Table A.5).

3 Tight Experiment: Effects on Beliefs and Directed Search

The previous section's conceptual framework and descriptive evidence suggest that jobseekers' beliefs about their skills might influence their search direction and outcomes. Here we use the tight experiment to study how jobseekers' skill beliefs and search direction react to new information about their skills. We collect rich data on beliefs and unique measures of skill-directed search using choices between jobs with different skill demands.

3.1 Experimental Design and Intervention

We ran the experiment during 34 day-long job search workshops attended by 373 jobseekers. Our main analysis uses the 278 jobseekers who have a unique CA, i.e., who score different quintiles for the communication and numeracy assessments. We impose this restriction because our measures of skill-directed job search, described in Section 3.4, can only be neatly defined for jobseekers with a unique skill CA. However, including these jobseekers in the sample and using an approximate definition of skill-directed job search for them produces similar treatment effects on our main outcomes (Table C.1).

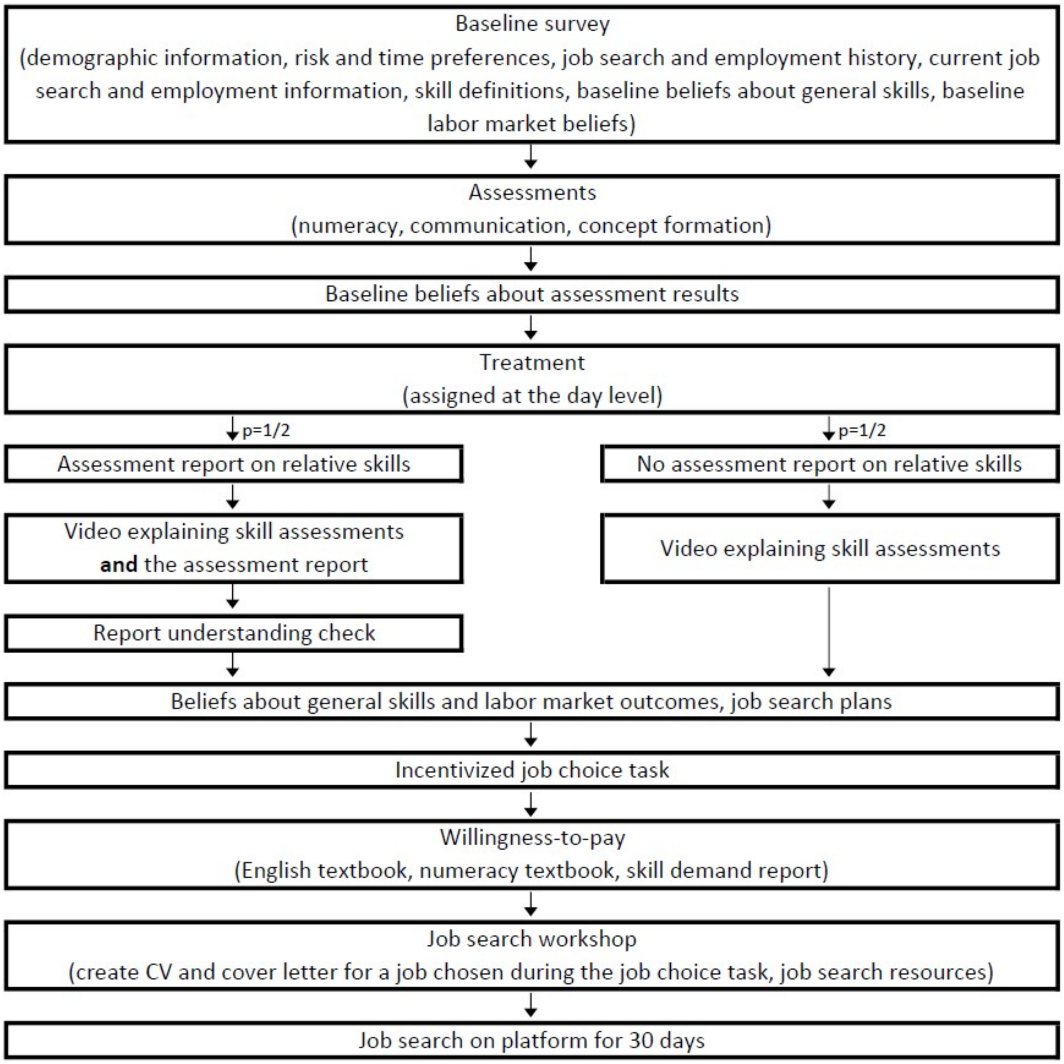
We randomized treatment at the workshop level to avoid spillovers between jobseekers. We assigned 17 workshops to treatment and 17 to control; the sample restriction discussed above does not drop any workshops. Treatment assignments are balanced on baseline covariates, for both the 278 and the 373 jobseekers (Table B.1).

Jobseekers knew they were participating in a research study but not an experiment and did not know the treatment assignment of their workshop. So treatment-control differences in outcomes should not reflect experimenter demand effects and attendance should not, and indeed does not, vary by treatment assignment. All jobseekers were reimbursed for transport and the time spent at the workshop, at above minimum wage.

The timeline of the day is shown in Figure 1. Jobseekers first completed a baseline survey on demographics, education, employment, job search, and beliefs. To measure beliefs about skills, we defined communication and numeracy skills and the concept of quintiles. We then asked jobseekers their beliefs about which quintile their general communication and numeracy skills are in, compared to the reference group. Jobseekers then took assessments of their numeracy, communication, and concept formation skills and completed another survey about their perceived performance on the assessments.

Jobseekers in treated workshops then received a report describing the assessments and their performance (Figure 2). For each skill, the report shows the quintile in which the jobseeker ranked on each assessment, compared to other jobseekers in the reference group. They watched a video that explained the skill assessments and how to interpret the report, particularly the quintiles. The video encouraged them to think about what jobs will value their skills. But it did not provide any information about which specific

Figure 1: Timeline of Activities during Tight Experiment Workshops



jobs value which skills or encourage them to apply to any specific job titles or types.

Jobseekers in control workshops did not receive a report. They watched a control video that was identical to the treatment video other than omitting the parts that explained how to interpret the report with assessment results. Both groups took the same assessments and answered the same skill-focused survey questions. This holds constant across groups any priming effects about the importance of skill levels and skill match and any learning from taking assessments.[17]

To facilitate comprehension, we intensively piloted reports and videos and gave jobseekers time to ask questions during and after the video. After the video, we asked treated jobseekers three understanding checks: 99% and 96% correctly reported the quintiles they

[17]Taking the assessments does generate some learning: the share of control group jobseekers with aligned CA beliefs about their skills increases by 10.1pp from the baseline to endline survey. This means that our intervention might generate larger effects relative to a counterfactual of jobseekers who were not assessed.

Figure 2: Sample Report

REPORT ON CANDIDATE COMPETENCIES
-Personal Copy-

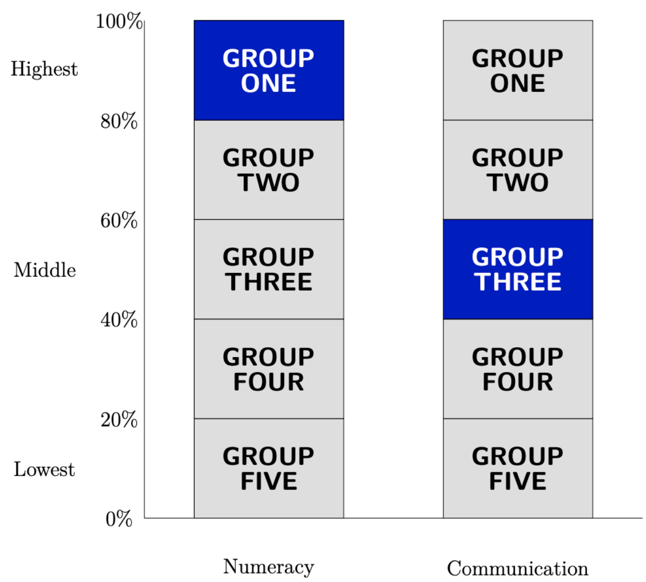
This report contains results from the assessments you took today. These results can help you learn about some of your strengths and weaknesses and inform your job search.

You completed assessments on English Communication (listening and reading comprehension) and Numeracy today.

1. The Numeracy test measures various maths abilities.
2. The Communication test measures English language ability through listening and reading comprehension.

Your results have been compared to a large group of young South African job seekers who have a matric certificate, have completed school in rural areas or townships around Johannesburg and have completed the same assessments.

You scored in the FIRST GROUP of these candidates for Numeracy and the THIRD GROUP for Communication.



Note: igure 2 shows an example of the reports given to treated jobseekers. Each report contains the jobseeker's assessment results but no identifying information (name, national identity number, etc.) and does not include any Harambee branding. "Completed school in rural areas or townships" is a common proxy in South Africa for attending school in a low-income area.

scored for respectively numeracy and communication, and 98% correctly reported the skill in which they scored higher. Belief updating does not differ by concept formation scores, suggesting fluid intelligence does not limit processing of this information.

After the videos, jobseekers completed another survey covering beliefs about their skills, beliefs about their labor market prospects, and job search plans. Then they completed several job search tasks we describe later. Appendix B contains a detailed description of the workshops and hyperlinks to the videos.

The report is designed to provide information only to the jobseekers themselves, not to prospective employers. The report does not include the jobseeker's name or any identifying information and has no Harambee branding. We show in Section 6 that information acquisition by firms is unlikely to explain the results of the experiment.

3.2 Specification

We estimate the effects of receiving information about one's relative skill ranking:

$$Y_{id} = T_d \cdot \beta + \mathbf{X}_{id} \cdot \Pi + \epsilon_{id}. \tag{3}$$

β, the average treatment effect, is the main object of interest. Y_{id} is the outcome for jobseeker i assessed on date d, T_d is a treatment indicator, and $\mathbf{X}_{id}$ is a vector of prespecified baseline covariates.[18] We use heteroskedasticity-robust standard errors clustered by workshop date, the unit of treatment assignment.

We also test whether treatment effects are larger for jobseekers whose baseline beliefs are misaligned with their CA, because these jobseekers receive more information from treatment. We estimate prespecified models of the form:

$$Y_{id} = T_d \cdot \alpha^{misaligned} + T_d \cdot Aligned_{id} \cdot \alpha^{diff} + Aligned_{id} \cdot \alpha + \mathbf{X}_{id} \cdot \Psi + \epsilon_{id}. \tag{4}$$

$Aligned_{id}$ is an indicator for jobseekers whose pre-treatment beliefs about their CA on the assessments match their assessment results (measurement details in Appendix A.3). We report the average treatment effect for jobseekers with misaligned baseline CA beliefs, whom we expect to be most affected by treatment ($\alpha^{misaligned}$); the average treatment effect for jobseekers with aligned baseline CA beliefs ($\alpha^{misaligned} + \alpha^{diff}$); and the difference between the treatment effects for the two subgroups (α^{diff}). We control for baseline values of the outcome, so our specifications capture the common belief updating models that regress posterior belief alignment on prior belief alignment, treatment, and their interaction (Haaland et al., 2023). We show later that our results are robust to using a continuous measure of misalignment between baseline beliefs and assessment results.

$Aligned_{id}$ is weakly correlated with other characteristics, so α^{diff} likely captures heterogeneity by baseline CA beliefs, not other characteristics. Specifically, Table A.7 shows using OLS, ridge, and lasso regressions that $Aligned_{id}$ is only weakly related to the assessment scores and is unrelated to gender, age, education, employment history, past feedback from actual and prospective employers, and family members' employment.

Both estimating equations, all baseline covariates, and most outcome measures are

[18] $\mathbf{X}_{id}$ contains age; a dummy for being female; dummies for only high school education, having a post-secondary certificate, and for having a post-secondary degree; dummies for each of the skill quintiles for each of numeracy and communication skills; a pre-treatment value of the outcome Y_{id} where available; and block fixed effects, to account for the fact that we randomize treatment within blocks of 4 sequential days. Results are unchanged if we add non-prespecified controls for baseline and past employment.

Table 2: Treatment Effects on Beliefs About Skills - Tight Experiment

	Aligned CA belief				Fraction aligned beliefs	
	Implied		Direct			
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.135***	0.208***	0.108**	0.261***	0.078***	0.036
	(0.035)	(0.050)	(0.048)	(0.068)	(0.026)	(0.028)
Treatment × Aligned CA belief (bl)		-0.137		-0.298***		0.080
		(0.082)		(0.093)		(0.050)
Aligned CA belief (bl)		0.586***		0.461***		-0.050
		(0.079)		(0.109)		(0.046)
Treatment effect: Aligned CA belief (bl)		0.072		-0.037		0.116***
		(0.058)		(0.058)		(0.041)
Control mean	0.475	0.475	0.532	0.532	0.183	0.183
Observations	278	278	278	278	278	278

Notes: Table 2 shows that treatment aligns jobseekers' beliefs about skills with their assessed skills in the tight experiment. "CA" stands for comparative advantage and "bl" stands for baseline. Columns show different outcomes: Cols. 1-4 show effects on dummies indicating if a jobseeker's belief about her CA in skills is aligned with her assessed CA. Cols. 1-2 use beliefs about skill quintiles to construct CA beliefs. Cols. 3-4 use a direct measure of CA beliefs. Cols. 5-6 show effects on the fraction of skills where her believed and assessed quintile are equal. Cols. 1, 3, and 5 show estimates from equation (3) and cols. 2, 4, and 6 show estimates from equation (4), with the treatment effect on jobseekers with aligned CA beliefs shown in the fourth row. Control variables are defined in footnote 18. Standard errors clustered at the treatment-day level shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

prespecified at https://www.socialscienceregistry.org/trials/10000/. We describe the relationship between the prespecified and final analysis in the Online Supplement.

We also estimate treatment effects on our main outcomes separately for jobseekers with CA in each skill and find no strong evidence of differences.[19]

3.3 Information about Skills Aligns Beliefs with Assessed Comparative Advantage

In the control group, 47.5% of jobseekers have aligned CA beliefs: they believe they rank in a higher quintile for the skill in which they score in a higher quintile on our assessments. Treated jobseekers are on average 13.5pp more likely to report aligned beliefs, a 28% increase on the control group mean (Table 2, column 1, $p < 0.001$).

Most of the treatment effect is driven by jobseekers with misaligned CA beliefs at baseline. Treatment increases the share of these jobseekers with aligned CA beliefs at endline by 21pp (44% of the control mean, $p < 0.001$). This is the estimate of $\alpha^{misaligned}$ from equation (4), shown in Table 2, column 2, row 1.

[19] Of our 278 jobseekers, 172 and 106 have CA in respectively communication and numeracy. CA is defined relative to a reference group of 12,000 jobseekers, so these 278 jobseekers need not have equal shares with CA in each skill. We do not view the unequal shares as a central factor for understanding our results for two reasons. First, treatment effect estimates are broadly similar for jobseekers with CA in each skill. Second, roughly equal shares of the jobseekers in the big experiment have communication and numeracy CA, and treatment effects are similar for outcomes measured in both experiments.

In contrast, treatment has modest effects on jobseekers with aligned beliefs at baseline. The share of this group with aligned beliefs increases by only 7.2pp ($p = 0.214$). This is the estimate of $\alpha^{misaligned} + \alpha^{diff}$ from equation (4), shown in Table 2, column 2, row 4. The difference in treatment effects between jobseekers with aligned and misaligned baseline beliefs is large – 13.7pp, or 28.8% of the control mean – but not quite statistically significant ($p = 0.105$). This is the estimate of α^{diff}, shown in Table 2, column 2, row 2.

These patterns are even stronger using an alternative measure of jobseekers' beliefs about their CA: directly asking in which skill they believe they have a CA. Using this measure, treatment increases the probability of holding aligned beliefs by 26pp ($p < 0.001$) for jobseekers with misaligned beliefs at baseline and has statistically significantly lower, near-zero effect on jobseekers with aligned beliefs at baseline (column 4).

Treatment shifts beliefs about their skill levels as well as beliefs about skill CA. 18% of control group jobseekers' beliefs about their skill levels match their assessed skill quintiles. Treatment increases this by 7.8pp, a 43% increase (column 5, $p = 0.005$).[20]

Treated jobseekers' post-treatment beliefs do not exactly match their assessment results. This is unsurprising, as beliefs about skills will naturally depend on multiple information sources. We discuss other information jobseekers use in Appendix A.5.

3.4 Job Search with Better-Aligned Beliefs about Skills

We show here that treatment shifts jobseekers' search toward jobs that align with their assessed CA across four measures of skill-directed search.

1) Job choice task: We design a novel incentive-compatible job search task in which participants make 11 choices between paired job adverts. In each pair, one job had been coded by recruiters as requiring more numeracy skills and one as requiring more communication skills. Participants were shown each pair of adverts, given time to read them, and asked to select one to apply to. Figure 3 shows an example pair and Table 3 shows all 22 job titles. Participants viewed the 11 pairs of job adverts in random order.

The adverts in the task are based on real job adverts posted on SAYouth.mobi. All jobseekers in the study used the platform, so choices between jobs represented real-life choices they often made. We considered all adverts for entry-level jobs in Johannesburg with no specialized education requirements. Among these, we selected a subset of job adverts with a clear numeracy or communication skill requirement, recognizing that not all jobs on the platform have such requirements. We asked 13 recruitment professionals with experience hiring for entry-level roles to rate each job on required communication and numeracy skills, transparency of skill requirements, expected wage, and overall desir-

[20] We show heterogeneous effects on this outcome by baseline CA beliefs (column 6) but do not focus on them, because baseline CA beliefs do not reflect individuals' scope to learn about their skill levels.

Table 3: Job Titles Used in the Job Choice Task - Tight Experiment

Numeracy job title	**Communication job title**
Receiving and dispatching clerk	Sales agent
Sales teller	Customer service agent
Stock controller	General administrator
Laundry assistant	Waiter / waitress
Cashier	Host/hostess
Data capturer	Front desk assistant
Restaurant till manager	Receptionist
Store cashier	Sales assistant
Cash teller	Recruitment administrator
Banking call center agent	Retail call center agent
Petrol attendant	Maintenance assistant

Notes: Table 3 shows the job titles for numeracy and communication-heavy jobs in each job pair in the job choice task in the tight experiment. Each job pair contains one relatively numeracy-heavy and one communication-heavy job title. The pairs were created by matching jobs with different skill requirements but similar wages and general appeal to workers to ensure sufficient variation in job choices. The pairs are based on a list of 28 jobs rated by HR professionals for their skill requirements, expected wage, overall and gender-specific desirability, and transparency of skill requirements.

ability (reflecting expected non-wage attributes). We then created 11 pairs of jobs. Within each pair, one job required more communication skill and one required more numeracy skill: averaging across all pairs, recruiters scored the job we defined as numeracy-heavy as needing 2.7 standard deviations higher numeracy skills.[21] But both jobs in each pair were similar in other ways: the average within-pair difference in expected wages was <5% of the mean wage and the average within-pair difference in desirability was even smaller. We removed employer names and locations and standardized length and format.

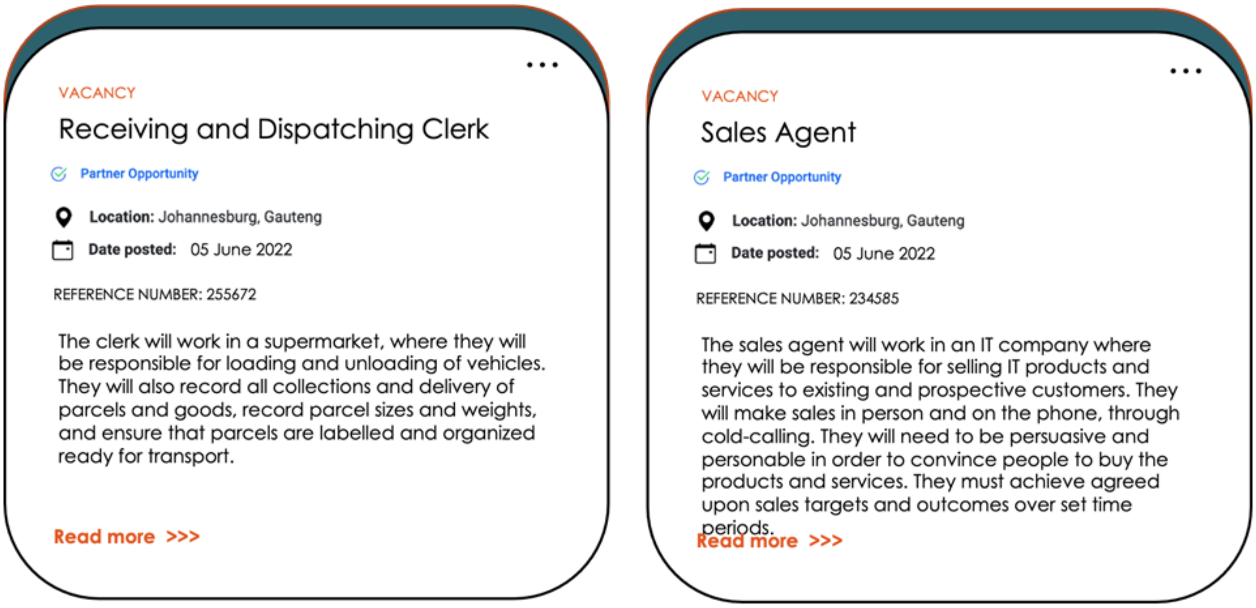
The job pairs thus differed primarily in their skill demands, while treatment gave information only on jobseekers' relative ranks on communication and numeracy skills. This allows us to measure how jobseekers' skill-directed job search changes in response to information about their skills, holding other job attributes constant. In non-experimental datasets, jobs with different skill demands may also differ on unobserved characteristics that drive observed search direction, making it difficult to study this question.

We incentivized jobseekers to respond truthfully in the job choice task in two ways. First, one pair contained live advertisements for jobs at a partner firm. We told jobseekers we would submit their application to the job they chose from this pair but did not tell them which pair it was. Second, we told jobseekers that after the workshop we would send them recommendations for entry-level job titles matching their choices in the task.

Our main prespecified outcome is the share of the 11 pairs in which the jobseeker

[21] We obtain a similar result using classifications from O*NET, a US-based dataset that classifies job titles based on their required English language and mathematics knowledge.

Figure 3: Sample Pair of Jobs from Job Choice Task



chose the job that required the skill aligned with her measured CA. In the control group, this share was 55%. Treatment raises this by 3.7pp (Table 4, column 1, $p = 0.285$). It is unsurprising that even treated jobseekers choose some jobs not aligned with their CA: they may have preferences over other job attributes or want a diversified application portfolio.

This average effect is driven by jobseekers with misaligned baseline CA beliefs, following the pattern we saw for treatment effects on aligned CA beliefs and matching the predictions from our conceptual framework. Specifically, for jobseekers with misaligned baseline beliefs about their CA, treatment increases the share of aligned choices by 8.8pp, 16% of the control mean (column 2, $p = 0.028$). Treatment does not change search direction for jobseekers with baseline aligned beliefs.

The search direction results suggest that jobseekers can interpret information in postings about the relative skill demands of different jobs. We run two tests to confirm this. First, we asked every jobseeker to score the communication and numeracy skills required by both jobs in a subset of pairs. This identifies the jobseeker's belief about which job is more communication-heavy. For 73% of the jobseeker$\times$pair data points, the jobseeker's belief matched the recruitment professionals' classification. Second, we explicitly revealed relative skill demand for the last two pairs of jobs (recall that the order in which the job pairs were shown was randomized). Jobseekers with initially misaligned beliefs align their search for job choices both with and without revealed skill demands, but substantially more when skill demands are revealed (Table C.6). This shows they have imperfect information about job skill requirements but enough for some skill-directed job search.

2) Application data from online search platform: We linked each jobseeker to their

Table 4: Treatment Effects on Search Direction - Tight Experiment

	% aligned (job choice)		Δ % aligned platform apps		Δ SMS click rate		Δ planned apps (w)		Aligned search index	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment	0.037	0.088**	0.063**	0.089*	0.071	0.157	1.420	4.746**	0.269**	0.603***
	(0.034)	(0.038)	(0.023)	(0.048)	(0.061)	(0.096)	(1.261)	(1.763)	(0.103)	(0.140)
Treatment × Aligned CA belief (bl)		-0.104**		-0.048		-0.163		-6.629**		-0.648***
		(0.039)		(0.080)		(0.128)		(2.651)		(0.205)
Aligned CA belief (bl)		0.165***		0.013		0.100		8.697***		0.727***
		(0.035)		(0.049)		(0.102)		(2.412)		(0.153)
Treatment effect:		-0.016		0.041		-0.005		-1.883		-0.045
Aligned CA belief (bl)		(0.034)		(0.044)		(0.081)		(1.807)		(0.131)
Control mean	0.550	0.550	0.007	0.007	-0.032	-0.032	4.331	4.331	-0.000	-0.000
Observations	278	278	278	278	278	278	278	278	278	278

Notes: Table 4 shows that informing jobseekers about their CA in skills aligns their search direction with their assessed CA in the tight experiment. "CA" stands for comparative advantage and "bl" stands for baseline. Aligned job search is defined as directing search effort toward jobs that mostly require the skill that aligns with jobseekers' assessed CA. Columns indicate different outcomes: the percentage of 11 incentivized job choices that are aligned with the measured CA of the jobseeker (cols. 1–2), the difference between the percentage of aligned and non-aligned applications on the online job search platform SAYouth.mobi (cols. 3–4), the difference in link click rates between aligned and non-aligned jobs sent to job seekers via text message (cols. 5–6), the difference between aligned and non-aligned planned applications for the 30 days after the workshop (cols. 7–8), and an inverse-covariance weighted average of the search alignment measures displayed in cols. 1–8 following Anderson (2008) (cols. 9–10). Even-numbered columns show heterogeneity by whether individuals have aligned CA beliefs at baseline. Control variables are defined in footnote 18. All variables marked (w) are winsorized at the 99^{th} percentile. Standard errors clustered at the treatment-day level shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

profile on SAYouth.mobi to observe their on-platform job search. In the 30 days after the workshop, the average participant in both the treatment and control groups started 15 job applications on the platform, slightly higher than in the preceding 30 days. The platform does not always record if applications are completed so we focus on started application.

We classified vacancies by skill demand where possible. There were 69,000 vacancies for Johannesburg posted online during our study period. Of these vacancies, we classified 14% as communication-heavy and 13% as numeracy-heavy jobs, suggesting roughly comparable demand for both skills among jobs posted online.[22] 23% of applications started by our jobseekers were to jobs classified as requiring either of these skills.

Treated jobseekers were more likely to start applications to jobs demanding skills that matched their measured CA. To show this, we calculate each jobseeker's number of applications to vacancies coded as requiring the skill aligned with their CA, subtract the number of applications to vacancies coded as requiring the opposite skill, and divide this difference by the number of applications to vacancies coded as requiring either skill. Treated jobseekers start 6.3pp more aligned than non-aligned applications (Table 4, column 3, $p = 0.010$). Again, this effect is driven by jobseekers with misaligned baseline CA beliefs, who have an effect of 8.9pp (column 4, $p = 0.074$). We obtained these data directly from the platform after the experiment, so experimenter demand effects are unlikely.

3) Clicks on links to real jobs: This measure captures whether jobseekers can conduct skill-directed job search using only job titles. We sent jobseekers three text messages with links to real job opportunities on SAYouth.mobi about a week after the workshop. We sent, in random order, one numeracy job, one communication job, and one job aligned with the skill demand of the majority of their choices in the job-choice task. The messages contained a greeting linking the message to the workshop they attended, a note that we found a job opportunity of interest within commuting distance of the workshop venue, the job title, the link to apply, and the #SAYouth hashtag (which Harambee regularly uses in messages to platform users). We tracked whether jobseekers clicked on these links.

Treatment increases this measure of skill-directed job search: the difference in click rates between jobs that are aligned and misaligned with the jobseeker's CA rises by 7pp (Table 4, column 5, $p = 0.254$). This is again driven by jobseekers with misaligned beliefs at baseline. However, estimates are less precise because this measure uses 3 choices per jobseeker, fewer than in the job choice task or on-platform search.

4) Planned applications to numeracy and communication jobs: After the treatment, before the job choice task, we surveyed participants about the number of applications they

[22]We classify a vacancy as numeracy-heavy if it contains one of 20 numeracy-heavy job titles, and similarly for communication. The lists of skill-specific job titles were created by the recruitment staff who classified vacancies for the job choice task. Most job postings omit wage information so we cannot construct wage-based measures of demand for communication and numeracy skills in this market.

planned to send to communication-heavy and numeracy-heavy jobs in the next 30 days. We calculate the number of planned applications to jobs aligned with their assessed CA, minus the number of planned applications to jobs focused on the other skill. Treatment increases this by 1.42 applications, a 33% increase on the control mean (Table 4, column 7, $p = 0.269$). This average effect is driven by a 4.8 application effect for jobseekers with misaligned baseline CA beliefs (column 8, $p = 0.011$). The result reflects both an increase in planned aligned applications and a reduction in planned misaligned applications. This measure is more susceptible to experimenter demand effects but produces similar results to the other measures.

Search direction index: We combine these four measures of skill-directed search into an index to avoid multiple hypothesis testing and to increase power (Anderson, 2008). We see a large, positive treatment effect of 0.27 standard deviations (Table 4, column 9, $p = 0.014$). This is entirely driven by jobseekers with misaligned baseline CA beliefs, for whom skill-directed search rises by 0.6 standard deviations (column 10, $p < 0.001$). Treatment closes roughly 80% of the skill-directed search gap between jobseekers with aligned versus misaligned baseline beliefs: control jobseekers with aligned baseline CA beliefs have an average value of 0.73 standard deviations (column 10, row 3), while treated jobseekers with misaligned baseline CA beliefs have an average of 0.60 (column 10, row 1).

3.5 Robustness Checks and Additional Results

We provide some additional results for readers interested in these specific topics.

Robustness checks for beliefs and search results: Hypothesis test results are robust to using a wild cluster bootstrap to account for having only 34 clusters and to accounting for multiple hypothesis testing (Table C.2). Treatment effect estimates are robust to replacing our binary measures of (mis)alignment between assessed skills and beliefs about skills with more continuous measures, both for the outcome variables and for $Aligned_{id}$ (Table C.3). They are also robust to controlling for baseline 'confidence' – believed quintile minus assessed skill quintile, averaged over the two skills – and its interaction with treatment (Table C.5). We discuss the role of confidence further in Section 5.

Additional belief measures: Appendix A.5 shows that updating of CA beliefs is at least partly explained by learning about the distribution of skills in the reference population, not only learning about one's own skills; that treatment updates under- more than overconfident beliefs, matching findings in other work (e.g. Eil & Rao 2011); and that treatment reduces the individual-specific variances of skill beliefs, which we measure by asking respondents the probability they are in each quintile for each skill. The Online Supplement shows that belief updating does not differ by gender so we pool genders throughout the rest of the paper.

Beliefs about returns to directed search: The model predicts that new information about jobseekers' relative skill ranks will also change their expectations about the relative returns to searching for jobs that require different skills. To test this, we measure jobseekers' expectations about the outcomes of applying to communication- and numeracy-heavy jobs, both in the job choice task and in their planned post-workshop search. For example, for a jobseeker with numeracy CA, we construct her expected wage for numeracy-heavy jobs minus her expected wage for communication-heavy jobs.

Treatment increases expected returns to most measures of skill-directed search, although some results are imprecisely estimated. Jobseekers' expectations about job-specific wages also predict choices in the job choice task. These results are consistent with the model. The Online Supplement gives full details of the measurement and results.

Earnings: We survey jobseekers by text message one month after the workshop asking them their total earnings in the preceding week. Treatment shifts the earnings distribution to the right and mean earnings are statistically significantly higher for the treatment group (Figure C.1). But, given the tight experiment's sample size, we view this as only suggestive evidence. To provide stronger evidence about the labor market outcomes of new information about skills, we turn to the big experiment.

4 Big Experiment: Effects on Beliefs, Search & Labor Market Outcomes

The tight experiment provides strong evidence that information about skills can shift CA beliefs and search direction and suggestive evidence this can raise earnings. For clearer evidence on labor market impacts, we now analyze another experiment with 4,389 participants and a follow-up period of 3.5 months.

4.1 Sample

The big experiment took place in the same location in 2016/17, five years before the tight experiment. We recruited for both experiments in the same way: contacting active jobseekers from the database of our partner Harambee, described in Section 2.2.

Table B.2 shows both samples have similar labor market participation. In the tight/big experiment, 33/37% had done some work or income-generating activity in the past seven days and 96/97% were actively searching for work. The average jobseeker submitted 10/9 job applications and spent 14/17 hours searching for work in the last seven days.

The samples differ in some demographic characteristics. Jobseekers in the big experiment, relative to the tight experiment, were younger (24 vs 26) and more were male (38 vs 33%). A similar share had finished high school, but slightly fewer in the big experiment had university degrees (17 vs 22%) and slightly more had shorter tertiary qualifications (22 vs 15%). Only 9% had ever held a permanent or long-term job (vs 25% in the tight experiment), perhaps because they were younger. These demographic differences mirror changes in Johannesburg's demographics between 2016 and 2021 (SSA, 2016, 2022).

Our main results are robust to accounting for these differences. Table C.2 shows that the main treatment effects in both experiments are robust to reweighting the two samples to have the same distribution of baseline demographics, education, and employment. The similar pattern of treatment effects between the two samples provides some evidence that the economic mechanisms we study generalize outside a single experimental sample.

4.2 Experimental Design

We assigned 2,114 jobseekers in 27 workshops to the treatment group and 2,274 jobseekers in another 27 workshops to the control group. Treatment was administered at the workshop (day) level to avoid spillovers. Treatment groups are balanced on covariates for both the full sample and the 96% recontacted for the endline survey and endline response rates are balanced across treatment groups (Table B.3).

The big and tight experiments have deliberately similar designs. The big experiment was also run during Harambee workshops that included skill assessments, surveys, and basic job search advice. Jobseekers in treated workshops received a report about their skill assessment results (Figure B.2); jobseekers in control workshops did not. Participants knew they were participating in a study but not that treatment differed by day, so treatment-control differences in outcomes should not reflect experimenter demand effects. Appendix B provides detailed descriptions of the workshops and interventions, including a figure showing the sequence of events in the workshops.

The experiments differ in four ways. First, the big experiment was designed to measure effects on labor market outcomes, so we studied a larger sample and collected outcome data later: on average 3.5 months after treatment.[23]

Second, in the big experiment we assessed and gave treated jobseekers information on six skills, rather than two. This is more similar to real-world settings, where jobseekers direct search across a broader range of job types with different skill demands. In settings like schools or job centers, jobseekers receive and must process information about their relative rank on multiple skills. In contrast, the tight experiment used fewer skills to allow simpler definitions of CA and clearer tests for skill-directed search.

We assessed communication, numeracy, concept formation, focus, grit, and planning. (Details in Appendix A.1). Due to time constraints, we only measured jobseekers' beliefs about their level of three skills: communication, numeracy, and concept formation. The reports given to treated jobseekers showed results for all six assessments and information

[23] We used phone surveys lasting on average 25 minutes, compensating respondents with mobile phone airtime. Phone and in-person surveys in this setting deliver similar labor market data (Garlick et al., 2020).

on what traits they measure (Figure B.2).[24] The six assessments differentiate jobseekers horizontally more than vertically because assessment results are weakly correlated across skills within candidate: 13 of the 15 pairwise correlations are <0.3. Most jobseekers' reports showed substantial variation across skills: 85% had at least one top tercile but only 1.7% had all six top terciles and 58% had both top and bottom terciles.

Third, we reported assessment results in terciles, not quintiles. The coarseness of terciles relative to quintiles and using six rather than two skills on the reports means that only 23% of jobseekers have a unique skill CA, compared to 75% in the tight experiment. In the tight experiment, we excluded jobseekers without a unique skill CA from the main analysis and included them only in robustness checks because the skill-directed search measures could not be sensibly defined for them, In the big experiment, we keep these jobseekers in all analyses because this experiment uses skill-directed job search measures that can be defined for them.

Fourth, there are small logistical differences that are unlikely to affect the core mechanism activated by the treatment. In the big experiment, each workshop had more candidates. The briefings after candidates received the report were delivered in person instead of by video, but with a script to ensure consistency across sessions. The job search assistance – advice about writing CVs and cover letters – was at the beginning of the day in the big experiment and the end of the day in the tight experiment.

4.3 Specification

As in the tight experiment, we estimate treatment effects using:

$$Y_{id} = T_d \cdot \delta + \mathbf{X}_{id} \cdot \Theta + \varepsilon_i \tag{5}$$

$$Y_{id} = T_d \cdot \gamma^{misaligned} + T_d \cdot Aligned_{id} \cdot \gamma^{diff} + Aligned_{id} \cdot \gamma + \mathbf{X}_{id} \cdot \Phi + \epsilon_{id}. \tag{6}$$

Y_{id} is the outcome for jobseeker i assessed on date d and $\mathbf{X}_{id}$ is a vector of prespecified controls.[25] As in the tight experiment, we cluster standard errors at the level of treatment (workshop date) and show that our main results are robust to using a wild cluster bootstrap and to adjusting for multiple hypothesis testing, following Benjamini et al. (2006) (Table C.2). The objects of interest are the average treatment effect, δ, and the average treatment effect for jobseekers with misaligned CA beliefs at baseline, $\gamma^{misaligned}$. We in-

[24] In piloting, jobseekers easily recognized the traits linked to the soft skills measures. For example, they did not use the term "grit" but did refer to the persistence required for repetitive, boring jobs.

[25] For the big experiment, the covariates and inference methods follow the preanalysis plan used by Carranza et al. (2022) but the outcome variables and heterogeneity analysis are not prespecified. See the Online Supplement for details. $\mathbf{X}_{id}$ contains baseline assessment results, skill beliefs, education, age, gender, employment, earnings, job offers, time and risk preferences, self-esteem, baseline values for the outcome where available, and fixed effects for the blocks of days within which treatment was randomized.

terpret $\gamma^{misaligned}$ with caution because this is a noisier proxy for the "dose" of information delivered in the big than tight experiment because we observe beliefs about only three of the six skills in the big experiment.

4.4 Treatment Aligns Beliefs and Search with Assessed Comparative Advantage

Receiving information about skills substantially increases alignment between assessed skills and skill beliefs. To show this, we ask candidates in which tercile they believe they ranked for each of the communication, numeracy, and concept formation assessments. Only 19.6% of control group jobseekers believe they ranked highest in the dimension in which they actually ranked highest on the assessment, i.e., have aligned CA beliefs. Treatment increases this by 13.9pp, a 70% increase from the control group mean (Table C.7, column 1, $p < 0.001$). This is in line with the effects on beliefs observed straight after treatment in the tight experiment.[26] This shows that jobseekers' updated beliefs persist over 3.5 months and that jobseekers can process more complex information about assessment results in multiple skills. See Appendix A.3 for measurement details and Appendix C for robustness checks using other belief measures.

Treatment also increases the fraction of beliefs about skill levels that align with assessment results by 14.2pp, a 37% increase from the control mean of 38.8pp (Table C.7, column 3, $p < 0.001$). As in the tight experiment, we report all analysis pooling genders because we observe little gender heterogeneity in baseline beliefs or belief updating. See the Online Supplement for detailed gender results.

Receiving information about skills also substantially increases alignment between assessed skills and search direction. To show this, we ask jobseekers to rank the importance of communication, numeracy, and concept formation skills for the jobs for which they are applying. Only 17% of control group jobseekers are searching for jobs that most value the skill in which they have an assessed CA. Treatment increases this by 5pp (column 5, $p < 0.001$). This result on self-reported skill-directed job search, measured 3.5 months after treatment, is qualitatively consistent with the four measures of skill-directed search in the tight experiment, collected in the first month after treatment.[27]

As in the tight experiment, the search direction result is driven entirely by jobseekers likely to get more information from treatment: those with misaligned CA beliefs at

[26] The effects on aligned CA beliefs are almost identical across the tight and big experiments. But the control group mean is higher in the tight than big experiment because the big experiment uses three skills, creating more ways for believed and assessed CA to differ. In both experiments, the control group's alignment rate is only slightly higher than random guessing, accounting for the possibility of ties: 20 versus 15% in the big experiment and 48 versus 40% in the tight experiment.

[27] The big experiment effect converts to a 0.14 standard deviation increase in self-reported skill-directed search, compared to a 0.27 standard deviation effect on the skill-directed search index in the tight experiment (Table 4, column 1). The effect size on the index is larger because the index averages over four measures, producing a smaller standard deviation and hence larger standardized effect size.

baseline (Table C.7). This pattern is less clear for the effect on aligned CA beliefs.

4.5 Treatment Improves Jobseekers' Outcomes in the Labor Market

The two experiments show strong evidence that giving jobseekers their skill assessment results better aligns their beliefs and search with their assessed skill CA. Here we show that this raises their earnings, as our model suggests. We then use 13 other labor market outcomes to explore mechanisms linking skill-directed job search to higher earnings.[28]

Treatment substantially increases earnings, by 6.5 USD PPP in the week preceding the endline survey (Table 5, panel A, column 1, $p = 0.020$). This is equivalent to a 26% increase from the control group mean or to moving from the 75^{th} to 79^{th} percentile of the unconditional earnings distribution. Treatment also raises earnings by 24% of the control group mean when we restrict the estimation sample to only employed participants (panel A, column 2, $p = 0.030$).[29] The earnings effects are within the range of effect sizes reported in related research, although toward the top of that range.[30] There is enough wage variation across entry-level jobs in this context that the magnitude of the effects on earnings is plausible, at only 0.16 standard deviations of control group earnings.[31]

The earnings effects are substantially higher for jobseekers who can get more information from treatment: those whose baseline CA beliefs are misaligned. Treatment increases weekly earnings for these jobseekers by 7.5 USD PPP, or 22.8 USD PPP conditional on employment (Table 5, panel B, columns 1 & 2, $p = 0.021$ & 0.027). These effects are almost three and two times as large, respectively, as the effects for jobseekers with baseline aligned CA beliefs. However, the relatively high variance of earnings measures means that the differences are not precisely estimated ($p = 0.37$ & 0.63). Recall from Section 4.2

[28] Our companion paper, Carranza et al. (2022), also reports 8 of the 45 treatment effects from this section: only average treatment effects, not heterogeneous treatment effects; and only on earnings, current and past employment, hours, hourly wages, job offers, and written contract.

[29] We use two earnings measures with complementary goals. The first earnings measure assigns zero values to the non-employed and the second assigns missing values to the non-employed. The first measure has a point mass at zero and a continuous distribution above zero, so treatment effects on this measure may be sensitive to scaling choices (Mullahy & Norton, 2022). The second measure may be sensitive to sample selection problems when treatment affects the probability of employment, but this is not a major factor in our application. We also show in Table C.4 that treatment effects on both earnings measures are robust across a range of alternative scalings. Quantile treatment on earnings are non-negative throughout the earnings distribution, although not always statistically significant (Figure C.2).

[30] Interventions that combine learning about skills with the option to signal skills have raised earnings by 11-34% (Abebe et al., 2021b; Bassi & Nansamba, 2022; Carranza et al., 2022). Bandiera et al. (2023) find an 11% rise in earnings from a matching intervention that they attribute to effects on beliefs about labor market prospects. Using a more model-based approach, Guvenen et al. (2020) estimate that moving from the bottom to the top decile of skill match quality raises earnings by 11%. Böhm et al. (2023) show that returns to skills vary by up to 35% log points between firms, showing scope for large returns to skill match.

[31] As an additional check, we estimate the standard deviation of earnings for workers in South Africa's Labor Force Survey, from Johannesburg, with < 3 months tenure, reweighted to match our sample's demographics and education. Our treatment effect on earnings is $< 10\%$ of this estimated standard deviation.

Table 5: Treatment Effects on Labor Market Outcomes - Big Experiment

	Earnings (w)		Worked		Earnings by source (w)				Tenure	Worked (timing)			Offers (w)	Formality	
	(1) Uncond.	(2) Cond.	(3) Any	(4) Hours (w)	(5) Per hour	(6) Wage emp.	(7) Self emp.	(8) New job	(9) Since bl	(10) Since bl	(11) Month 1	(12) Month 2	(13)	(14) Wr. contract	(15) Reg. pay
Panel A: Average effects															
Treatment	6.517**	20.393**	0.009	0.603	0.295**	6.758**	-0.410	4.523**	0.024	0.030**	0.023*	0.007	0.015	0.017*	0.016
	(2.712)	(9.404)	(0.013)	(0.545)	(0.130)	(2.560)	(0.708)	(2.254)	(0.039)	(0.014)	(0.013)	(0.015)	(0.017)	(0.010)	(0.013)
Observations	4196	1280	4204	4142	4183	4174	4174	4176	4183	4205	4201	4204	4140	4184	4196
Panel B: Het. effects															
Treatment	7.467**	22.783**	0.013	0.614	0.355**	7.276**	0.171	7.251**	0.050	0.039**	0.037**	0.009	0.020	0.018	0.011
	(3.230)	(10.302)	(0.014)	(0.635)	(0.149)	(3.023)	(0.903)	(2.789)	(0.042)	(0.015)	(0.015)	(0.018)	(0.017)	(0.011)	(0.014)
Treatment × Aligned	-4.697	-10.017	-0.017	0.080	-0.304	-2.831	-2.904	-11.489**	-0.110	-0.033	-0.057	-0.015	-0.011	-0.002	0.026
comp adv belief (bl)	(5.261)	(20.776)	(0.031)	(1.208)	(0.251)	(5.015)	(2.528)	(5.636)	(0.087)	(0.032)	(0.042)	(0.034)	(0.035)	(0.022)	(0.027)
Aligned comp adv belief (bl)	2.796	14.738	-0.014	-0.942	0.244	0.054	1.262	6.363*	-0.016	0.013	0.045	0.018	0.017	-0.023	-0.017
	(3.577)	(13.183)	(0.019)	(0.785)	(0.178)	(3.596)	(1.671)	(3.546)	(0.049)	(0.017)	(0.028)	(0.024)	(0.030)	(0.015)	(0.017)
Treatment effect:	2.770	12.766	-0.004	0.694	0.051	4.445	-2.733	-4.239	-0.059	0.006	-0.019	-0.006	0.009	0.016	0.037
Aligned comp adv belief (bl)	(4.140)	(19.098)	(0.027)	(1.025)	(0.219)	(4.074)	(2.102)	(4.484)	(0.079)	(0.031)	(0.038)	(0.029)	(0.036)	(0.018)	(0.025)
Observations	4122	1248	4130	4071	4111	4101	4101	4103	4110	4131	4127	4130	4071	4111	4122
Control mean	25.424	85.826	0.309	8.794	1.267	18.463	4.184	21.095	0.582	0.671	0.465	0.437	0.182	0.120	0.219

Notes: Table 5 shows treatment effects on labor market outcomes in the big experiment. Panel A shows average treatment effects. Panel B shows heterogeneous treatment effects for jobseekers with and without aligned CA beliefs at baseline. Columns indicate different outcomes: unconditional and conditional earnings in the seven days before the endline survey (cols. 1 & 2), any work in the last seven days (col. 3), hours worked in the last seven days (col. 4), hourly earnings in the last seven days (col. 5), earnings from wage employment (col. 6), earnings from self employment (col. 7), earnings from jobs started after baseline (col. 8), tenure since baseline in current job (col. 9), worked at any point since baseline (col. 10), worked in month 1 after baseline (col. 11), worked in month 2 after baseline (col. 12), job offers in the last 30 days (col. 13), has written contract (col. 14), has regular payment frequency (col. 15). Control variables are described in footnote 25. All monetary figures are reported in 2021 USD in purchasing power parity terms. All variables marked (w) are winsorized at the 99^{th} percentile. Standard errors clustered at the treatment-day level shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

that the baseline CA beliefs are a noisier proxy in the big than tight experiment, so some imprecision in this analysis is not surprising.

How might a treatment-induced increase in skill-directed job search generate this large rise in earnings? We answer this question in three steps. First, we show that the earnings rise is largely explained by jobseekers earning higher hourly starting wages. Treatment has a modest positive but not statistically significant effect on employment at the time of the endline survey and no effect on hours worked, showing that the earnings effect is entirely explained by a 25% increase in hourly earnings (Table 5, columns 3–5). This increase in hourly earnings is explained entirely by higher wage earnings, not self-employment earnings (columns 6–7). This increase in hourly wages occurs mainly in jobs that start after treatment (column 8). This shows that effects are driven more by higher hourly wages in new jobs than by renegotiation with existing employers. Higher hourly wages in new jobs must be mechanically explained by higher starting wages and/or wage growth with tenure. But treatment does not raise tenure (column 9) and the 3.5 month period from baseline and endline provides limited scope to accumulate higher tenure. This leaves higher hourly starting wages as a more likely explanation.

Second, we provide suggestive evidence that treated jobseekers follow a different path in the labor market between baseline and endline, which might explain their higher wages. These jobseekers are 3pp more likely to work at all between treatment and endline, largely due to a 2.3pp higher probability of employment in the first month after treatment (Table 5, columns 10–12). This might arise if treatment helps jobseekers apply to jobs that better match their skills, so firms are more willing to hire them or offer higher wages, although we cannot directly test this because we do not observe the skill demands of the jobs they apply to in the big experiment. They do not all keep these jobs until the endline, as shown by the near-zero effect on tenure noted above. But they do continue with on-the-job search: the share actively searching at endline is 69–70% for all four combinations of employed $\times$ treated. On-the-job search, potentially from better-matched jobs, might allow treated jobseekers to either attract higher starting wages in new jobs or accept only higher wage offers than their current job. They do not receive more job offers (column 13), suggesting a role for higher wage offers or more selective acceptances.[32] This pattern of earlier employment, on-the-job search, and rising earnings is consistent with models of wage ladders with on-the-job search (e.g. Krolikowski 2017).[33]

[32] Treatment increases reservation wages but the effect is not precisely estimated. This imprecision may reflect the substantial challenges to measuring reservation wages in surveys, discussed by Feld et al. (2022).

[33] One extra month of work experience would not normally generate a 25% increase in earnings. However, strong job ladder effects might be possible from short-term work experience in a new job with better skill match. For example, Godlonton (2020) experimentally estimates an even higher return to short-term work experience in a new occupation for young jobseekers.

The idea that treatment allows different, better paid paths in the labor market is consistent with two other findings. Treated jobseekers are 1.7pp more likely to have formal written contracts and 1.6pp more likely to have a fixed payment frequency, such as a weekly salary (Table 5, columns 14 & 15, $p = 0.089$ & 0.182). Both effects are substantial increases relative to control group means of respectively 12 and 21%, both features are reported as desirable by jobseekers in our sample, and both features are associated with higher wages in national labor force survey data.

Third, we provide suggestive evidence that higher wages reflect *match quality* rather than *job quality* effects. Higher match quality would occur if treated and control jobseekers apply to the same job types on average but, within this pool of jobs, treated jobseekers applied more to jobs that better match their skills and hence offer better terms such as higher wages. Higher job quality would occur if some job types offer better terms to all workers conditional on their skill match, and treated jobseekers applied more to them. We find limited scope for a job quality mechanism, as there is high overlap in the types of jobs to which treated and control jobseekers apply: 96.7% of job titles on SAYouth.mobi that received an application from *any* jobseeker in the tight experiment received applications from *both* treated and control jobseekers. Furthermore, treated jobseekers do not apply to jobs with higher values of multiple job quality proxies. This pattern holds for quality proxies of jobs in the job choice task – salary or offer probability, based on expert assessments and control group beliefs – and quality proxies of jobs on SAYouth.mobi – provision of training or non-wage amenities, based on job advert text (Table D.4). These null results suggest that treatment does not induce jobseekers to apply to higher quality jobs, leaving higher match quality as a more likely explanation for higher wages.

We conclude that treatment may raise earnings through more skill-directed search that allows jobseekers to move onto better-matched job ladders. However, we acknowledge that our evidence for more skill-directed search and higher earnings is stronger than our evidence for the specific mechanisms linking skill-directed search and earnings.[34]

5 Beliefs about Skill Levels and Choice of Search Effort

We do not find evidence consistent with an alternative model, in which jobseekers are on average overconfident about the *level* of their skills, learn from treatment that they have lower skills than they thought, increase their search effort, and hence raise their earnings. We summarize this evidence and model here and provide details in Appendix D.

We find few treatment effects on search effort in either experiment, suggesting it does

[34] There may exist a feedback loop between beliefs and labor market experience: treatment $\rightarrow$ beliefs about CA $\rightarrow$ skill-directed job search $\rightarrow$ labor market outcomes $\rightarrow$ beliefs about CA. Such a feedback loop would be relevant for interpreting the magnitudes of the treatment effects on labor market outcomes, which would then reflect both direct effects and indirect effects through this feedback loop.

not explain the relationship between treatment and labor market outcomes.[35] In the tight experiment, treatment effects are small and not statistically significant on six different measures of search effort: a survey question on post-workshop planned applications, time spent on a job search task during the workshop, and four measures of job search on the SAYouth.mobi platform after the workshop (Table D.2).[36] In the big experiment, we estimate precise near-zero effects on self-reported number of applications, time spent and money spent on search (Table D.3).

Treatment does lower jobseekers' beliefs about their skill levels in both experiments (Tables D.2 and D.3, column 1). This occurs because treatment shifts jobseekers' beliefs about their skill levels toward their assessed skill levels, and more jobseekers have baseline skill beliefs above their assessed skills than below.

A generalized version of our conceptual framework shows how treatment can lower average beliefs about skills without changing search effort. In this framework, jobseekers endogenously choose their total search effort level based on their expected search outcomes. When treatment lowers believed skill level, the framework predicts two search effort responses. There is a substitution effect: jobseekers search less because the expected return to each unit of search effort is lower. There is also an "income" effect: jobseekers search more because more search is needed to achieve the same labor market outcome. The net effect on search effort can be negative, zero, or positive.

We also see no evidence that jobseekers who receive negative news about their skill levels respond by applying to jobs with worse attributes that they view as more attainable. Treated and control jobseekers apply to jobs with similar values of quality proxies in both the job choice task and on SAYouth.mobi (Table D.4).

6 Additional Mechanisms

In this section, we evaluate three other mechanisms that might account for treatment effects on labor market outcomes. These are not mutually exclusive with the directed search mechanism. We find little evidence for any of these three mechanisms. We briefly summarize the evidence here and show details in the Online Supplement.

Self-esteem: Treatment has near-zero effects on self-esteem in the big experiment, in both a text message survey 2–3 days after treatment and the endline phone survey 3.5

[35]We merely argue shifts in search effort are unlikely to explain the treatment effects on labor market outcomes in this study. We recognize search effort can play an important role in labor market outcomes in other ways, covered in the review by Mueller & Spinnewijn (2023). As we note above, our treatment does not lower reservation wages, a mechanism some papers in that review link to search beliefs and effort.

[36]Treatment provides information both about skill levels and CA. To isolate the role of information about skill levels, we repeat this analysis for a group of jobseekers who cannot get information about their CA from treatment: those with tied communication and numeracy quintiles. Effects on search effort for these jobseekers are similar to the effects for our main sample (Table D.1).

months after treatment, using questions from the Rosenberg (1965) scale. This suggests general beliefs about self-worth do not respond to new skill information and are unlikely to affect search behavior or labor market outcomes.

Human capital investment: We find that jobseekers might make skill investments in response to new information about skills, but that this behavior is unlikely to explain the labor market effects we find. In the big experiment, treatment has near-zero effects on enrollment in both formal and vocational education, limiting scope for education investment to drive the labor market effects. Education investment might, of course, change over longer time horizons or at other margins.

We find suggestive evidence in the tight experiment that jobseekers might change skill investments in response to new information about their skills. Treatment reduces willingness-to-pay (WTP) for a numeracy workbook, mostly for people who learn that they have numeracy CA, suggesting that jobseekers might prefer to invest in skills where they are relatively weak. WTP for a communication workbook is unaffected by treatment.

Skill information transmission to firms: If jobseekers share assessment results with firms during job applications, this might lead to firm-side learning about jobseekers' skills. We view this mechanism as unlikely to explain labor market outcomes. Labor market effects are driven by the jobseekers who say they did *not* use the report in applications, although we interpret this result cautiously because this analysis conditions on report use, a post-treatment outcome. We have conflicting evidence from the big experiment on how often jobseekers share results with firms: only 0.8% of treated jobseekers included their assessment report when applying to vacancies we created for another experiment but 29% of treated jobseekers self-report at endline that they ever included a copy of their assessment results with any application.

The assessment results jobseekers receive are deliberately designed not to be credible to firms. They do not show the jobseeker's name or national identity number, so firms cannot verify that the report is linked to that job applicant. They include no information about Harambee, the source of the assessments, or the value of skills. They are printed in black and white on low-quality paper. None of the 15 hiring managers we interviewed during piloting said they would view these reports as credible.[37]

7 Concluding Reflections

Our paper's main contribution is providing evidence of a relatively understudied type of job search friction: misdirected search due to imperfect information about skill CA. To study this question, we run two field experiments that randomize jobseekers' access

[37]In contrast, Carranza et al. (2022) study effects of using certificates of assessment results designed to be credible to firms. These are branded by Harambee and contain the jobseeker's name and ID number.

to information about their skills in multiple dimensions and collect unique measures of beliefs about skills, skill-directed job search, and labor market outcomes. We show that new information about jobseekers' skill CA can shift their beliefs, redirect search toward jobs that better match their CA, and improve their labor market outcomes.

We end the paper by reflecting on three topics that audiences often press us to discuss more, although we do not view as core parts of the paper. These reflections are deliberately brief, tentative, and intended to point to directions for future research.

7.1 Scale and General Equilibrium Considerations

The effects of job search interventions can depend on their scale due to interactions between jobseekers or between jobseekers and firms. We present an informal framework for thinking about scale and general equilibrium implications of skill-directed search. We provide empirical evidence about parts of this framework, which suggests that the effects of our intervention on jobseekers' labor market outcomes need not shrink with scale.

There are two obvious channels for general equilibrium effects of job search interventions. First, more skill-directed job search might **raise match quality** or **lower screening costs** for firms, as they receive better-matched applications. These mechanisms can increase aggregate labor demand according to both classic matching models (e.g. Mortenson & Pissarides 1994) and recent experiments (e.g. Algan et al. 2022). Section 4.5 provided indirect evidence our intervention raised match quality.

Second, changes in search might generate **search congestion**: more applications sent in total or to some job types, leading to lower job offer probabilities and potentially lower wage offers. This is unlikely for the type of intervention we study. Information about skill CA is inherently differentiated and shifts different jobseekers to apply for different job types, rather than shifting total applications toward one job type. We observe exactly this pattern in the tight experiment. In the job choice task, treatment spreads applications more evenly between communication and numeracy jobs. And treatment does not raise any search effort measure, including total on-platform applications to communication or numeracy jobs. See the Online Supplement for detailed results. In contrast, research that has found search congestion effects has studied interventions providing less differentiated information about high-demand sectors or locations, encouraging search effort, or providing job placement services (e.g. Altmann et al. 2022; Crépon et al. 2013; Johnston & Mas 2018; LaLive et al. 2022).[38]

We tentatively conclude that there is no clear evidence that the type of intervention we study would generate smaller effects at larger scales. However, a more conclusive answer

[38] Information about skill CA as we define it may generate search congestion if communication-heavy jobs are substantially more common than numeracy-heavy jobs or vice versa. However, we showed in Section 3.4 that the two types of jobs are roughly equally common in online job postings.

would require research designed to evaluate general equilibrium effects.

General equilibrium effects are also unlikely to affect our experiments' internal validity. The big experiment involved only 4,400 people in a city of roughly 8 million people and 2 million employed workers (SSA, 2016); they were sampled from all around the city, not one small geographic area; workshops were spread over seven months, not concentrated in time; and participants were not guided to apply to specific jobs or search in specific areas. This limits the scope for competition between jobseekers in the experiment.

7.2 Generalizability

Several features of our sample and context may be important for understanding which types of jobseekers in which types of labor markets might face imperfect information about skill CA and misdirected job search.

Sample: We deliberately study a sample of active jobseekers, a policy-relevant population for job search assistance policies. Our sample contains people who (1) are aged 18–34, (2) are from disadvantaged backgrounds, (3) have registered for job search assistance, and (4) have chosen to attend our workshops. The last two criteria screen out many non-searchers, arguably a less relevant population for job search interventions.

It's ambiguous whether jobseekers who chose not to attend the workshops might have more or less information about their skill CA than jobseekers who chose to attend, and hence whether their gains from treatment would be larger or smaller. Roughly 50% of jobseekers we invite choose to attend the workshop and attendance is weakly related to jobseeker characteristics: attendees are slightly older but are similar on other demographics, education, and recent job search effort. It's very unlikely that jobseekers make attendance decisions based directly on their expected gains from treatment because the workshop invitations don't provide information about the treatment. However, jobseekers may make attendance decisions based on factors correlated with their current information about their skill CA or gains from treatment. On the one hand, employed people might chose not to attend the workshop because their opportunity cost of time is higher and might also have lower gains from treatment if their employment is already providing information about their skill CA. On the other hand, people who did not attend because they have become discouraged from searching intensively might have higher gains from treatment if they have even less information about their skill CA.

The first two criteria are likely to, on average, screen out people with more education, more network access, and more work experience. These people might have better information about their skill CA and hence have smaller gains from treatment.

Context: Many jobseekers in this context enter the labor market with imperfect information about their CA because schools give noisy feedback on their skills (Lam et al., 2011). Similar conditions hold in many low- and middle-income countries (Gust et al., 2024; Pritchett, 2013). Learning about CA may be slow because high unemployment limits learning about skills through work experience in this context. This factor may also be relevant in the 40 countries facing youth unemployment above 25% (ILO, 2022).

More broadly, jobseekers might struggle to evaluate their skill fit with different jobs when navigating unfamiliar labor markets due to industrial displacement, migration, or structural transformation (e.g. Huckfeldt 2022; Robinson 2018). Misdirected search may be particularly important when search is costly for jobseekers or screening mismatched applicants is costly to firms (Abebe et al., 2021a; Fernando et al., 2023; Hensel et al., 2022).

7.3 How to Provide Information about Skill Comparative Advantage?

Our results show private gains for jobseekers who acquire more information about their skill CA. We show in the Online Supplement that the average treatment effect on earnings may exceed the average variable cost of the assessment operation under plausible assumptions. Firms may also gain from skill-directed job search if it raises match quality or lowers screening costs, in line with work by Abebe et al. (2021a) and Algan et al. (2022).

Given these gains, could private actors profitably supply skill assessments to jobseekers? Assessment allows substantial scale economies, particularly if run on online job search and matching platforms. Many platforms, including SAYouth.mobi, already offer skill assessments to jobseekers (LinkedIn, 2023). However, market failures are possible: prospective private skill assessors might face large fixed costs of developing assessments and building brand credibility, and jobseekers with firmly-held beliefs about their skills might not pay for assessments, even if these beliefs are inaccurate.

Governments also have avenues for providing information about skill CA to jobseekers. A strong education system can in principle provide graduates with reliable information about their CA, reducing the need for market-based provision. But the accuracy of information acquired during schooling may decrease over time, both as people age and as the labor market evolves. Government-funded job search counseling services sometimes include skill assessments but researchers have not yet studied the effects of this specific component of job search counseling (McCall et al., 2016).

Future work might examine the economics of both public- and private-sector actors providing information about CA across different types of skills.

References

Abebe, G., A. S. Caria & E. Ortiz-Ospina (2021a) "The Selection of Talent: Experimental and Structural Evidence from Ethiopia," *American Economic Review*, 111 (6), 1757–1806.

Abebe, G., S. Caria, M. Fafchamps, P. Falco, S. Franklin & S. Quinn (2021b) "Anonymity or Distance? Job Search and Labour Market Exclusion in a Growing African City.," *Review of Economic Studies*, 88 (3), 1279–1310.

Abebe, G., S. Caria, M. Fafchamps, P. Falco, S. Franklin, S. Quinn & F. Shilpi (2022) "Matching Frictions and Distorted Beliefs: Evidence from a Job Fair Experiment," Working Paper.

Abel, M., R. Burger & P. Piraino (2020) "The Value of Reference Letters: Experimental Evidence from South Africa," *American Economic Journal: Applied Economics*, 12 (3), 40–71.

Acemoglu, D. & D. Autor (2011) "Skills, Tasks and Technologies: Implications for Employment and Earnings," in Card, D. & O. Ashenfelter eds. *Handbook of Labor Economics*, 1043–1171: Elsevier.

Adams-Prassl, A. & A. Andrew (2023) "Revealed Beliefs and the Marriage Market Return to Education," Working Paper.

Adams-Prassl, A., T. Boneva, M. Golin & C. Rauh (2023) "Perceived Returns to Job Search," *Labour Economics*, 80, 102307.

Alfonsi, L., O. Bandiera, V. Bassi, R. Burgess, I. Rasul, M. Sulaiman & A. Vitali (2020) "Tackling youth unemployment: Evidence from a labor market experiment in Uganda," *Econometrica*, 88 (6), 2369–2414.

Alfonsi, L., M. Namubiru & S. Spaziani (2022) "Meet Your Future: Experimental Evidence on the Labor Market Effects of Mentors," Working Paper.

Algan, Y., B. Crépon & D. Glover (2022) "Are Active Labor Market Policies Directed at Firms Effective? Evidence from a Randomized Evaluation with Local Employment Agencies," Working Paper.

Altmann, S., A. Falk, S. Jäger & F. Zimmermann (2018) "Learning About Job Search: A Field Experiment With Job Seekers in Germany," *Journal of Public Economics*, 164, 33–49.

Altmann, S., A. Glenny, R. Mahlstedt & A. Sebald (2022) "The Direct and Indirect Effects of Online Job Search Advice," IZA Discussion Paper 15830.

Altonji, J., E. Blom & C. Meghir (2012) "Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers,," *Annual Review of Economics*, 4 (1), 185–223.

Altonji, J., P. Arcidiacono & A. Maurel (2016) "Chapter 7 - The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects," in Hanushek, E. A., S. Machin & L. Woessmann eds. *Handbook of the Economics of Education*, 5, 305–396: Elsevier.

Altonji, J. & C. Pierret (2001) "Employer Learning and Statistical Discrimination," *The Quarterly Journal of Economics*, 116 (1), 313–335.

Anderson, M. L. (2008) "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 103 (484), 1481–1495.

Arcidiacono, P., E. Aucejo, A. Maurel & T. Ransom (2016) "College Attrition and the Dynamics of Information Revelation," NBER working paper 22325.

Arcidiacono, P., P. Bayer & A. Hizmo (2010) "Beyond Signaling and Human Capital: Education and the Revelation of Ability," *American Economic Journal: Applied Economics*, 2 (4), 76–104.

Aucejo, E. & J. James (2021) "The Path to College Education: The Role of Math and Verbal Skills," *Journal of Political Economy*, 129 (10), 2905–2946.

Baley, I., A. Figueiredo & R. Ulbricht (2022) "Mismatch Cycles," *Journal of Political Economy*, 130 (11).

Bandiera, O., V. Bassi, R. Burgess, I. Rasul, M. Sulaiman & A. Vitali (2023) "The Search for Good Jobs: Evidence from a Six-Year Field Experiment in Uganda," *Journal of Labor Economics*, forthcoming.

Bandiera, O., N. Parekh, B. Petrongolo & M. Rao (2022) "Men Are from Mars, and Women Too: A Bayesian Meta-analysis of Overconfidence Experiments," *Economica*, 89 (S1), 38–70.

Banerjee, A. & S. Sequeira (2023) "Learning by Searching: Spatial Mismatches and Imperfect Information in Southern Labor Markets," *Journal of Development Economics*, 164, 103111.

Bassi, V. & A. Nansamba (2022) "Screening and Signalling Non-Cognitive Skills: Experimental Evidence from Uganda," *The Economic Journal*, 132 (642), 471–511.

Batra, H., A. Michaud & S. Mongey (2023) "Online Job Posts Contain Very Little Wage Information," NBER working paper 31984.

Bazzi, S., L. Cameron, S. G. Schaner & F. Witoelar (2021) "Information, Intermediaries, and International Migration," NBER Working Paper 29588.

Beam, E. (2016) "Do Job Fairs Matter? Experimental Evidence from the Philippines," *Journal of Development Economics*, 120, 32–40.

Behaghel, L., S. Dromundo Mokrani, M. Gurgand, Y. Hazard & T. Zuber (2022) "Encouraging and Directing Job Search: Direct and Spillover Effects in a Large Scale Experiment," Banque de France working paper 900.

Belot, M., P. Kircher & P. Muller (2019) "Providing Advice to Jobseekers at Low Cost: An Experimental Study on Online Advice," *Review of Economic Studies*, 86 (4), 1411–1447.

——— (2022) "Do the Long-term Unemployed Benefit from Automated Occupational Advice during Online Job Search?", IZA Discussion Paper 15452.

Benjamini, Y., A. Krieger & D. Yekutieli (2006) "Adaptive Linear Step-Up Procedures That Control the False Discovery Rate," *Biometrika*, 93 (3), 491–507.

Berry, J., R. Dizon-Ross & M. Jagnani (2022) "Not Playing Favorites: Parents Equalize Their Children's Opportunities," NBER Working Paper 26732.

Bobba, M. & V. Frisancho (2022) "Self-Perceptions About Academic achievement: Evidence From Mexico City," *Journal of Econometrics*, 231 (1), 58–73.

Böhm, M., K. Esmkhani & G. Gallipoli (2023) "Firm Heterogeneity in Skill Returns," IZA discussion paper 16644.

Boudreau, L., R. Heath & T. McCormick (2023) "Migrants, Experience, and Working Conditions in Bangladeshi Garment Factories," *Journal of Economic Behavior and Organization*, forthcoming.

Caria, S., K. Orkin, A. Andrew, R. Garlick, R. Heath & N. Singh (2024) "Barriers to Search and Hiring in Urban Labour Markets," *VoxDevLit*, 10 (1).

Carranza, E., R. Garlick, K. Orkin & N. Rankin (2022) "Job Search and Hiring with Limited Information about Workseekers' Skills," *American Economic Review*, 112 (11), 3547–83.

Chakravorty, B., A. Y. Bhatiya, C. Imbert, M. Lohnert, P. Panda & R. Rathelot (2023) "Impact of the COVID-19 Crisis on India's Rural Youth: Evidence From a Panel Survey and an Experiment," *World Development*, 168, 106242.

Clark, D. & P. Martorell (2014) "The Signaling Value of a High School Diploma," *Journal of Political Economy*, 122 (2), 282–318.

Conlon, J. J., L. Pilossoph, M. Wiswall & B. Zafar (2018) "Labor Market Search With Imperfect Information and Learning," NBER Working Paper 24988.

Cortés, P., J. Pan, L. Pilossoph & B. Zafar (2023) "Gender Differences in Job Search and the Earnings Gap: Evidence from the Field and Lab," *The Quarterly Journal of Economics*, forthcoming.

Crépon, B., E. Duflo, M. Gurgand, R. Rathelot & P. Zamora (2013) "Do Labor Market Policies Have Displacement Effects? Evidence from a Clustered Randomized Experiment," *The Quarterly Journal of Economics*, 128 (2), 531–580.

Delaney, J. & P. Devereux (2021) "High School Rank in Math and English and the Gender Gap in STEM," *Labour Economics*, 69, 1019–1069.

Diamond, A. (2013) "Executive Functions," *Annual Review of Psychology*, 64, 135–168.

Dizon-Ross, R. (2019) "Parents' Beliefs about Their Children's Academic Ability: Implications for Educational Investments," *American Economic Review*, 109 (8), 2728–2765.

Duckworth, A. (2016) *Grit: The Power of Passion and Perseverance*, New York, NY, US: Scribner/Simon & Schuster.

Eil, D. & J. M. Rao (2011) "The Good News-Bad News Effect: Asymmetric Processing of Objective Information About Yourself," *American Economic Journal: Microeconomics*, 3 (2), 114–38.

Falk, A., D. Huffman & U. Sunde (2006) "Self-Confidence and Search," IZA Discussion Paper 2525.

Feld, B., A. Osman & A. Nagy (2022) "What Do Jobseekers Want? Comparing Methods to Estimate Reservation Wages and the Value of Job Attributes," *Journal of Development Economics*, 159.

Fernando, N., N. Singh & G. Tourek (2023) "Hiring Frictions and the Promise of Online Job Portals: Evidence from India," *American Economic Review: Insights*, forthcoming.

Franco, C. (2019) "How does relative performance feedback affect beliefs and academic decisions?", Working paper.

Fredriksson, P., L. Hensvik & O. N. Skans (2018) "Mismatch of Talent: Evidence on Match Quality, Entry Wages, and Job Mobility," *American Economic Review*, 108 (11), 3303–3338.

Garlick, R., K. Orkin & S. Quinn (2020) "Call Me Maybe: Experimental Evidence on Frequency and Medium Effects in Microenterprise Surveys," *The World Bank Economic Review*, 34 (2), 418–443.

Gneezy, U., A. Rustichini & A. Vostroknutov (2010) "Experience and Insight in The Race Game," *Journal of Economic Behavior and Organization*, 75 (2), 144–155.

Godlonton, S. (2020) "Employment Exposure: Employment and Wage Effects in Urban Malawi," *Economic Development and Cultural Change*, 68 (2).

Groh, M., D. McKenzie, N. Shammout & T. Viswanath (2015) "Testing the Importance of Search Frictions and Matching through a Randomized Experiment in Jordan," *IZA Journal of Labor Economics*, 4 (7).

Gust, S., E. Hanushek & L. Woessmann (2024) "Global Universal Basic Skills: Current Deficits and Implications for World Development," *Journal of Development Economics*, 166.

Guvenen, F., B. Kuruscu, S. Tanaka & D. Wiczer (2020) "Multidimensional Skill Mismatch," *American Economic Journal: Macroeconomics*, 12 (1), 210–244.

Haaland, I., C. Roth & J. Wohlfart (2023) "Designing Information Provision Experiments," *Journal of Economic Literature*, 61 (1), 3–40.

Hardy, M. & J. McCasland (2023) "Are Small Firms Labor Constrained? Experimental Evidence from Ghana," *American Economic Journal: Applied Economics*, 15 (2), 253–284.

Harrison, G. & J. List (2004) "Field Experiments," *Journal of Economic Literature*, 42, 1009–1055.

He, Q. & P. Kircher (2023) "Updating about Yourself by Learning about the Market: The Dynamics of Beliefs and Expectations in Job Search," NBER working paper 31940.

Heath, R. (2018) "Why Do Firms Hire Using Referrals? Evidence from Bangladeshi Garment Factories," *Journal of Political Economy*, 126 (4), 1691–1746.

Hensel, L., T. G. Tekleselassie & M. Witte (2022) "Formalized Employee Search and Labor Demand," Working Paper.

Hoffman, M. & S. Burks (2020) "Worker Overconfidence: Field Evidence and Implications for Employee Turnover and Firm Profits," *Quantitative Economics*, 11, 315–348.

Huckfeldt, C. (2022) "Understanding the Scarring Effect of Recessions," *American Economic Review*, 112 (4), 1273–1310.

Huffman, D., C. Raymond & J. Shvets (2022) "Persistent Overconfidence and Biased Memory: Evidence from Managers," *American Economic Review*, 112 (10), 3141–75.

ILO (2022) "International Labour Organization's Global Employment Trends for Youth 2022," Geneva.

Ioannides, Y. M. & L. D. Loury (2004) "Job Information Networks, Neighborhood Effects, and Inequality," *Journal of Economic Literature*, 42 (4), 1056–1093.

Jäger, S., C. Roth, N. Roussille & B. Schoefer (2023) "Worker Beliefs About Outside Options," Working Paper.

Jepsen, C., P. Mueser & K. Troske (2016) "Labor Market Returns to the GED Using Regression Discontinuity Analysis," *Journal of Political Economy*, 124 (3), 621–649.

Johnston, A. C. & A. Mas (2018) "Potential Unemployment Insurance Duration and Labor Supply: The Individual and Market-Level Response to a Benefit Cut," *Journal of Political Economy*, 126, 2480–2522.

Jones, S. & R. Santos (2022) "Can Information Correct Optimistic Wage Expectations? Evidence From Mozambican Job-seekers," *Journal of Development Economics*, 159, 102–987.

Kahn, L. & F. Lange (2014) "Employer Learning, Productivity, and the Earnings Distribution: Evidence from Performance Measures," *Review of Economic Studies*, 84 (1), 1575–1613.

Kelley, E. M., C. Ksoll & J. Magruder (2023) "How do Digital Platforms Affect Employment and Job Search? Evidence from India ," *Journal of Development Economics*, forthcoming.

Kerr, A. (2017) "Tax(i)ing The Poor? Commuting Costs in South African Cities," *South African Journal of Economics*, 85 (3), 321–340.

Kroft, K. & D. Pope (2014) "Does Online Search Crowd Out Traditional Search and Improve Matching Efficiency? Evidence from Craigslist,," *Journal of Labor Economics*, 32 (2), 259–303.

Krolikowski, P. (2017) "Job Ladders and Earnings of Displaced Workers," *American Economic Journal: Macroeconomics*, 9 (2), 1–31.

LaLive, R., C. Landais & J. Zweimuller (2022) "Market Externalities of Large Unemployment Insurance Extension Programs," *American Economic Review*, 105, 3564–3596.

Lam, D., C. Ardington & M. Leibbrandt (2011) "Schooling as a Lottery: Racial Differences in School Advancement in Urban South Africa," *Journal of Development Economics*, 95 (2), 121–136.

Le Barbanchon, T., L. Hensvik & R. Rathelot (2023) "How Can AI Improve Search and Matching? Evidence from 59 Million Personalized Job Recommendations," Working paper.

Lentz, R., J. Maibom & E. Moen (2022) "Competitive or Random Search?", Working Paper.

LinkedIn (2023) "LinkedIn Skill Assessments," https://www.linkedin.com/help/linkedin/answer/a507734, Date accessed: 2023/04/21.

Lise, J. & F. Postel-Vinay (2020) "Multidimensional Skills, Sorting, and Human Capital Accumulation," *American Economic Review*, 110 (8), 2328–76.

MacLeod, W. B., E. Riehl, J. Saavedra & M. Urquiola (2017) "The Big Sort: College Reputation and Labor Market Outcomes," *American Economic Journal: Applied Economics*, 9 (3), 223–261.

Malmendier, U. & G. Tate (2015) "Behavioral CEOs: On the Role of Managerial Overconfidence," *Journal of Economic Perspectives*, 29 (4), 37–60.

Mas, A. & A. Pallais (2017) "Valuing Alternative Work Arrangements," *American Economic Review*, 107 (12), 3722–3759.

McCall, B., J. Smith & C. Wunsch (2016) "Government-sponsored Vocational Education for Adults," in Hanushek, E., S. Machin & L. Woessmann eds. *Handbook of the Economics of Education*, 5, 479–652: Elsevier.

Moore, D. A. & P. J. Healy (2008) "The Trouble With Overconfidence.," *Psychological Review*, 115 (2), 502.

Mortenson, D. & C. Pissarides (1994) "Job Creation and Job Destruction in the Theory of Unemployment," *Review of Economic Studies*, 61, 397–415.

Mueller, A. I. & J. Spinnewijn (2023) "Expectations Data, Labor Market, and Job Search," in Bachmann, R., G. Topa & W. van der Klaauw eds. *Handbook of Economic Expectations*, Chap. 22, 677–713: Academic Press.

Mueller, A. I., J. Spinnewijn & G. Topa (2021) "Job Seekers' Perceptions and Employment Prospects: Heterogeneity, Duration Dependence and Bias," *American Economic Review*, 111, 324–363.

Mullahy, J. & E. C. Norton (2022) "Why Transform Y? A Critical Assessment of Dependent-Variable Transformations in Regression Models for Skewed and Sometimes-Zero Outcomes," NBER Working Paper 30735.

Nielsen, E. (2023) "How Sensitive are Standard Statistics to the Choice of Scale?", Working Paper.

Pallais, A. (2014) "Inefficient Hiring in Entry-Level Labor Markets," *American Economic Review*, 104 (11), 3565–3599.

Papageorgiou, T. (2014) "Learning your comparative advantages," *Review of Economic Studies*, 81 (3), 1263–1295.

Pillay, A. L. (2020) "Prioritising Career Guidance and Development Services in Post-Apartheid South Africa," *African Journal of Career Development*, 2 (1), 1–5.

Posner, M. & G. DiGirolamo (1998) "Executive Attention: Conflict, Target Detection, and Cognitive Control," in Parasuraman, R. ed. *The Attentive Brain*, 401–423: MIT Press.

Pritchett, L. (2013) *The Rebirth of Education: Schooling Ain't Learning*, Washington, DC: Center for Global Development.

Raven, J. & J. Raven (2003) "Raven Progressive Matrices," in McCallum, R. ed. *Handbook of Nonverbal Assessment*, 223–237, Boston: Springer.

Robinson, C. (2018) "Occupational Mobility, Occupation Distance, and Specific Human Capital," *Journal of Human Resources*, 53 (2), 513–551.

Rosenberg, M. (1965) *Society and the Adolescent Self-Image*: Princeton University Press.

Roy, A. D. (1951) "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3 (2), 135–146.

Saltiel, F. (2022) "Multidimensional Skills and Gender Differences in STEM Majors," *Economic Journal*, 133 (651), 1217–1246.

Sanders, C. & C. Taber (2012) "Life-Cycle Wage Growth and Heterogeneous Human Capital," *Annual Review of Economics*, 4 (1), 399–425.

Santos-Pinto, L. & L. E. de la Rosa (2020) "Overconfidence in Labor Markets," *Handbook of Labor, Human Resources and Population Economics*, 1–42.

Schoer, V., M. Ntuli, N. Rankin, C. Sebastiao & K. Hunt (2010) "A Blurred Signal? The Usefulness of National Senior Certificate (NSC) Mathematics Marks as Predictors of Academic Performance at University Level," *Perspectives in Education*, 28 (2), 9–18.

Sockin, J. & A. Sojourner (2023) "What's the Inside Scoop? Challenges in the Supply and Demand for Information about Job Attributes," *Journal of Labor Economics*, forthcoming.

Spinnewijn, J. (2015) "Unemployed but Optimistic: Optimal Insurance Design With Biased Beliefs," *Journal of the European Economic Association*, 13 (1), 130–167.

SSA (2016) *Quarterly Labor Force Survey Quarter 4 2016*, Pretoria: Statistics South Africa.

——— (2022) *Quarterly Labor Force Survey Quarter 3 2022*, Pretoria: Statistics South Africa.

Stinebrickner, R. & T. Stinebrickner (2014) "A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout," *Review of Economic Studies*, 81, 426–472.

Stroop, J. R. (1935) "Studies of Interference in Serial Verbal Reactions," *Journal of Experimental Psychology*, 18 (3), 643–662.

Subramanian, N. (2022) "Workplace Attributes and Women's Labor Supply Decisions: Evidence from a Randomized Experiment," Working Paper.

Taylor, T. (1994) "A Review of Three Approaches to Cognitive Assessment, and a Proposed Integrated Approach Based on a Unifying Theoretical Framework," *South African Journal of Psychology*, 24.

Vyborny, K. ⓡ R. Garlick ⓡ N. Subramanian ⓡ E. Field (2023) "Why Don't Jobseekers Search More? Barriers and Returns to Search on a Job Matching Platform," Working Paper.

Wheeler, L., R. Garlick, E. Johnson, P. Shaw & M. Gargano (2022) "LinkedIn(to) Job Opportunities: Experimental Evidence from Job Readiness Training," *American Economic Journal: Applied Economics*, 14 (2), 101–25.

Wiswall, M. & B. Zafar (2015) "Determinants of College Major Choice: Identification Using an Information Experiment," *Review of Economic Studies*, 82 (2), 791–824.

Wu, L. (2021) "Partially Directed Search in the Labor Market," Working Paper.

Jobseekers' Beliefs about Comparative Advantage and (Mis)Directed Search: Online Appendices Not for Print Publication

Appendix A describes our skills and beliefs measurements, how firms use skills in hiring, and additional treatment effects on skill beliefs. Appendix B describes the experimental protocols. Appendix C shows robustness checks. Appendix D shows our search effort model and results, building on Section 5. The Online Supplement contains further information that is less central to the paper but is included for interested readers: descriptive statistics on the skill assessment results; evidence that treatment increases expected returns to skill-directed job search, building on Section 3.5; results relating to mechanisms discussed in Section 6; evidence that there is limited gender heterogeneity in skill beliefs or treatment effects on skill beliefs; an explanation of how our final analysis relates to our pre-analysis plan; and details of the cost-benefit calculations discussed in Section 7.3.

A Skills and Skill Beliefs: Measurement and Uses

A.1 Skill Measurement

Tight Experiment: We measured three skills. The *numeracy* assessment is developed and used by a South African retail chain to screen applicants for cashier roles. The *communication* assessment is developed by a local adult education provider to measure jobseekers' English listening and reading comprehension at a high school level. The *concept formation* assessment is adapted from Ravens' matrices to measure jobseekers' fluid intelligence: their conceptual reasoning and the rate at which they learn (Raven & Raven, 2003). The Online Supplement shows that assessment results have bell-shaped distributions without floor or ceiling effects and that all three measures have pairwise correlations of 0.31-0.38.

Big Experiment: We measured six skills. The *communication* and *concept formation* assessments are identical to the tight experiment and the *numeracy* assessment uses mostly the same questions. The *focus* measure is a computerized, color-based Stroop task (Stroop, 1935). It evaluates inhibitory control, controlling one's attention and guiding thought and action to achieve a goal (Diamond, 2013; Posner & DiGirolamo, 1998). We assess jobseekers' *grit* using the self-reported 8-item scale from Duckworth (2016). It rates jobseekers' willingness to work on difficult tasks and persevere to achieve long-term goals. The *planning* measure is a Hit 15 task (Gneezy et al., 2010). It measures how jobseekers are able to search for relevant information and anticipate the consequences of actions. In the Online Supplement, we show that scores are have pairwise correlations of 0.05–0.44.

The modest positive correlations between scores on different assessments in both experiments show that the assessments horizontally differentiate candidates based on their relative skills rather than only vertically differentiating them on a single skill dimension.

A.2 Firms' Use and Observation of Skills

In the Online Supplement we provide detailed evidence that the skills we measured are important in this labor market. We summarize this evidence briefly here. First, the communication, concept formation, and numeracy assessments have been used to screen jobseekers by our partner, the Harambee Youth Employment Accelerator. By 2016, Harambee had been contracted by South African firms to screen roughly 160,000 prospective workers using these assessments. Second, we used an incentivized choice experiment to show that firms vary in their valuation of communication and numeracy skills and value both highly relative to some forms of education. This experiment involved hiring managers from 67 firms choosing between between job applicant profiles with different skill and education levels, with their choices used to recommend job applicants to their firms on an online matching platform. Third, we conducted a measurement exercise to show that firms can partially observe assessed skills and value applicants whose skill profile matches their job requirements. This exercise was embedded into the hiring process of one firm, whose hiring staff rated the communication and numeracy skills of applicants to a mix of communication- and numeracy-heavy jobs and shortlisted them for interviews.

A.3 Skill Beliefs Measurement

We measure skill beliefs at the skill×individual level in terms of quintiles (tight experiment) or terciles (big experiment). In the tight experiment, we measure beliefs about numeracy and communication skills. In the big experiment, we also measure beliefs about concept formation skills. We measure two types of skill beliefs: beliefs about *assessment results* and beliefs about *general* skills. Table A.1 contains the exact wording of each question. The strong positive relationship between the two belief measures that we show below suggests that jobseekers view the assessments as relevant to their general skills.

Before eliciting beliefs, we define skills as follows: "Numeracy means working with numbers. It includes using addition, subtraction, multiplication, and division to solve real problems involving money, time, and quantities. For example, if a box holds 18 cans of tuna, can you calculate how many cans of tuna there are in 9 boxes? Communication means reading, writing, and listening in English. It includes understanding your coworkers and customers when they explain problems they have and explaining how to solve these problems. These are not skills about how to treat other people, just English skills."

Beliefs about general skills: We measure each jobseeker's beliefs about their general skills as their beliefs about their skills relative to the skills of the reference group in a specific domain, abstracting from specific assessment results. We see these skill beliefs as being *most relevant for search decisions* because they capture general, labor-market-relevant skills that rather than performance on our specific assessments. Thus, we use these beliefs to define our preferred outcome measure of aligned CA beliefs.

In the tight experiment, we measure beliefs about general skills twice: before assessments and after treatment. (Figure 1 shows the experiment timeline.) We use general skill beliefs for all descriptive statistics and to define our main belief outcome measures. In the big experiment, we only measure beliefs about general skills once: in the endline survey, for a random subsample of participants, only to use as a robustness check. Using this measure, treatment raises aligned CA beliefs by 7.6 pp ($p = 0.003$) and fraction aligned beliefs by 10.8pp ($p < 0.001$). These are only slightly smaller than the effects on beliefs about assessment results, our main belief outcomes in the big experiment (Table C.7).

Beliefs about assessment results: We also measure jobseekers' beliefs about their quintiles in the assessments they completed in both experiments. We follow Haaland et al. (2023) in viewing these beliefs as a *proxy for the information content of the intervention* for each individual. Jobseekers whose baseline beliefs differ from their assessment results will learn that their actual performance differed from their beliefs and may respond more to treatment. Hence, we estimate heterogeneous treatment effects in both the tight and big experiments using a dummy variable for jobseekers whose baseline beliefs matched their skill CA in the assessments. We also use a continuous version for robustness checks.

In the tight experiment, we asked about beliefs about assessment results once: after assessments but before treatment (see Figure 1). We use these measures for all heterogeneity analysis by baseline beliefs. We ask the same question again right after the treatment administration for the treatment group only to check whether they understood the results on the report. We report this understanding check on page 17. Our results are similar when we estimate heterogeneous treatment effects using beliefs about general skills or beliefs about assessment results because the two measures are highly correlated (ρ=0.68).[39]

In the big experiment, we asked beliefs about assessment results twice: after the assessments but before treatment, and at endline roughly 3.5 months later. We use these beliefs both as outcomes and for heterogeneity analysis for the big experiment because we only measured beliefs about general skills for a subsample of jobseekers. However, treatment effects on beliefs about general skills are similar, as we note above.

A.4 Skill Belief Descriptive Statistics

Skill beliefs, assessment results, and high school graduation exam results: Table A.2 shows that prior beliefs about skills are strongly correlated with assessment results for numeracy but not for communication in the tight experiment. Table A.3 shows relationships between jobseekers' self-reported scores in English and mathematics in the high

[39] Similarly, regressing domain-specific general skill beliefs on beliefs about assessment results produces coefficients of 0.39–0.52, with or without controls for assessment results, demographics, and education.

Table A.1: Measurement of Beliefs about Comparative Advantage

Description	Survey question
Panel A: Tight experiment	
Beliefs about general skills, pre-treatment (most likely quintile)	Think about 100 people who are jobseekers from Johannesburg aged 18-34 with a matric from a township or rural school. Imagine that we rank everyone according to their [numeracy/communication] skills, from lowest to highest. We create five equal size groups. The first group are the 20 people with the strongest [numeracy/communication] skills. The second group are the 20 people with the next best skills – they are less good than the top 20, but better than the other 60 people. The fifth group are the 20 people with less strong numeracy skills than the other 80. Out of these five groups we just talked about, what group do you think you are most likely to be in based on your [numeracy/communication] skills?
Beliefs about general skills belief, post treatment (most likely quintile)	Think about 100 people who are jobseekers from Johannesburg aged 18-34 with a matric from a township or rural school. Imagine that we rank everyone according to their [numeracy/communication] skills, from lowest to highest. This ranking is based on overall [numeracy/communication] skills, not only the numeracy skills that were tested in the Numeracy assessment you just took. We create five equal size groups. The first group are the 20 people with the strongest [numeracy/communication] skills. The second group are the 20 people with the next best skills – they are less good than the top 20, but better than the other 60 people. The fifth group are the 20 people with less strong [numeracy/communication] skills than the other 80. Out of these five groups we just talked about, what group do you think you are most likely to be in based on your [numeracy/communication] skills?
Beliefs about assessment results, pre- and post-treatment (most likely quintile)	Think about 100 people who are jobseekers from Johannesburg aged 18-34 with a matric from a township or rural school. Imagine that we rank everyone according to their results on the [numeracy/communication] assessment. We create five equal size groups. The first group are the 20 people with the highest numeracy results. The second group are the 20 people with the next best results – they are less good than the top 20, but better than the other 60 people. The fifth group are the 20 people with lower strong numeracy results than the other 80. Out of these five groups we just talked about, what group do you think you are most likely to be in based on your [numeracy/communication] assessment result?
Panel B: Big experiment	
Beliefs about general skills, post-treatment (most likely tercile)	Remember that people who come to Harambee are from Johannesburg, are aged 18-34 and have a matric from a township or rural school. So that should be the group you're picturing. If we ranked candidates by their [numeracy/communication/concept formation] skills, do you think you are in the top third, middle third or bottom third of Harambee candidates?
Beliefs about assessment results, pre-treatment (most likely tercile)	Now think about all the people who are in the room with you. They are all jobseekers from Johannesburg aged 18-34 with a matric from a township or rural school and have done the Harambee assessments. Imagine we line everyone up according to what score they got, from lowest to highest. Then we divide the group into three. The lower third are the people who got the lowest scores. The top third are the people who got the highest scores. The middle third are the rest of the people. Would you be in the top third, middle third or bottom third of people on the [numeracy/communication/concept formation] test?
Beliefs about assessment results, post-treatment (most likely tercile)	Do you remember the assessments you took at Harambee during Phases 1 and 2? [wait for yes]. Now I want you to imagine other Harambee candidates who have also taken these assessments. Remember that people who come to Harambee are from Johannesburg, are aged 18-34 and have a matric from a township or rural school. So that should be the group you're picturing. Imagine we look at everyone's assessment scores, and we make three groups: One group for people with the lowest scores, one group for people with the highest scores, and one group for people in the middle. Each group contains one third of the people who took the assessment. Keep this scenario in your mind, and answer the following questions. Remember that this will not have any impact on your progress with Harambee. These answers are only for research purposes and will be kept confidential. Off the top of your head, do you think you are in the top third, middle third or bottom third of people on the [numeracy/communication/concept formation] test?

Notes: **Table A.1 displays the exact wording of our questions to jobseekers about their beliefs about their skills. Note that we construct beliefs about skill CA from the beliefs about terciles/quintiles of different skills.**

school graduation exam (matric) and their assessment results and beliefs. The matric results correlate with our assessments and the score differences between the respective subjects on those exams positively correlate with the measured CA in the tight experiment, although the latter relationship is imprecisely estimated (columns 1-3). The high school graduation exam results also predict jobseekers' beliefs (columns 4-7).

Skill beliefs, search activities, and expected search outcomes: Table A.4 shows that jobseekers' beliefs about their CA correlate strongly with search direction in the control group of the big experiment, with or without controls for demographics, education, work experience, and measured CA. In contrast, the same table shows that measured CA is weakly associated with search direction, with or without controls for demographics, education, work experience, and believed CA.

Table A.5 shows that jobseekers with higher beliefs about their skills expect shorter search durations and higher wages. This might reflect unobserved jobseeker heterogeneity, so it does not necessarily show a causal relationship. As an additional check, we asked jobseekers in the big experiment their expected search duration and earnings conditional on finding a job and then their expectations for another jobseeker who had better numeracy skills but was otherwise identical to themselves. Jobseekers expect that the other hypothetical jobseeker will search for 0.74 fewer months than themselves (24% of the mean, $p = 0.02$) and earn 118 USD more (13% of the mean, $p < 0.001$). We did not ask this question about communication skills due to time constraints.

Dynamics of skill beliefs: Beliefs persist from baseline to endline in the big experiment's control group, even for active jobseekers, suggesting that jobseekers learn at best slowly about their skills via search (Table A.6).

Other correlates of skill beliefs: In the tight experiment baseline, having aligned CA beliefs is associated with assessment results but unrelated to demographics, education, and labor market exposure (Table A.7).

Table A.2: Association between Assessed and Believed Skill Quintiles - Tight Experiment

	Skill quintile beliefs	
	Numeracy (1)	Communication (2)
Numeracy quintile	0.188***	-0.011
	(0.049)	(0.034)
Communication quintile	-0.061	0.018
	(0.040)	(0.032)
Dep var. mean	2.367	3.259
Observations	278	278

Notes: **Table A.2 shows that jobseekers' beliefs about their score quintiles correlate with their assessment results for numeracy but not communication.** No control variables are included. Heteroskedasticity-robust standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.3: Association between High School Graduation Exam Results and Assessed and Believed Skills - Tight Experiment

	Asessed skills			Beliefs about skills			
	Skill quintile		Comp. adv.	Skill quintile		Comp. adv.	
	Num. (1)	Com. (2)	Num. (3)	Num. (4)	Com. (5)	Num. (6)	Com. (7)
Matric: Math score	1.451***	0.594		1.782***	-0.360		
	(0.506)	(0.609)		(0.399)	(0.261)		
Matric: English score	0.219	1.219**		-0.355	1.016***		
	(0.400)	(0.477)		(0.335)	(0.223)		
Matric: Δ Math score - English score			0.192			0.308***	-0.487***
			(0.146)			(0.108)	(0.139)
Dep var. mean	1.540	2.173	0.378	2.367	3.259	0.137	0.579
Observations	263	263	263	263	263	263	263

Notes: **Table A.3 shows that self-reported scores in English and mathematics in the high school graduation exam (matric) correlate positively with jobseekers' assessment results and their baseline beliefs about skills.** Cols. 1, 2, 4, and 5 show the relationship between math and English matric scores and assessed skill quintiles (cols. 1 and 2) and beliefs about skill quintiles (cols. 4 and 5). Matric scores are rescaled to range from 0 to 1. So, for example, the coefficient in column 4 row 1 shows that moving from the lowest to highest possible matric grade in math is associated with a belief that numeracy is 1.78 quintiles higher. No further control variables are included. The sample size is 263 because 15 jobseekers cannot recall their exam results. Heteroskedasticity-robust standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.4: Association between Skill Beliefs and Search Direction - Big Experiment Control Group

	Target numeracy jobs			Target communication jobs		
	(1)	(2)	(3)	(4)	(5)	(6)
Believed numeracy CA	0.097***		0.091***	-0.086***		-0.084***
	(0.018)		(0.018)	(0.021)		(0.022)
Believed communication CA	-0.074***		-0.067***	0.082***		0.088***
	(0.021)		(0.020)	(0.024)		(0.023)
Numeracy CA		0.029	0.021		-0.018	-0.012
		(0.018)	(0.018)		(0.022)	(0.022)
Communication CA		-0.068***	-0.060***		0.054**	0.045**
		(0.019)	(0.019)		(0.023)	(0.023)
Control mean	0.222	0.222	0.222	0.471	0.471	0.471
Observations	2179	2179	2179	2179	2179	2179

Notes: **Table A.4 shows that jobseekers' beliefs about CA correlate positively with skill-directed search in the control group of the big experiment.** "CA' stands for comparative advantage. Dependent variables are dummies indicating that jobseekers rate numeracy (columns 1-3) or communication (columns 4-6) as the most important skill for the jobs that they applied to in the last 30 days, measured at endline. Independent variables are dummies assessed and believed skill CA. All specifications include controls for age, gender, having worked in a wage job, as well as dummies for three education categories. Heteroskedasticity-robust standard errors are in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.5: Association between Skill Beliefs and Beliefs About Search Outcomes - Big Experiment Control Group

	E[search duration] (months, w)		E[wage] (w)	
	(1)	(2)	(3)	(4)
Average skill tercile belief (z-scored)	-0.127***	-0.129***	38.318***	33.755***
	(0.047)	(0.046)	(9.518)	(9.429)
Control mean	2.718	2.718	892.045	892.045
Observations	2148	2144	2183	2179
Controls	No	Yes	No	Yes

Notes: **Table A.5 shows that jobseekers' beliefs about their skills correlate positively with their beliefs about the returns to search in the control group in the big experiment.** All specifications control for average standardized skill levels. Columns 2 and 4 further include controls for age, gender, having worked in a wage job, as well as dummies for three education categories. Winsorized variables (w) are winsorized at the 99^{th} percentile. All monetary values are in 2021 USD in purchasing power parity terms. Heteroskedasticity-robust standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.6: Development of Skill Beliefs Over Time - Big Experiment Control Group

	Aligned CA belief			% aligned skill belief		
	(1)	(2)	(3)	(4)	(5)	(6)
Endline	-0.004	-0.001	-0.015	0.009	0.009	0.004
	(0.012)	(0.017)	(0.014)	(0.009)	(0.013)	(0.011)
Endline × Above median search effort		-0.008			0.000	
		(0.024)			(0.019)	
Above median search effort		0.035**			0.003	
		(0.017)			(0.013)	
Endline × Worked last 7 days			0.034			0.015
			(0.027)			(0.020)
Worked last 7 days			-0.005			0.013
			(0.019)			(0.014)
Constant	0.200***	0.183***	0.202***	0.379***	0.377***	0.375***
	(0.008)	(0.012)	(0.010)	(0.006)	(0.009)	(0.008)
Observations	4405	4315	4315	4456	4365	4365

Notes: **Table A.6 shows that the misalignment of skill beliefs persists over time in the control group of the big experiment even for those who are employed or have above-median search activity.** Estimation is at the survey round times jobseeker level and is restricted to the control group. "CA" stands for comparative advantage. Cols. 1 to 3 show results for aligned beliefs about CA. Cols. 4 to 6 show results for the fraction of aligned skill beliefs. Heteroskedasticity-robust standard errors are in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.7: Association Between Baseline Aligned Comparative Advantage Belief and Other Baseline Characteristics - Tight Experiment

	Aligned CA belief		
	OLS	Ridge	Lasso
	(1)	(2)	(3)
Age	-0.003	-0.003	0.000
	(0.034)		
Female	0.004	0.005	0.000
	(0.028)		
Has completed secondary education	-0.008	0.005	0.000
	(0.053)		
Has post-secondary certificate	-0.029	-0.011	0.000
	(0.044)		
Has post-secondary diploma	0.028	0.021	0.000
	(0.040)		
Has post-secondary degree	-0.063	-0.034	-0.024
	(0.040)		
Employed in any form at baseline	0.013	0.009	0.000
	(0.029)		
Total work experience at baseline (years)	0.052	0.024	0.000
	(0.036)		
Ever held multiyear job	-0.021	-0.009	0.000
	(0.027)		
Ever received feedback during job search	0.002	-0.002	0.000
	(0.029)		
Ever received feedback during employment	-0.036	-0.014	0.000
	(0.030)		
Any HH member employed	0.037	0.026	0.004
	(0.029)		
Has numeracy comparative advantage	-0.164***	-0.093	-0.138
	(0.049)		
Numeracy assessment score (%)	-0.002	-0.024	0.000
	(0.040)		
Communication assessment score (%)	0.067	0.071	0.053
	(0.040)		
Concept formation assessment score (%)	0.021	0.016	0.000
	(0.030)		
Constant	0.486***	0.486	0.486
	(0.027)		
# jobseekers	278	278	278
% of outcome variation explained	0.222	0.117	0.129
p: all coefficients = 0	0.000	.	.

Notes: **Table A.7 shows that baseline CA beliefs are unrelated to most observed jobseeker characteristics.** It displays coefficients from regressions with baseline data of an indicator for aligned CA belief on age, gender, education categories (omitting less than completed high school), multiple measures of work experience, multiple proxies for access to information about skills, and assessment results. All variables are standardized to have mean zero and standard deviation one to make effect sizes comparable. Cols. 1, 2, and 3 show the results from OLS, ridge, and LASSO regressions respectively. Heteroskedasticity-robust standard errors shown in parentheses in column 1 but omitted in columns 2 and 3 because standard errors are not well defined for ridge and LASSO regressions. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.5 Skill Belief Treatment Effects

Treatment in the tight experiment reduces the jobseeker-level variance of beliefs about skill levels, measured by asking jobseekers the probability that their assessment results fall into each quintile (Table A.8, column 1).

Treatment in the tight and big experiments updates underconfident skill beliefs more than overconfident skill beliefs. Treatment increases the share of skill beliefs that match assessment results by reducing the shares of underconfident skill beliefs (columns 2 & 4) and overconfident skill beliefs (columns 3 & 5). The relative reduction in underconfident beliefs (31% and 28% of the control means in the tight and big experiment respectively) is significantly larger than the relative reduction in overconfident beliefs (3.5% and 22% of the control mean in the tight and big experiment respectively). This aligns with lab evidence on asymmetric belief updating (Eil & Rao, 2011).

Jobseekers might update their beliefs about their skill CA through two channels: updated beliefs about (1) their own skill levels or (2) the distribution of other jobseekers' skills. To evaluate these channels, we ask jobseekers how many questions they answered correctly on each assessment, after they took the assessments but before any treatment. We construct a dummy for **accurate score ranking**, equal to one if and only if the jobseeker

Table A.8: Treatment Effects on Belief Variance and Over-/Underconfidence - Both Experiments

	Tight experiment (quintiles)			Big experiment (terciles)	
	Variance	Underconfident	Overconfident	Underconfident	Overconfident
	(1)	(2)	(3)	(4)	(5)
Treatment	-0.049	-0.057***	-0.022	-0.043***	-0.101***
	(0.054)	(0.016)	(0.018)	(0.005)	(0.007)
Treatment effect/ control mean		-0.306***	-0.035	-0.284***	-0.220***
		(0.087)	(0.028)	(0.033)	(0.016)
p[Treat/mean(uc)]=p[Treat/mean(oc)]			0.001		0.000
Control mean	0.644	0.187	0.629	0.150	0.461
Observations	278	278	278	4205	4205

Notes: **Table A.8 shows that the variance of beliefs weakly decreases and that underconfident beliefs update more than overconfident beliefs.** Cols. 1–3 show results for the tight experiment. Cols. 4–5 show results for the big experiment. Col. 1 shows the treatment effect on the variance of the deviation of skill quintile beliefs from assessed quintiles, averaged across communication and numeracy skills. The variance of skill beliefs is measured by asking probabilities that assessment results fall in each quintile. Cols. 2–4 show treatment effects on dummies for believing that one's skills are lower than one's assessment results (cols. 2 & 4) or higher than one's assessment results (cols. 3 & 5), averaged across skill domains. All specifications include randomization block fixed effects and prespecified baseline covariates described in footnotes 18 and 25. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

correctly identifies the assessment on which she answered more questions correctly. We then identify jobseekers who have an accurate score ranking and a misaligned CA belief. These jobseekers can only update their CA beliefs due to channel (2); channel (1) is shut down for them. Treatment substantially updates CA beliefs for this subset of jobseekers ($p = 0.001$). This shows that learning about the distribution of other jobseekers' skills at least partially explains the updating of CA beliefs.

Treatment also increases the weight that jobseekers' beliefs place on assessment results relative to other sources of information. To show this, we run a regression adapted from the standard Bayesian learning specification. In the control group, we regress an indicator for believing one has communication CA on two variables: the difference between one's communication and numeracy quintiles and the difference between one's high school graduation exam scores in English and mathematics, both transformed to have mean zero and standard deviation one. The two coefficient estimates capture the relative weights placed on these two information sources in jobseekers' beliefs. We then repeat this exercise for the treatment group. Treatment increases the weight on assessment results from 0.31 to 0.72, with standard errors 0.25 and 0.07 and $p = 0.122$ for the test of equality.[40] Results are similar for numeracy CA beliefs and with controls for demographics, education, and baseline work experience. We view this exercise as an interesting extension rather than core result for two reasons. First, the scales of the assessment results and graduation exam results are different and we cannot convert the latter into quintiles because the full distribution of exam results is not public. Second, the regression approach we use cannot be easily adapted for our main object of interest, aligned CA beliefs.

Finally, we note that treatment effects on our main outcomes are robust to conditioning on beliefs about CA in high school graduation exams and their interaction with treatment. Adding these variables to the heterogeneous treatment effects specification (4) has almost no impact on treatment effects on aligned CA beliefs for jobseekers with baseline misaligned beliefs (20.8 versus 20.4pp) or baseline aligned beliefs (7.1 versus 6.4pp) or treatment effects on the aligned search index for jobseekers with baseline misaligned beliefs (0.603 versus 0.649 SDs) or baseline aligned beliefs (-0.045 versus 0.01 SDs). This shows that treatment effects are driven by new information provided by assessment results conditional on the prior signal from high school.

B Protocol and Intervention Details

This appendix summarizes the protocol and intervention details for both experiments.

[40]Control group jobseekers' beliefs are unsurprisingly correlated with their assessment results, reflecting the many common factors that may influence their beliefs and assessment results. However, the relationship varies substantially across jobseekers, so the control group belief weight has a large standard error.

B.1 Tight Experiment

The data collection for the tight experiment ran for August – October 2022 in downtown Johannesburg. We recruited 373 participants using the user database of our implementation partner, the Harambee Youth Employment Accelerator. We contacted users who were active on our partner's platform SAYouth.mobi in the past month, lived within commuting distance of our field office, had completed at least high school, and were at most 35 years old. Using this contact list, surveyors called potential participants to establish eligibility (currently searching for work and not a full-time student) and invite eligible respondents to a day-long job search workshop. We offered 150 Rand (roughly 21 USD PPP) mobile phone airtime for compensation, which is equivalent to 6.5 hours work at the national minimum wage. The structure of data collection is shown in Figure 1.

When participants arrived at the venue, they received information about the schedule of the day, had breakfast, and were matched to a surveyor. The surveyor sought informed consent and started administering the surveys, programmed in Survey CTO, on a tablet. The surveyors were instructed to provide further explanations and translate the questions as needed, and to tailor the pace of the surveys to the needs of the participants.

The baseline survey collected participants' demographic information, their employment and job search history, baseline beliefs about their skills and their labor market prospects, and their risk and time preferences. This survey was followed by three assessments: communication, numeracy, and concept formation, in that order. Participants had 30 minutes for each of the communication assessment and numeracy assessment and 15 minutes to complete the concept formation test. After the assessments, the surveyor administered a short survey about participants' beliefs about their assessment performance.

On treatment days, the surveyors handed over the printed reports (Figure 2) to participants who then watched a video on the tablet with headphones. The video explained how participants should interpret the report, used several hypothetical examples for further explanation and prompted participants to review their own report and ask the surveyor any questions that they had. On control days, participants still viewed a minimally modified version of the video that omitted the explanation of the results. The scripts and the video were thoroughly piloted to ease participants' understanding. The treatment video is available at `https://bit.ly/3EoVoNL` and the control video is available at `https://bit.ly/3srwLgj`. The corresponding scripts are available at `https://bit.ly/45zthqu`. Baseline covariates are balanced across treatment arms (Table B.1).

After the treatment and a lunch break, the surveyors collected participants' beliefs about their skills and future labor market outcomes, and they administered the job choice task. In the job choice task, participants were asked to choose between two realistic jobs.

Table B.1: Balance Table - Tight Experiment

	Restricted sample					Full sample				
	Control (1)	Treatment (2)	Δ (3)	p(Δ = 0) (4)	N (5)	Control (6)	Treatment (7)	Δ (8)	p(Δ = 0) (9)	N (10)
Panel A: Demographics										
Black African	1.00	1.00	0.00	.	278	0.99	0.99	0.01	0.41	372
Male	0.33	0.32	-0.01	0.74	278	0.28	0.29	0.00	0.92	372
Age	26.42	26.40	-0.02	0.84	278	26.89	26.79	-0.09	0.91	372
Completed secondary education only	0.61	0.60	-0.01	0.68	278	0.62	0.58	-0.04	0.31	372
University degree / diploma	0.19	0.24	0.04	0.16	278	0.19	0.21	0.02	0.60	372
Any other post-secondary qualification	0.16	0.14	-0.01	0.46	278	0.15	0.17	0.02	0.42	372
Panel B: Labor market background										
Any work in last 7 days	0.35	0.31	-0.04	0.41	278	0.35	0.32	-0.03	0.49	372
Has worked in permanent wage job before	0.23	0.27	0.04	0.50	278	0.24	0.27	0.03	0.44	372
Earnings in USD (last 7 days, w)	46.28	43.53	-2.75	0.85	277	47.45	50.52	3.07	0.69	371
Written contract	0.09	0.16	0.06	0.04	278	0.11	0.18	0.07	0.03	372
Panel C: Search behavior										
Any job search in last 30 days	0.96	0.96	-0.01	0.72	278	0.96	0.96	-0.00	0.90	372
# applications (last 7 days, w)	11.31	8.69	-2.62	0.10	278	11.46	10.49	-0.97	0.54	372
Search expenditure in USD (last 7 days, w)	23.98	21.47	-2.51	0.25	278	24.42	22.65	-1.77	0.28	372
Hours spent searching (last 7 days, w)	13.75	13.90	0.15	0.92	278	14.06	14.23	0.17	0.88	371
# job offers (last 30 days, w)	0.14	0.21	0.07	0.11	278	0.20	0.23	0.03	0.51	372
Panel D: Search alignment with CA										
Δ planned apps (w, aligned - misaligned)	0.94	0.53	-0.41	0.55	278	0.57	0.40	-0.17	0.77	372
Δ % platform apps (aligned - misaligned)	-0.00	-0.00	0.00	0.73	278	-0.00	-0.00	0.00	0.74	372
Panel E: Skills beliefs										
Aligned belief about CA	0.47	0.50	0.04	0.68	278	0.43	0.46	0.03	0.69	369
Fraction aligned belief domains	0.18	0.26	0.08	0.12	278	0.18	0.24	0.06	0.17	369

Notes: **Table B.1 shows that big experiment covariates are balanced across treatment arms.** "CA" stands for comparative advantage. "(w)" denotes variables winsorized at the 99^{th} percentile. All monetary values are in 2021 USD in purchasing power parity terms. Cols. 1–5 show statistics for the baseline sample. Cols. 6–10 show statistics for the participants reached at endline: 96% of the control group and 95.5% of the treatment group. P-values are from regressions of each covariate on treatment and randomization block fixed effects using standard errors clustered by workshop day.

Each job pair contained jobs with opposite skill demand (one communication- and one numeracy-heavy job) based on 13 recruiters' prior evaluation of the jobs. The jobs in the pairs were matched on several important dimensions: expected desirability, job-offer probability, and salary (as assessed by the recruiters), as well as location. These were all entry-level jobs that did not require certifications or equipment to ensure that all participants could reasonably apply for them. See Table 3 for the pairs of job titles. The job descriptions and their layout followed the design and style of the SAYouth.mobi platform. The job descriptions were presented side by side in a printed booklet to allow for easy comparisons. See Figure 3 for an example of one pair. Participants were shown the pairs in the booklet, read the descriptions and were asked to pick the job that they were most interested in applying to. Participants made decisions for the same set of 11 job pairs in a randomized order. The last two job pairs explicitly included the main skill (communication or numeracy) that the job required. For 5 of the 11 job pairs, after participants made their choice, we asked them the job offer probability and expected wage for each job if they were to apply to it.

We incentivized the job choices in two ways. First, one job pair of the 11 pairs included real job opportunities, and we submitted participants' application materials to the job that participants picked for this real pair. Participants were informed about this incentive, but they did not learn which pair was the real pair during the workshop. As a second incentive, participants received a list of job titles at the end of the workshop that matched their most preferred skill according to their choices to assist their job search.

In the next survey module, we measured participants' willingness-to-pay using incentivized multiple price lists for three products: a document that revealed the skill demand of a set of common jobs as per the rating of HR experts, and self-study materials to improve each of communication and numeracy skills. Participants completed a practice round before the elicitation. Further details of the WTP protocol and results are included in the Online Supplement.

In the next session we collected information to populate a CV template for each participant. After the workshop we delivered this CV to the participant and to their chosen job from the choice task.

The final session of the workshop measured search effort. Participants had the opportunity to spend up to 15 minutes to write an optional short cover letter to the employer of their chosen job from the choice task. Participants could choose to end this task early or skip it entirely. The message, along with participants' CV that we created, were both delivered to the employer, whose HR team could then evaluate the candidates. After this part, participants completed the check out procedure and received their compensation.

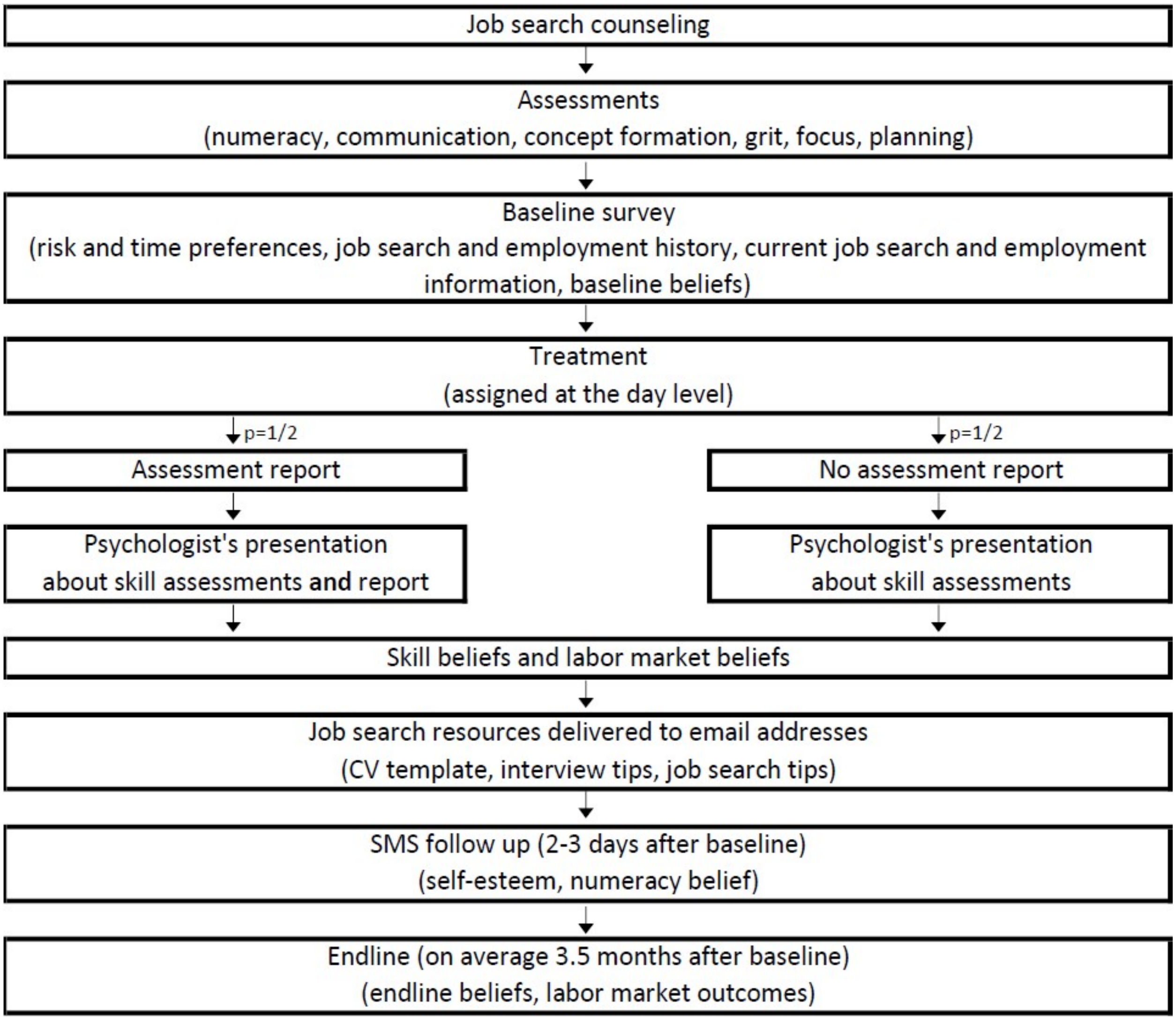
B.2 Big Experiment

The recruitment, data collection procedure and intervention were similar in the tight and big experiments. Here we highlight the main differences between the two experiments. Figure B.1 shows the order of activities during big experiment workshops.

The big experiment took place in the same labor market, but earlier in time, in 2016/17. We again recruited participants from the contact list of the Harambee Youth Employment Accelerator who were aged 18–34, had completed secondary school, had at most 12 months of formal work experience, and were from disadvantaged backgrounds. The workshops took place at the Harambee office at the time, in downtown Johannesburg.

As in the tight experiment, jobseekers completed surveys and assessments. Treatment was randomized at the day level. Treated participants received a report (Figure B.2) showing results in terciles for six skills, versus results in terciles for two skills in the tight experiment. Participants received their reports privately in sealed envelopes and then attended a group briefing from an industrial psychologist, covering similar content to the tight experiment video. The materials used for the briefing are available at

Figure B.1: Big Experiment Design



Notes: **Figure 1 shows the order of activities during each workshop in the big experiment.**

https://bit.ly/44oOp1v and the script is available at https://bit.ly/3YLaDK6.

We conducted two follow-up surveys. A short SMS survey 2-3 days after the workshop, and a longer phone survey on average 3.5 months after the workshop. 96% of the control group and 95.5% of the treatment group were successfully surveyed at endline. Respondents were compensated with mobile phone airtime.

Summary statistics for the big experiment sample (Table B.2) are similar to those for the tight experiment sample (Table 1). Baseline covariates in the big experiment are balanced across treatment arms in the baseline sample and the sample surveyed at endline (Table B.3).

Figure B.2: Sample Treatment Report - Big Experiment

REPORT ON CANDIDATE COMPETENCIES
-Personal Copy-

This report contains results from the assessments you took at Harambee in Phase 1 and Phase 2. These results can help you learn about some of your strengths and weaknesses and inform your job search.

You completed assessments on English Communication (listening, reading and comprehension) and Numeracy today in Phase 2. In Phase 1, you completed a Concept Formation assessment which asked you to identify patterns.

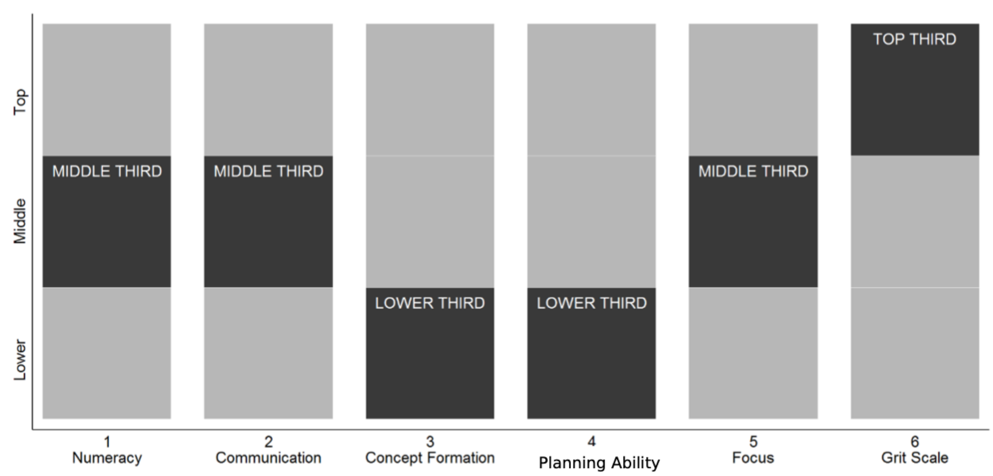
1. The Numeracy tests measure various maths abilities. Your score is the average of the two maths tests you did today at Harambee.
2. The Communication test measures English language ability through listening, reading and comprehension.
3. The Concept Formation test measures the ability to understand and solve problems. Candidates with high scores can generally solve complex problems, while lower scores show an ability to solve less complex problems.

You also did some games and questionnaires to measure your soft skills:

4. The Planning Ability Test measures how you plan your actions in multi-step problems. Candidates with high scores generally plan one or more steps ahead in solving complex problems.
5. The Focus Test looks at your ability to pick out which information is important in confusing environments. Candidates with high scores are able to focus on tasks in distracting situations.
6. The Grit Scale measures candidates' determination when working on difficult problems. Candidates with high scores spend more time working on the problems rather than choosing to pursue different problems.

Your results have been compared to a large group of young South African job seekers who have a matric certificate, are from socially disadvantaged backgrounds and have been assessed by Harambee.

You scored in the MIDDLE THIRD of candidates assessed by Harambee for Numeracy, MIDDLE THIRD for Communication, LOWER THIRD for Concept Formation, LOWER THIRD for Planning Ability, MIDDLE THIRD for Focus and TOP THIRD for the Grit Scale.



DISCLAIMER
Please note that this is a confidential assessment report and is intended for use by the person specified above. Assessment results are not infallible and may not be entirely accurate.

Notes: **Figure B.2 shows an example of the reports given to treated jobseekers in the big experiment.** Each report contains the jobseeker's assessment results but no identifying information and no branding by Harambee.

Table B.2: Summary Statistics - Big Experiment

	(1) Mean	(2) Median	(3) Min	(4) Max	(5) SD	(6) Obs.
Panel A: Demographics						
Black African	0.98	1.00	0.00	1.00	0.12	4389
Male	0.38	0.00	0.00	1.00	0.48	4389
Age	23.67	23.14	18.04	35.08	3.28	4389
Completed secondary education only	0.61	1.00	0.00	1.00	0.49	4389
University degree / diploma	0.17	0.00	0.00	1.00	0.37	4389
Any other post-secondary qualification	0.22	0.00	0.00	1.00	0.41	4389
Panel B: Labor market background						
Any work in last 7 days	0.37	0.00	0.00	1.00	0.48	4389
Has worked in permanent wage job before	0.09	0.00	0.00	1.00	0.29	4377
Earnings in USD (last 7 days, winsorized)	31.26	0.00	0.00	476.00	75.72	4389
Panel C: Search behavior						
Any job search in last 7 days	0.97	1.00	0.00	1.00	0.17	4389
# applications (last 30 days, winsorized)	9.34	5.00	0.00	90.00	12.85	4346
Search expenditure in USD (last 7 days, winsorized)	30.97	20.40	0.00	204.00	33.05	3995
Hours spent searching (last 7 days, winsorized)	17.06	8.00	0.00	96.00	19.68	4273
# job offers (last 30 days, winsorized)	0.80	0.00	0.00	20.00	2.66	4335
Panel D: Skills beliefs						
Aligned belief about CA	0.20	0.00	0.00	1.00	0.40	4312
Fraction of aligned belief domains	0.38	0.33	0.00	1.00	0.31	4378

Notes: **Table B.2 shows big experiment summary statistics.** "CA" stands for comparative advantage. Winsorization is at the 99^{th} percentile. All monetary values are in 2021 USD purchasing power parity terms.

Table B.3: Balance Table - Big Experiment

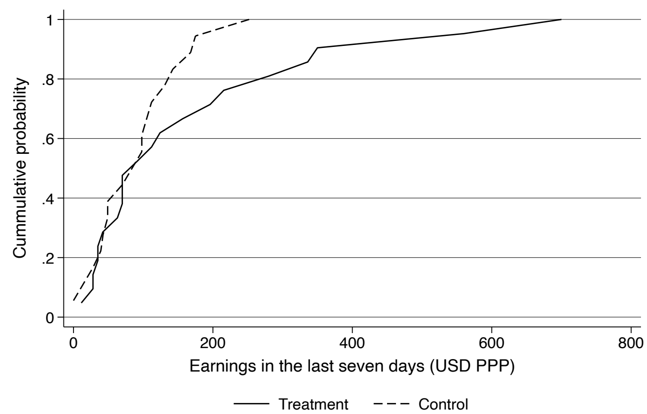
	Full sample					Non-attrited sample				
	Control (1)	Treatment (2)	Δ (3)	p(Δ = 0) (4)	N (5)	Control (6)	Treatment (7)	Δ (8)	p(Δ = 0) (9)	N (10)
Panel A: Demographics										
Black African	0.98	0.99	0.00	0.93	4389	0.98	0.99	0.00	0.98	4206
Male	0.39	0.36	-0.02	0.04	4389	0.39	0.36	-0.03	0.03	4206
Age	23.55	23.79	0.25	0.07	4389	23.53	23.80	0.26	0.05	4206
Completed secondary education only	0.62	0.59	-0.02	0.20	4389	0.62	0.59	-0.03	0.17	4206
University degree / diploma	0.16	0.18	0.02	0.29	4389	0.15	0.18	0.02	0.24	4206
Any other post-secondary qualification	0.21	0.22	0.01	0.61	4389	0.22	0.23	0.01	0.63	4206
Panel B: Labor market background										
Any worked in last 7 days	0.36	0.39	0.02	0.22	4389	0.36	0.39	0.03	0.19	4206
Has worked in permanent wage job before	0.09	0.09	-0.00	0.61	4377	0.10	0.09	-0.01	0.35	4195
Earnings in USD (last 7 days, w)	30.09	32.52	2.43	0.13	4389	29.84	32.49	2.65	0.11	4206
Panel C: Search behavior										
Any job search in last 7 days	0.97	0.97	0.01	0.07	4389	0.97	0.98	0.01	0.02	4206
# applications (last 30 days, w)	9.31	9.37	0.06	0.95	4346	9.13	9.27	0.14	0.79	4165
Search expenditure in USD (last 7 days, w)	32.09	31.14	-0.95	0.34	3912	32.02	31.15	-0.87	0.42	3747
Hours spent searching (last 7 days, w)	17.35	16.74	-0.61	0.32	4273	17.24	16.56	-0.68	0.28	4093
# job offers (last 30 days, w)	0.77	0.84	0.06	0.64	4335	0.78	0.87	0.09	0.44	4152
Panel D: Skills beliefs										
Aligned belief about CA	0.20	0.21	0.01	0.49	4312	0.20	0.21	0.01	0.67	4132
Fraction aligned belief domains	0.38	0.38	-0.00	0.49	4378	0.38	0.37	-0.00	0.45	4196

Notes: **Table B.3 shows that big experiment covariates are balanced across treatment arms.** "CA" stands for comparative advantage. "(w)" denotes variables winsorized at the 99^{th} percentile. All monetary values are in 2021 USD in purchasing power parity terms. Cols. 1–5 show statistics for the baseline sample. Cols. 6–10 show statistics for the participants reached at endline: 96% of the control group and 95.5% of the treatment group. P-values are from regressions of each covariate on treatment and randomization block fixed effects using standard errors clustered by workshop day.

C Additional Treatment Effects

Figure C.1 shows treatment increases earnings in the tight experiment for jobseekers with misaligned beliefs about their CA at baseline, who receive more information from treatment ($p = 0.057$). In contrast, treatment and control earnings do not differ for jobseekers with aligned baseline beliefs about their CA.

Figure C.1: Cumulative Distributions of Earnings by Treatment Group - Tight Experiment



Notes: **Figure C.1 shows that treatment increases earnings in the tight experiment for jobseekers with misaligned baseline CA beliefs.** This figure uses only employed jobseekers, as treatment does not affect the employment rate. Earnings are measured in a text message survey one month after the workshop, which asks about all sources of labor market earnings in the preceding week. Earnings are measured in 2021 USD purchasing power parity terms and are winsorized at the 99^{th} percentile.

C.1 Robustness Checks

This subsection replicates main treatment effect estimates from the both experiments using different samples, methods, and variable definitions. Our main analysis in the tight experiment omits jobseekers without a unique CA because some measures of skill-directed search are not sensibly defined for them. Including these jobseekers with their skill-directed search coded as zeroes does not substantially change treatment effects on search direction and beliefs about skills (Table C.1). Belief, search, and labor market results in both experiments are also robust to estimating p-values using a wild cluster bootstrap, estimating sharpened q-values to account for multiple hypothesis testing, and reweighting the samples in the two experiments to have the same distribution of baseline covariates (Table C.2). Treatment effects on beliefs are robust to using non-binary measures of alignment between assessment results and beliefs about skills, both as outcomes and for heterogeneity analysis (Table C.3). Treatment effects on earnings are robust to different transformations (Table C.4) and are non-negative throughout the distribution (Figure C.2).

Table C.1: Treatment Effects on Beliefs and Aligned Search Direction - Tight Experiment Including Jobseekers Without Unique Comparative Advantage

	Beliefs						Aligned search direction									
	Al. CA belief (implied)		Al. CA belief (direct)		Fraction al. beliefs		% aligned (job choice)		Δ % al. platform apps		Δ SMS click rate		Δ planned apps (w)		Al. search index	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Treatment	0.120***	0.180***	0.092**	0.213***	0.089***	0.064***	0.032	0.062**	0.053***	0.063**	0.071	0.122*	1.121	3.363***	0.241***	0.436***
	(0.037)	(0.055)	(0.037)	(0.047)	(0.020)	(0.018)	(0.025)	(0.026)	(0.017)	(0.030)	(0.046)	(0.063)	(0.878)	(1.140)	(0.078)	(0.094)
Treatment ×		-0.130		-0.278***		0.057		-0.070**		-0.022		-0.113		-5.190**		-0.445***
Al. CA belief (bl)		(0.089)		(0.066)		(0.041)		(0.029)		(0.060)		(0.098)		(2.162)		(0.156)
Al. CA belief (bl)		0.447***		0.331***		-0.039		0.111***		0.010		0.070		6.113***		0.504***
		(0.081)		(0.073)		(0.037)		(0.026)		(0.035)		(0.076)		(1.872)		(0.121)
Effect:		0.050		-0.065		0.121***		-0.008		0.041		0.009		-1.827		-0.009
Al. CA belief (bl)		(0.061)		(0.047)		(0.036)		(0.029)		(0.039)		(0.070)		(1.613)		(0.118)
Control mean	0.446	0.446	0.467	0.467	0.171	0.171	0.416	0.416	0.005	0.005	-0.024	-0.024	3.272	3.272	-0.000	-0.000
Observations	368	368	368	368	368	368	368	368	368	368	368	368	368	368	368	368

Notes: **Table C.1 shows that treatment effects on beliefs and aligned search direction are robust to including jobseekers who do not have a unique skill CA.** The outcome measures are identical to those shown in Tables 2 and 4. All search direction measures are coded as zero for jobseekers with tied skill quintiles. "CA" stands for comparative advantage in skills. Aligned job search is defined as directing search effort toward jobs that mostly require the skill that aligns with jobseekers' measured CA. Columns indicate different outcomes: a dummy indicating if a jobseeker's belief about her CA in skills based on quintile beliefs is aligned with her assessed CA (cols. 1–2), a dummy indicating if jobseeker's directly measured CA beliefs align with her assessed CA (cols. 3–4), the fraction of skills where her believed and assessed quintile equal (cols. 5–6), the percentage of 11 incentivized job choices that are aligned with the measured CA of the jobseeker (cols. 7–8), the difference between the percentage of aligned and non-aligned applications on the online job search platform SAYouth.mobi (cols. 9–10), the difference in link click rates between aligned and non-aligned jobs sent to job seekers via SMS (cols. 11–12), the difference between aligned and non-aligned planned applications for the 30 days after the workshop (cols. 13–14), and the inverse covariance-weighted average of the measures displayed in cols. 7–14 following Anderson (2008) (cols. 15–16). Even-numbered columns show heterogeneity by whether individuals have aligned CA beliefs at baseline. All regressions include prespecified control variables listed in footnote 18 and a dummy variable equal to one iff the jobseeker has no CA, which is prespecified but cannot be used when we restrict the sample to jobseekers with unique CA. Winsorized variables (w) are winsorized at the 99th percentile. Standard errors clustered at the treatment-day level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.2: Robustness Checks for Main Outcomes - Comparing Tight and Big Experiments

	Tight experiment		Big experiment					
	Aligned CA belief	Al. search index	Al. CA belief	Al. search	Worked last 7 days	Worked since bl	Uncond. earnings	Cond. earnings
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Robustness for inference								
Treatment	0.135	0.254	0.139	0.050	0.009	0.030	6.517	20.393
clustered p-value	(0.001)	(0.017)	(0.000)	(0.000)	(0.453)	(0.033)	(0.020)	(0.035)
wild-bootstrapped p-value	[0.005]	[0.011]	[0.000]	[0.000]	[0.850]	[0.168]	[0.004]	[0.003]
q-value	{0.003}	{0.021}	{0.001}	{0.001}	{0.061}	{0.026}	{0.021}	{0.026}
Panel B: Reweighting to match other sample								
Treatment	0.149***	0.254**	0.144***	0.054***	0.015	0.028*	6.202**	18.393*
	(0.044)	(0.102)	(0.011)	(0.010)	(0.013)	(0.014)	(2.803)	(9.614)
Control mean	0.475	0.000	0.196	0.165	0.309	0.671	25.424	85.826
Observations	278	278	4118	4205	4204	4205	4196	1280
Number of clusters	34	34	54	54	54	54	54	54

Notes: **Table C.2 shows that the main results are robust to different hypothesis testing approaches and to reweighting data to match the demographics across experiments. Panel A** shows p-values from cluster-robust standard errors, p-values from a wild cluster bootstrap with 10,000 replications, and sharpened q-values that control the false discovery rate across tests on all outcomes shown in the table, following Benjamini et al. (2006). **Panel B** shows the same treatment effects estimated with the experimental samples reweighted to have the same distribution of baseline covariates. The covariates used to estimate weights are all variables displayed in the summary statistics Tables 1 and B.2 that are consistently measured in both experiments: dummies for being Black, male, having only completed secondary education, having a post-secondary degree, or having a post-secondary certificate, having worked in the last seven days, and having ever worked in a wage job, as well as continuous measures of earnings, the number of job applications, job search hours and expenditure, and job offers in the last seven days. Cols. 1–2 show effects on the main outcomes of the tight experiment. Cols. 3–8 show effects on the main outcomes of the big experiment. Control variables for the tight experiment are listed in footnote 18. Control variables for the big experiment are listed in footnote 25. Standard errors clustered at the treatment-day level in parentheses in panel B. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ in Panel B.

Table C.3: Heterogeneous Treatment Effects by Nonbinary Baseline Comparative Advantage Beliefs - Tight Experiment

	Main outcomes				Additional belief outcomes			
	Aligned search index		Aligned CA belief		Degree of CA alignment		Overall alignment of belief levels	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.269**	0.842*	0.152***	0.444***	0.063***	0.324***	0.073***	0.138**
	(0.103)	(0.414)	(0.044)	(0.142)	(0.012)	(0.091)	(0.012)	(0.062)
Treatment × Baseline degree of CA belief alignment		-0.681		-0.348*		-0.315***		-0.078
		(0.504)		(0.176)		(0.106)		(0.072)
Baseline degree of CA belief alignment		1.595***		1.125***		0.828***		0.045
		(0.495)		(0.215)		(0.071)		(0.050)
Control mean	-0.000	-0.000	0.475	0.475	0.816	0.816	0.511	0.511
Observations	278	278	278	278	278	278	278	278

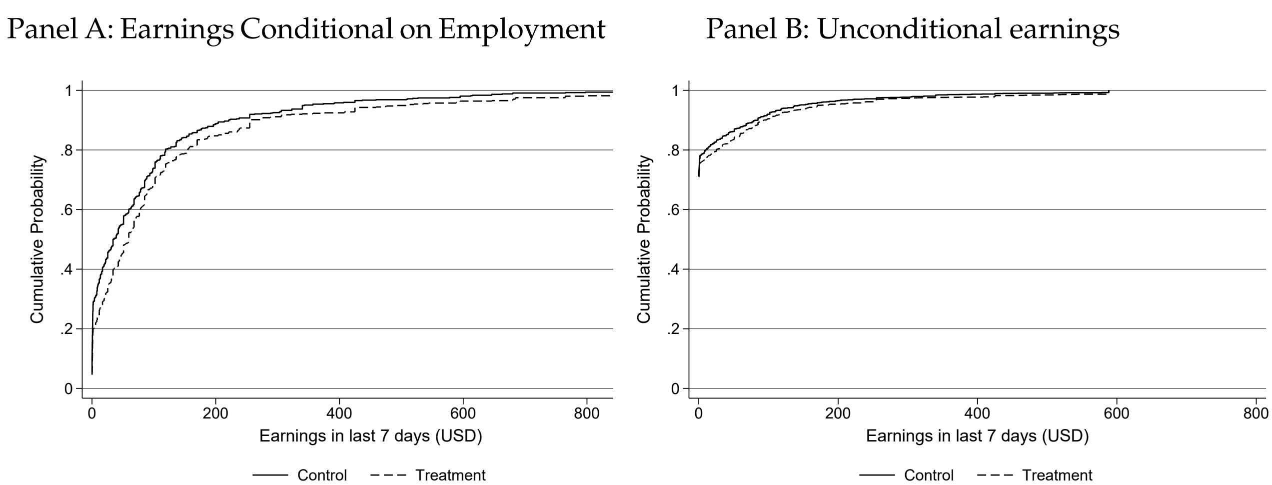
Notes: **Table C.3 shows that average and heterogeneous treatment effects are robust to using nonbinary measures of aligned skill beliefs in the tight experiment.** The outcomes in cols 1–4 are the same as those used in the main analysis. "Degree of CA belief alignment" captures the degree to which comparative advantage beliefs are aligned. It is defined as $1 - max\{0, abs(Q_{low} - Q_{high} + 1)\}/5$, where Q_{low} is the jobseeker's believed quintile for the skill in which they actually score lower and Q_{high} is the jobseeker's believed quintile for the skill in which they actually score higher. It is equal to 0 for jobseekers with aligned CA beliefs (matching the binary variable), is between 0 and 1 for all other jobseekers, and is increasing in "how much" the jobseekers' beliefs about their skill quintiles would need to change to match their CA in the assessment results. This is the outcome in cols. 5–6 and is used for heterogeneity analysis in all even-numbered columns. The mean of this measure is 0.83 and the standard deviation is 0.22. "Overall alignment of belief levels" is defined as one minus the average difference between beliefs about skill quintiles and assessed quintiles divided by the maximum possible difference. A value of 1 indicates perfectly aligned beliefs while a value of 0 indicates maximally misaligned beliefs. This is the outcome in cols 7–8. Control variables are listed in footnote 18. Standard errors clustered at the treatment-day level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.4: Treatment Effects on Different Transformations of Earnings - Big Experiment

	Unconditional earnings					Conditional earnings				
	raw	wins	z	ln(earn+1)	ihs	raw	wins	z	ln(earn+1)	ihs
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment	9.360**	6.517**	0.097**	0.115**	0.127**	25.100**	20.393**	0.159**	0.229**	0.244**
	(3.610)	(2.712)	(0.037)	(0.053)	(0.060)	(11.348)	(9.404)	(0.072)	(0.093)	(0.101)
Control mean	27.080	25.424	-0.000	0.954	1.109	88.109	85.826	-0.000	3.105	3.608
Observations	4196	4196	4196	4196	4196	1280	1280	1280	1280	1280

Notes: **Table C.4 shows that treatment effects on earnings in the seven days before the endline survey are robust to a range of transformations.** Columns show effects on different transformations of the unconditional earnings variable (cols. 1-5) and the earnings variable conditional on doing any work in the seven days before the endline survey (cols. 6-10). Outcomes in each column are raw earnings (cols. 1 and 6), earnings winsorized earnings at the 99^{th} percentile (cols. 2 and 7), earnings standardized with respect to the control group mean and standard deviation (cols. 3 and 8), the natural logarithm of earnings plus one (cols. 4 and 9), and the inverse hyperbolic sine transformation of earnings (cols. 5 and 10). Control variables are listed in footnote 25. Earnings are measured in 2021 USD purchasing power parity terms. Standard errors clustered at the treatment-day level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure C.2: Cumulative Distributions of Earnings by Treatment Group - Big Experiment



Notes: **Figure C.2 shows that treatment effects on earnings in the big experiment reflect a rightward shift in the entire distribution.** It shows the cumulative distribution function of earnings in the last seven days in the big experiment for the treatment and control groups. Earnings are measured in 2021 USD purchasing power parity terms and are winsorized at the 99^{th} percentile. Panel A shows earnings conditional on having worked in the last seven days. Panel B shows unconditional earnings for the full sample.

C.2 Heterogeneous Treatment Effects

This section shows heterogeneous treatment effects mentioned in the text of the main paper. Table C.5 shows heterogeneous treatment effects on beliefs about skills, search direction, and search effort by baseline aligned CA beliefs and baseline confidence in the tight experiment. These show that the heterogeneous treatment effects by baseline aligned CA beliefs are broadly robust to allowing effects to vary also by baseline confidence. Table C.6 shows heterogeneous treatment effects on search direction in the tight experiment's job choice task by whether skill requirements for specific job pairs were revealed. This shows that effects on search direction are larger when skill requirements are revealed but are positive even when they are not revealed. Table C.7 shows average and heterogeneous treatment effects on beliefs about skills and search direction by baseline aligned CA belief in the big experiment. These are consistent with the results of tight experiment.

Table C.5: Heterogeneous Treatment Effects on Main Outcomes by Baseline Aligned Comparative Advantage Beliefs and Confidence - Tight Experiment

	Aligned CA belief (1)	Aligned search index (2)	Search effort index (3)
Treatment	0.216***	0.550***	-0.192
	(0.066)	(0.155)	(0.177)
Treatment × Aligned CA belief (bl)	-0.161	-0.651**	0.274
	(0.112)	(0.247)	(0.298)
Aligned CA belief (bl)	0.737***	0.539**	-0.053
	(0.110)	(0.211)	(0.226)
Treatment × Average confidence (bl)	-0.003	0.084	0.081
	(0.079)	(0.179)	(0.175)
Average confidence (bl)	0.072	0.061	0.201
	(0.058)	(0.142)	(0.143)
Treatment × Aligned CA belief (bl)	0.013	-0.065	-0.167
× Average confidence (bl)	(0.111)	(0.208)	(0.230)
Control mean	0.475	0.000	0.000
Observations	278	278	278

Notes: **Table C.5 shows that the main results are robust to interacting the treatment with both baseline aligned CA beliefs and baseline confidence in skills.** The magnitudes of the interaction terms in row 2 are very similar to the estimates from regressions that exclude baseline confidence, shown in Tables 2 and 4. The estimated coefficients on the three-way interaction term (row 6) are small and not statistically significant. "CA" stands for comparative advantage in skills and "bl" stands for baseline. Baseline confidence levels are defined as the average baseline deviation of beliefs about skill quintiles from measured quintiles (across numeracy and communication), divided by the control group standard deviation. Positive values indicate overconfidence and negative values indicate underconfidence. Columns indicate different outcome variables: a dummy indicating aligned CA beliefs (col. 1), the aligned search index (col. 2), and the search effort index (col. 3). Control variables are listed in footnote 18. Standard errors clustered at the treatment-day level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.6: Heterogeneous Treatment Effects on Search Direction for Jobs With and Without Revealed Skill Demands - Tight Experiment

	Choice aligned with measured comp. adv.	
	(1)	(2)
Treatment	0.039	0.023
	(0.032)	(0.034)
Treatment × Skill req. revealed	0.130**	-0.114
	(0.063)	(0.085)
Treatment effect:	0.169***	-0.091
Skill req. revealed	(0.060)	(0.077)
Control mean	0.550	0.550
Observations	1573	1485
Sample: jobseekers with baseline CA beliefs	misaligned	aligned

Notes: **Table C.6 shows that, for jobseekers with misaligned baseline CA beliefs, treatment effects on search direction in the job choice task are stronger when jobs' skill requirements are revealed.** The outcome variable is a dummy equal to one if job choices in the job choice task are aligned with jobseekers' assessed CA. The effects are estimated at the job-pair × individual level for jobseekers with misaligned baseline beliefs about CA (col. 1) and for jobseekers with aligned baseline beliefs (col. 2). Controls include randomization block fixed effects, job pair and job pair order fixed effects, and prespecified baseline covariates described in footnote 18. "CA" stands for comparative advantage in skills. Standard errors clustered at the treatment-day level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.7: Treatment Effects on Beliefs about Skills & Search Direction - Big Experiment

	Beliefs				Search direction	
	Aligned CA belief		Fraction aligned beliefs		Aligned search	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.139***	0.142***	0.142***	0.134***	0.050***	0.060***
	(0.011)	(0.011)	(0.009)	(0.009)	(0.010)	(0.012)
Treatment × Aligned CA belief (bl)		-0.014		0.043**		-0.057*
		(0.036)		(0.020)		(0.032)
Aligned CA belief (bl)		0.157***		-0.008		-0.013
		(0.027)		(0.013)		(0.024)
Treatment effect:		0.129***		0.177***		0.003
Aligned CA belief (bl)		(0.034)		(0.020)		(0.028)
Control mean	0.196	0.196	0.388	0.388	0.165	0.165
Observations	4118	4118	4195	4131	4205	4131

Notes: **Table C.7 shows that treatment effects on skill beliefs and search direction in the big experiment are broadly consistent with the tight experiment.** "CA" stands for skill comparative advantage and "bl" for baseline. Columns show different outcomes: a dummy for aligned CA beliefs (cols. 1 & 2), the fraction of aligned skill beliefs (cols. 3 & 4), and a dummy indicating if self-reported search direction is aligned with skill CA (col. 5 & 6). Control variables are listed in footnote 25. Standard errors clustered at the treatment-day level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D Search Effort

This appendix provides a more detailed description and interpretation of the search effort results and conceptual framework summarized in Section 5. Jobseekers are, on average, overconfident about their skill relative to the reference population: in the tight experiment control group, 63% of beliefs about skill quintiles are above assessment results and only 19% are below; these shares are 46% and 15% in the big experiment (Table A.8). Thus, jobseekers receive on average negative news about their skill levels. If jobseekers react to this information by changing search effort, this could also affect labor market outcomes.

Treatment effects on beliefs about skill levels: We first document that, on average, jobseekers negatively update their beliefs about their skill levels. In the tight experiment, treatment reduces jobseekers' average beliefs about their level of skills by an average of 0.08 quintiles or 0.15 standard deviations over the two skills (Table D.2, panel A, column 1, $p = 0.025$). In the big experiment, treatment reduces jobseekers' average beliefs about their level of skills by an average of 0.11 terciles or 0.3 standard deviations over the three skills (Table D.3, panel A, column 1, $p < 0.001$).

These effects unsurprisingly vary by baseline beliefs about skill levels. We estimate:

$$Y_{id} = T_d \cdot \beta_1 + T_d \cdot confidence_i \cdot \beta_2 + confidence_i \cdot \beta_3 + \mathbf{X}_{id} \cdot \Gamma + \varepsilon_i \tag{7}$$

where $confidence_i$ is the believed skill level minus the skill assessment result divided by the control group standard deviation.[41] We find that $\hat{\beta}_2 < 0$, so treatment effects on belief levels are more negative for jobseekers with higher levels of baseline confidence: one standard deviation higher baseline confidence is associated with a treatment effect that is 0.26 quintiles more negative in the tight experiment and 0.125 terciles more negative in the big experiment (Tables D.2 and D.3 respectively, panel B, column 1, $p < 0.001$).

Conceptual framework: These updated beliefs about skill levels might influence search effort. To model this, we replace the assumption of fixed total search effort $\bar{E}$ from Section 2.1 with the assumption that jobseekers choose the levels of search for both communication-heavy and numeracy-heavy jobs, E_C and E_N. This gives a utility function with three arguments: the expected outcome of search for communication-heavy jobs, $V_C(S_C, S_N, E_C)$, the expected outcome of search for numeracy-heavy jobs, $V_N(S_C, S_N, E_N)$, and a constraint function capturing the alternative use of time or money allocated to search effort, $A(E_C, E_N)$. For simplicity, we discuss the case of a monetary constraint: $A = Y - P \cdot E_C - P \cdot E_N$, where Y is the jobseeker's unearned income, and P is the price of

[41] In the tight experiment, this uses quintiles and averages over two skills; in the big experiment this uses terciles and averages over three skills. We use the term "confidence" to refer to jobseekers' beliefs about their level of skill relative to other jobseekers, not the precision of their beliefs, which we do not measure at baseline. Some researchers refer to this type of overconfidence as "overplacement" (Moore & Healy, 2008).

job search relative to a numeraire consumption good. But we could instead use a time constraint $A = T - E_C - E_N$, where T is the jobseeker's time endowment; a constraint function incorporating time and money; or an intertemporal budget constraint. The jobseeker's problem becomes

$$\max_{E_C,E_N} U\left(V_C\left(S_C,S_N,E_C\right),V_N\left(S_C,S_N,E_N\right),Y-P\cdot E_C-P\cdot E_N\right). \tag{8}$$

As in Section 2.1, we assume utility is an increasing concave function of all three arguments, expected search outcomes are increasing concave functions of skill and search effort, and search effort and skill are more complementary within than across dimensions.

In this framework, increasing the believed level of either skill has an ambiguous effect on total search effort. To see this, note that a fall in the believed level of communication skill S_C lowers the expected marginal productivity of search for communication-heavy jobs, $\frac{\partial V_C}{\partial E_C}$. This has two effects. First, a substitution effect, which causes the jobseeker to substitute away from search for communication-heavy jobs and toward both search for numeracy-heavy jobs and alternative activities. Second, an income effect: it lowers the expected outcome for any level of search effort, so the jobseeker has to search more for either communication- or numeracy-heavy jobs to maintain the same expected income. The net effect is a increase in search for numeracy-heavy jobs, an ambiguous effect on search for communication-heavy jobs, and hence an ambiguous effect on total search effort.[42]

Treatment effects on search effort: We find little evidence that treatment affects search effort in either experiment. In the tight experiment, treatment effects are negative on five of our six search effort measures but all effects are small – less than 11% of the control group mean – and none is statistically significant.[43] Treatment lowers an index of these search effort measures by 0.08 standard deviations (Table D.2, panel A, column 2, $p = 0.47$) and a prespecified index of search effort on the SAYouth.mobi platform by 0.1 standard deviations (Table D.2, panel A, column 6, $p = 0.29$). In the big experiment, treatment has a tiny effect of 0.003 standard deviations on an index of search effort measures (Table D.3, panel A, column 2, $p = 0.92$). Treatment effects on the three components of this index – applications submitted, hours and money spent searching – are positive but

[42] The framework has a similar structure to the standard static labor supply model. In that model, a lower wage decreases work effort because the return to work is lower (substitution effect) but increases work effort to afford the same consumption level (income effect). Abebe et al. (2022) also show that raising expected job search outcomes has an ambiguous effect on search effort using a frictional matching model.

[43] Our six search effort measures are: planned applications in the upcoming month, asked during the workshop; time spent drafting a cover letter during the workshop; click rate on three text messages with links to job adverts sent after the workshop; and three measures of job search on the SAYouth.mobi platform in the month after the workshop: days active, jobs viewed, and applications submitted. The planned applications, text messages, and job applications are described in Section 3.4. The cover letter is a task measure of real search effort: the time jobseekers choose to spend writing a cover letter for a real job application at the end of the workshop, on a tablet we provided, rather than leaving early with their incentive.

tiny ($< 3\%$ of the control group mean) and none is statistically significant.

Treatment effects on search effort also do not vary substantially by jobseekers' baseline confidence levels in either experiment. We estimate equation (7) and show heterogeneous treatment effects by baseline confidence about skills in even-numbered columns in Table D.2 for the tight experiment and Table D.3 for the big experiment.

In the tight experiment, none of the interaction terms are statistically significant, the effect sizes are mostly small, and the signs of the interaction terms vary across search effort measures. The interaction effect on the main search effort index for the tight experiment is a tiny 0.02 (Table D.2, panel A, column 2, $p = 0.87$). This implies that a jobseeker with a one standard deviation higher confidence level at baseline has just a 0.02 standard deviation higher treatment effect on search effort. The platform-based measure has a somewhat larger interaction effect of 0.14 standard deviations but it is still not statistically significant (Table D.2, panel B, column 6, $p = 0.25$). The key search direction results are also robust to including the interaction with baseline confidence and baseline CA beliefs in the same regression (Table C.5).

Similarly, treatment effects on search effort in the big experiment do not vary by jobseekers' baseline confidence levels. For the search effort index, the interaction term is a tiny 0.03 standard deviations (Table D.3, panel B, column 2, $p = 0.41$). For the index components, the interaction effects are positive but small and not statistically significant.

Treatment provides information both about skill levels and CA. To isolate the role of information about skill levels, we repeat our analysis for a group of jobseekers who cannot get information about their CA from treatment: those with tied communication and numeracy quintiles. These jobseekers are excluded from our main analysis because some measures of skill-directed job search are not sensibly defined for them. But the measures of search effort can be defined for them. Effects on search effort for these jobseekers are not substantially different to the effects for our main analysis sample (Table D.1). However, effects are naturally less precisely estimated for this smaller group of jobseekers.

In addition to adjusting search effort, jobseekers who receive negative news about their skills could also redirect their search effort to jobs that they believe are less desirable and hence less competitive. However, we find no evidence of this pattern. Treatment does not lead jobseekers to chose jobs in the job choice task that experts expect to offer lower salaries (Table D.4, columns 1–3), that control group jobseekers expect to offer lower salaries (columns 4–6), or that control group jobseekers expect to have lower offer probabilities (columns 7–9). Nor does treatment lead jobseekers to apply on SAYouth to jobs that are less likely to offer training or non-wage benefits (columns 10–15). Across all these outcomes, treatment effects are small (mostly $< 1\%$ of the control group means) and not statistically significant. These patterns hold for the average jobseeker (columns

Table D.1: Heterogeneous Treatment Effects on Search Effort Among Sample Without Clear Comparative Advantage - Tight Experiment

	Index		Planned apps (w)		Drafting time(w)		SMS click rate		# apps (w)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment	-0.138	-0.231	-3.137	-11.038	0.578	0.495	0.015	0.002	-7.512*	-7.625*
	(0.186)	(0.326)	(4.744)	(9.505)	(1.701)	(2.590)	(0.069)	(0.099)	(4.004)	(4.411)
Treatment × Avg. confidence (bl)		0.100		6.875		0.038		0.021		0.066
		(0.261)		(5.807)		(1.816)		(0.083)		(5.263)
Avg. confidence (bl)		0.089		-6.526		-0.404		0.090		-0.435
		(0.283)		(7.646)		(2.033)		(0.107)		(5.587)
Control mean	-0.000	-0.000	50.674	50.674	7.417	7.417	0.639	0.639	17.750	17.750
Observations	94	94	91	91	89	89	94	94	94	94

Notes: **Table D.1 shows that there are no significant treatment effects on prespecified search effort among jobseekers who don't have a clear CA in skill and hence learn less about their skill CA from treatment.** "Avg. confidence (bl)" measures jobseekers' baseline skill beliefs minus their assessment results, both relative to other jobseekers, with a detailed definition in footnote 41. Columns indicate different outcome variables: an inverse covariance-weighted average of the search effort measures used in all other columns (cols. 1–2), the number of planned applications in the 30 days after the workshop (cols. 3–4), the time spent drafting a cover letter during the workshop in minutes (cols. 5–6), the click rate for three text messages with links to job adverts we sent to jobseekers after the workshop (cols. 7–8), and the number of applications sent on the job search platform in the 30 days after the workshop (cols. 9–10). "(w)" denotes variables winsorized at the 99^{th} percentile. Control variables are listed in footnote 18. Standard errors clustered at the treatment-day level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

1, 4, 7, 10, 13), for jobseekers who receive negative news about their skill levels (columns 3, 6, 9, 12, 15), and for jobseekers with misaligned CA beliefs at baseline (columns 2, 5, 8, 11, 14). These results also illustrate the point made in Section 4.5 that treatment does not lead jobseekers to apply to jobs that offer better terms to all applicants, suggesting that skill-directed job search instead raises earnings by improving firm-worker match quality.

Treatment effects on labor market outcomes by baseline confidence: Treatment effects on labor market outcomes in the big experiment do not vary substantially by baseline confidence about skills. The interaction effects are 0.000 for employment (Table D.3, panel B, column 6, $p = 1$) and 0.1 on earnings (panel B, column 7, $p = 0.97$). Other effects shown in the table are similarly small and not statistically significant.

Conclusion: This analysis suggests that search effort is unaffected by treatment and hence is unlikely to explain the treatment effects on labor market outcomes. This might arise because the negative treatment effect on believed skill level produces offsetting substitution and income effects on search effort. This does not, of course, imply that search effort plays no role in determining labor market outcomes in this or other settings, as discussed in the review by Mueller & Spinnewijn (2023).

Table D.2: Heterogeneous Treatment Effects on Skill Beliefs and Search Effort by Baseline Confidence - Tight Experiment

	Belief	Main search effort				Platform search effort			
	Ave. skill quint. (0-4)	Main index	Planned # apps (w)	Drafting time (w)	SMS click rate	Platform index	Days active	Adverts clicked (w)	# apps (w)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Average effects									
Treatment	-0.082**	-0.077	-3.854	-0.530	0.003	-0.100	-0.723	-0.157	-1.430
	(0.035)	(0.103)	(2.555)	(0.591)	(0.032)	(0.093)	(0.477)	(1.525)	(1.553)
Panel B: Het. effects									
Treatment	0.044	-0.102	-3.574	-0.377	-0.055	-0.235	-1.451**	-2.784	-2.074
	(0.057)	(0.145)	(4.124)	(0.925)	(0.053)	(0.157)	(0.598)	(3.136)	(2.376)
Treatment ×	-0.122***	0.019	-0.429	-0.205	0.062	0.142	0.759	2.773	0.626
Average confidence (bl)	(0.039)	(0.111)	(2.936)	(0.771)	(0.045)	(0.121)	(0.449)	(2.560)	(2.068)
Average confidence (bl)	-0.220**	0.196	4.032	1.280	-0.059	-0.066	-0.278	-2.268	1.239
	(0.102)	(0.124)	(2.675)	(0.900)	(0.046)	(0.111)	(0.530)	(2.272)	(1.916)
Control mean	2.693	-0.000	37.878	8.828	0.635	-0.000	6.014	9.043	15.187
Observations	278	278	278	267	278	278	278	278	278

Notes: **Table D.2 shows that treatment effects on search effort in the tight experiment do not vary significantly by baseline confidence levels.** "Avg. confidence (bl)" measures jobseekers' baseline skill beliefs minus their assessment results, both relative to other jobseekers, with a detailed definition in footnote 41. Columns show different outcome variables: average belief about communication and numeracy skill quintiles from the tight experiment (col. 1), a prespecified inverse covariance-weighted average of the search effort measures used in cols. 3, 4, 5, and 9 (col. 2), the number of planned applications in the 30 days after the workshop (col. 3), the time spent drafting a cover letter during the workshop in minutes (col. 4), the click rate for three text messages with links to job adverts we sent to jobseekers after the workshop (col. 5), an inverse covariance-weighted average of the search effort measures in cols. 7 to 9 (col. 6), the number of days jobseekers were active on the platform in the 30 days following the workshop (col. 7), the number of job adverts jobseekers clicked on on the platform in the 30 days following the workshop (col. 8), and the number of applications sent on the job search platform in the 30 days after the workshop (cols. 9). "(w)" denotes variables winsorized at the 99^{th} percentile. Control variables are listed in footnote 18. Standard errors clustered at the treatment-day level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D.3: Heterogeneous Treatment Effects on Skill Beliefs, Search Effort, and Labor Market Outcomes by Baseline Confidence - Big Experiment

	Belief	Search effort				Labor market outcomes		
	Ave. skill					Worked	Earnings (w)	
	terc. (0-2)	Index	# apps (w)	Hours (w)	Expend.(w)	last 7d	uncond.	cond.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Average effects								
Treatment	-0.108***	0.003	0.062	-0.319	0.151	0.009	6.517**	20.393**
	(0.010)	(0.032)	(0.466)	(0.333)	(0.798)	(0.013)	(2.712)	(9.404)
Control mean	1.447	-0.000	11.682	11.151	21.446	0.309	25.424	85.826
Observations	4195	4205	4145	4086	3739	4204	4196	1280
Panel B: Het. effects								
Treatment	-0.007	-0.017	-0.042	-0.571	-0.355	0.010	6.461*	22.187*
	(0.011)	(0.038)	(0.560)	(0.524)	(0.839)	(0.017)	(3.456)	(12.669)
Treatment ×	-0.125***	0.030	0.158	0.321	0.770	-0.000	0.100	-1.425
Average confidence (bl)	(0.012)	(0.036)	(0.457)	(0.509)	(0.732)	(0.014)	(2.755)	(10.586)
Average confidence (bl)	-0.064***	-0.034	-0.090	-0.537	-0.274	-0.019	-4.044	-5.571
	(0.016)	(0.049)	(0.663)	(0.679)	(0.900)	(0.017)	(3.640)	(12.538)
Control mean	1.446	-0.000	11.716	11.083	20.878	0.309	25.424	85.826
Observations	4131	4131	4075	4017	3681	4130	4122	1248

Notes: **Table D.3 shows that treatment effects on search effort and labor market outcomes in the big experiment do not vary significantly by baseline confidence levels.** "Avg. confidence (bl)" measures jobseekers' baseline skill beliefs minus their assessment results, both relative to other jobseekers, with a detailed definition in footnote 41. Columns indicate different outcome variables: average belief about communication, numeracy, and concept formation skill terciles from the tight experiment (col. 1), an inverse covariance-weighted average of the search effort measures used in all other columns (col. 2), the number of applications in the 30 days before the endline survey (col. 3), the number of hours spent searching for jobs in the same 30 days (col. 4), job search expenditure in the same 30 days (col. 5), a dummy indicating any work for pay in the seven days before the endline survey (col. 6), unconditional earnings in the last seven days (col. 7), earnings in the last seven days conditional on working (col. 8). "(w)" denotes variables winsorized at the 99^{th} percentile. Control variables are described in footnote 25. All monetary values are measured in 2021 USD in purchasing power parity terms. Standard errors clustered at the treatment-day level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table D.4: Heterogeneous Treatment Effects on Applications to Jobs with Better Attributes - Tight Experiment

	Average salary of job chosen in job choice task						Perceived offer prob.			Share of applied to jobs					
	Experts' assessment			Control group's assessment			Control group's assessment			Offering training			Offering amenities		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Treatment	2.583	1.626	3.655	-0.931	0.754	-0.976	0.001	-0.001	0.001	0.003	0.002	-0.039	-0.002	-0.001	-0.003
	(2.228)	(2.886)	(3.251)	(3.791)	(6.908)	(5.968)	(0.003)	(0.004)	(0.004)	(0.025)	(0.040)	(0.040)	(0.002)	(0.003)	(0.003)
Treatment × Aligned CA belief (bl)		1.998			-2.468			0.002			0.005			-0.002	
		(3.315)			(10.134)			(0.007)			(0.049)			(0.004)	
Aligned CA belief (bl)		-4.350			-10.659			0.010			-0.033			0.004	
		(2.584)			(7.905)			(0.007)			(0.046)			(0.004)	
Treatment × Average confidence (bl)			-1.008			-0.068			-0.000			0.044			0.001
			(2.242)			(4.804)			(0.003)			(0.030)			(0.003)
Average confidence (bl)			-2.495			3.226			-0.002			-0.027			0.001
			(2.248)			(5.789)			(0.004)			(0.034)			(0.003)
Control mean	661.434	661.434	661.434	845.472	845.472	845.472	0.549	0.549	0.549	0.206	0.206	0.206	0.006	0.006	0.006
Observations	278	278	278	278	278	278	278	278	278	278	278	278	278	278	278

Notes: **Table D.4 shows that treatment does not lead jobseekers to apply to jobs with better or worse attributes, even for jobseekers who receive more negative information about their skills from treatment.** "CA" stands for comparative advantage and "Avg. confidence (bl)" measures jobseekers' baseline skill beliefs minus their assessment results, both relative to other jobseekers, with a detailed definition in footnote 41. Columns show different outcomes. Cols 1–9 show attributes of the jobs chosen in the job choice task: the average monthly salary in 2021 USD in purchasing power parity terms assessed by the hiring experts (cols. 1–3) or by jobseekers in the control group (cols. 4–6), the perceived job offer probability assessed by jobseekers in the control group (cols. 7–9). Cols 10–15 show attributes of jobs applied for on the SAYouth.mobi platform: the share of applied for jobs that mention the offer of training (cols. 10–12) and amenities (pension, medical, and unemployment insurance; cols. 13–15) in the job descriptions. Control variables are listed in footnote 18. Standard errors clustered at the treatment-day level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.