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

Andrea Kiss ⓡ Robert Garlick ⓡ Kate Orkin ⓡ Lukas Hensel

September 11, 2023

Abstract

Worker sorting into tasks and occupations has long been recognized as an important feature of labor markets. This sorting may be inefficient if jobseekers have inaccurate beliefs about their skills and therefore apply to jobs that don't match their skills, behavior we show is common for young South African jobseekers. We run two field experiments that give jobseekers their results from standardized assessments of job-relevant skills. This redirects their search toward jobs that value skills where they score relatively highly without raising their search effort, using measures from administrative, incentivized task, and survey data. This also raises earnings and job quality, consistent with inefficient sorting due to limited information.

*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, Vittorio Bassi, Michele Belot, Stefano Caria, Clement Imbert, David Huffman, Jeremy Magruder, Barbara Petrongolo, Chris Roth, Duncan Thomas, and Basit Zafar, as well as seminar and conference participants at AFE, Bocconi, CEPR, Carnegie Mellon, Columbia, Duke, Essex, FIU, IFLAME, NBER, Oxford, Peking, PHBS, Renmin, Pittsburgh, and Royal Holloway. 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, Maryana Iskander, Khodani Nematei, Laura Poswell, Neil Rankin, Mosoue Sekonyela, and Rob Urquhart. This study has been approved by 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://doi.org/10.1257/rct.1631-8.0 and https://doi.org/10.1257/rct.10000-1.0. This paper is 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.

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 job types. However, jobseekers may have imperfect information about how well their skills match what different job types require. Research in both developed and developing economies has documented imperfect information about other aspects of job search but has not studied jobseekers' beliefs about their skill match with different job types (Abebe et al., 2022; Abel et al., 2020; Bandiera et al., 2023; Beam, 2016; Belot et al., 2019, 2022b; Boudreau et al., 2022; Conlon et al., 2018; Cortés et al., 2023; Jones & Santos, 2022; Kelley et al., 2020; Wiswall & Zafar, 2015).

This paper shows that jobseekers can have imperfect information about their skills, distorting how they direct job search, and leading to worse labor market outcomes. We formalize this idea in a simple model of directed job search. Jobseekers have multidimensional skills and beliefs about their **skill comparative advantage**: their relative rank in different skill dimensions, compared to other jobseekers from similar backgrounds. These beliefs influence their **skill-directed job search**: the way they allocate effort across searching for jobs with different skill demands. In this model, imperfect information about skill comparative advantage can distort skill-directed job search.

We provide causal evidence from two field experiments to support this idea. We work with a job search assistance agency in Johannesburg, South Africa. The agency helps young jobseekers, from disadvantaged backgrounds, mostly with only secondary education, who are searching mainly for non-specialist, entry-level, service-sector jobs. We invite jobseekers to day-long job search assistance workshops run with the agency. In the workshops, we measure jobseekers' skills using established psychometric assessments of communication and numeracy – two general-purpose, job-relevant skills that we verify firms in this context value in prospective workers. We also measure jobseekers' beliefs about their levels of these skills. Many jobseekers' beliefs about their skill comparative advantage do not match their assessment results, these beliefs persist over time, and the beliefs predict which jobs they apply for.[1] These descriptive patterns are consistent with our model and help to motivate our experimental analysis.

In both experiments, we randomly assign half of the workshops to give participants

[1]We define each jobseeker's comparative advantage as the skill dimension in which they rank highest relative to other jobseekers from similar backgrounds who take the assessments. This follows the few job search studies using multidimensional skill measures (e.g. Guvenen et al. 2020) and common practice in education research (Altonji et al., 2016). We discuss trade-offs relative to other definitions later in the paper.

information about their communication and numeracy assessment results, both expressed as rankings relative to a large group of jobseekers from similar backgrounds. We interpret the treatments as providing only jobseeker-facing information because it is difficult for participants to credibly share information with firms and we show that sharing is rare.

Our first experiment (N = 278) studies skill-directed job search using unusually rich survey, task, and administrative data on beliefs and search. Treatment shifts participants' skill beliefs substantially closer to their measured skills, including their beliefs about their comparative advantage. Treated participants are also more likely to apply to jobs whose skill demands match their comparative advantage, using multiple prespecified measures of skill-directed job search: a novel incentivized task in which participants choose between applying to jobs with different skill requirements, applications on a job search platform, and survey data.[2]

Our second experiment (N = 4,389) studies labor market outcomes using survey data for a larger sample collected on average 3.5 months after the workshops. Treatment increases jobseekers' weekly earnings and hourly wages by respectively 6.5 and 0.3 USD PPP (both roughly 25% of the control group means) and increases the probability of formal employment. But we find at most weak evidence for a higher employment rate. Treatment effects on jobseekers' comparative advantage beliefs and skill-directed job search are qualitatively similar to the first experiment, although using different measures.

The two experiments show a consistent picture of this labor market: getting more information about their skills shifts jobseekers' perceived comparative advantage toward their measured comparative advantage, redirects their search toward jobs aligned with their measured comparative advantage, and leads to higher-quality employment. In contrast, we find limited evidence for a plausible alternative economic mechanism: that treatment shifts beliefs about skill levels and therefore shifts search effort and labor market outcomes. We do see the first part of this mechanism in our data: treatment lowers the average jobseeker's belief about their skill level, because untreated jobseekers are on average overconfident about their skill level 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 in both experiments. This is consistent with our model, which shows that the effect of skill level beliefs on search effort has a theoretically ambiguous sign.

The two experiments provide complementary types of evidence, from the same location, with participants recruited in the same way, several years apart. Given the two experiments' different strengths, we call them the "tight" and "big" experiments respec-

[2] The job search task provides a novel way to use job application choices to measure jobseekers' valuation of skill match relative to other vacancy characteristics. This builds on work using choice experiments to study job or education choices (Adams-Prassl & Andrew, 2023; Mas & Pallais, 2017; Wiswall & Zafar, 2015).

tively. We use this light-hearted, non-standard terminology because standard taxonomies do not capture the differences between the two experiments.[3] Together, the two experiments give us exogenous variation in jobseekers' beliefs about their skills and the unique combination of data needed for this research: jobseeker-level data on multidimensional skills, beliefs about these skills, and job search activities and outcomes; vacancy-level data on skill demands; and jobseeker*vacancy-level data on application decisions.[4] This adds to a small, recent literature showing the value of combining multiple research designs and data sources to understand both a job search process and the outcomes of that process (Carranza et al., 2022; Cortés et al., 2023; Field et al., 2023).

More generally, what labor market conditions might lead to inaccurate comparative advantage beliefs and hence to poorly skill-directed job search? Our participants enter the labor market with limited information about their comparative advantage because schools give noisy feedback on their skills (Lam et al., 2011). The same is true in many education systems (Pritchett, 2013). Our participants learn slowly about their comparative advantage because high unemployment limits learning through work experience, the standard mechanism in many models (e.g. Guvenen et al. 2020). Similarly, high and persistent youth unemployment has become common in many countries (ILO, 2022). More generally, jobseekers might struggle to evaluate their fit with different jobs when navigating new or rapidly changing labor markets due to industrial displacement, migration, or structural transformation (Huckfeldt, 2022; Robinson, 2018). This can be particularly important when search is costly for jobseekers or screening mismatched applicants is costly to firms (Abebe et al., 2021a; Algan et al., 2022; Fernando et al., 2023; Hensel et al., 2022).

This paper provides the first direct evidence that jobseekers' beliefs about their comparative advantage in skills influence how they direct search to different job types and therefore influence their labor market outcomes. This contributes to two literatures. First, we extend research on the relationships between jobseekers' beliefs, search activities, and search outcomes, reviewed by Mueller & Spinnewijn (2022).[5] This literature focuses on

[3]Both are field rather than lab or lab-in-the-field experiments because they feature real jobseekers engaging in real job search activities. Both have elements of what Harrison & List (2004) call framed field experiments and natural field experiments.

[4]No existing data sources provide all this information. 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.

[5]This includes research into the co-evolution of search and search-related beliefs in panel data (Adams-Prassl et al., 2023; Conlon et al., 2018; Mueller et al., 2021; Spinnewijn, 2015); experiments providing information about labor market conditions (Altmann et al., 2018; Jones & Santos, 2022); experiments on search subsidies, matching services, or mentoring programs that influence multiple outcomes including jobseekers' beliefs about their labor market prospects (Abebe et al., 2022; Alfonsi et al., 2022; Bandiera et al., 2023; Banerjee & Sequeira, 2023; Beam, 2016; Kelley et al., 2020; Wheeler et al., 2022); and research on jobseek-

jobseekers' beliefs about the level of their labor market prospects, captured by job offer arrival rates or wage offer distributions. Despite the centrality of sorting on skill comparative advantage in labor economics (Roy, 1951), the literature on belief-based job search has not studied beliefs about skills or skill comparative advantage. And it rarely studies how search is directed across different job types or beliefs about the outcomes from searching for different job types. Our experimental design is most similar to recent papers showing that encouraging jobseekers to apply to different occupations can redirect job search and improve some labor market outcomes (Altmann et al., 2022; Belot et al., 2019, 2022a). We complement this work by focusing on skills rather than occupations and directly measuring the beliefs that link these treatments to search decisions. We also show how to directly measure the entire causal chain – skills, skill beliefs, skill-directed job search, and search outcomes – while working in a setting without government administrative data.

Second, our work complements research on skill (mis)match, reviewed by Sanders et al. (2012). This literature includes work documenting relationships between worker*job skill match and wages, as well as work quantifying the long-term earnings losses due to workers searching with imperfect information about their skill match with different jobs (e.g. Baley et al. 2022; Fredriksson et al. 2018; Guvenen et al. 2020; Lise & Postel-Vinay 2020). These papers do not observe beliefs or search, instead inferring these processes from dynamic search and matching models fitted to rich longitudinal employment data. Our measures of skills, beliefs about skills, search, and short-term labor market outcomes, combined with experimental variation in beliefs, allow us to directly observe behavior assumed by these models. Within this literature, our findings are most consistent with the modeling framework of Baley et al. (2022), which features jobseekers directing search based on inaccurate beliefs about their multidimensional skills. The limited information and misdirected search we document might help to explain cross-country variation in job turnover rates, which Donovan et al. (2022) link to variation in matching frictions.

A related literature focuses on firm-side limited information about jobseekers' skills.[6] Within this literature, our work is most closely related to our companion paper, Carranza

ers' beliefs about attributes of specific jobs (e.g. Bazzi et al. 2021; Boudreau et al. 2022; Chakravorty et al. 2023; Sockin & Sojourner 2023; Subramanian 2022). The formation and consequences of beliefs about skills have also been studied outside of job search: decisions in the workplace (e.g. Hoffman & Burks 2020; Huffman et al. 2022; Malmendier & Tate 2015), education investment (e.g. Berry et al. 2022; Bobba & Frisancho 2020), and in lab settings, including some studies of tasks that mimic job search (e.g. Falk et al. 2006). Skill comparative advantage beliefs have not been studied outside education settings Altonji et al. (2016).

[6] 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, 2017; Kahn & Lange, 2014), references and referrals (Abel et al., 2020; Heath, 2018; Ioannides & Loury, 2004; Pallais, 2014), education qualifications (Alfonsi et al., 2017; 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).

et al. (2022). That paper uses multiple firm-facing experiments to study how firms react to new information about jobseekers' skills. It also reports the labor market outcomes from our big experiment, as part of an argument that giving information about jobseekers' skills to jobseekers alone versus jobseekers and firms produces different outcomes. Unlike the companion paper, we focus deeply on jobseeker-side limited information. We add the tight experiment, with its detailed measures of skill-directed job search and comparative advantage beliefs, and add a model of belief-based skill-directed job search and additional mechanism measures from the big experiment not used in Carranza et al. (2022). Combining the tight and big experiments allows us to study the full causal chain from comparative advantage beliefs, through skill-directed job search, to labor market outcomes, which the companion paper does not do.

Our findings also relate to extensive work on active labor market programs in both developed and developing economies, reviewed by Card et al. (2018) and McKenzie (2017). The average earnings gain accumulated from treatment to follow-up is roughly 1.8 times the average variable cost of our intervention. This suggests similar interventions can be cost-effective active labor market programs, potentially run through job search platforms.

Section 2 of the paper describes our model, context, and patterns of skill, skill beliefs, and job search in our sample. In Section 3 we show the relationship between comparative advantage beliefs and skill-directed job search in the tight experiment. In Section 4 we show the relationship between comparative advantage 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 discuss additional possible mechanisms in Section 6.

2 Economic Environment

We begin with a conceptual framework to define skill comparative advantage and how limited information about jobseekers' skill comparative advantage can distort skill-directed job search and worsen labor market outcomes. We then describe our context, sample, and skill assessments. We then report four key descriptive patterns that inform our conceptual framework: jobseekers' skills vary across 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 suggest that jobseekers have different skills, these skills are valued in different jobs, but limited information about their own skills can distort their search decisions.

In the conceptual framework and tight experiment, we study jobseekers' beliefs about comparative advantage over communication and numeracy and their search over jobs demanding these skills. Two skill dimensions allow simplicity and provide the minimum needed to study comparative advantage and skill-directed search. These skills suit our research question: they are general-purpose 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 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 has a similar spirit to recent models of "partially directed job search," where jobseekers try to direct search to highest-wage vacancies but face uncertainty about wages (Lentz et al., 2022; Wu, 2021). We use a partial equilibrium framework focusing on jobseekers' search and beliefs, treating labor demand and wage posting as fixed, because we observe limited firm-side data. We do not model dynamic belief updating in response to search outcomes, because we show in Section 2.6 the jobseekers we study learn little about their skills during search over the timeframe of our study.

Skills and comparative advantage: In our framework, each jobseeker has communication and numeracy skill levels S_C and S_N. In our empirical work, we rank each jobseeker's communication and numeracy skills using their scores on standardized assessments relative to a large group of jobseekers from similar backgrounds. Many education systems generate similar rankings. We focus on ranks because absolute scores are sensitive to the way assessments are scaled (Nielsen, 2023).

We classify a jobseeker as having a **comparative advantage in communication** if they rank higher in the distribution of communication than numeracy skills and vice versa. This definition aligns with the spirit of standard definitions of comparative advantage in trade. There, a country has a comparative advantage in the product it can produce at lowest opportunity cost. Here, a jobseeker has a comparative advantage in the skill where she ranks higher, because she can supply it to the market at a lower opportunity cost in terms of the other skill.[7] While other definitions are possible, our definition takes advantage of our multidimensional skill measures and follows common practice when researchers have access to similar data in education (reviewed by Altonji et al. 2016) and, more rarely, in labor research (e.g. Guvenen et al. 2020).[8]

[7] For analytical simplicity, we assume the two skills have the same distributions, so a jobseeker has higher communication than numeracy skill iff she has a higher rank in the communication than numeracy skill distribution and vice versa.

[8] Our definition approximates each jobseeker's rank against other jobseekers likely to be competing for similar jobs, although we do not observe the jobseeker's rank in the skill distribution of applicants for any specific job. Alternative approaches estimate occupation-specific wages by education and use this to define comparative advantage in occupations based on education level (e.g. Acemoglu & Autor 2011; Gibbons et al. 2005). These approaches can price the value of different types of workers in different types of jobs, which we do not. But their wage estimates are conditional on the way workers currently sort into

Search over jobs demanding different skills: Each job demands primarily communication or primarily numeracy skills. Jobseekers split fixed total search effort $\bar{E}$ between search for communication jobs E_C and numeracy jobs E_N. Searching for a type j job yields outcome $V_j(S_C, S_N, E_j)$, which is a reduced-form expression for the expected present value of a job offer scaled by the probability of an offer. We make three assumptions. First, that V_j is increasing and concave in all three arguments and that $\partial V_j / \partial S_j > \partial V_j / \partial S_i > 0$ 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 the gross utility from job search $U(V_C, V_N)$ is increasing and concave in both arguments, which allows jobseekers to value the outcomes of both types of search, without modeling 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 non-pecuniary 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 (1), $\frac{\partial V_C}{\partial E_C}$ 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 (3) will rise more than the right-hand side. The jobseeker will then increase E_C and decrease E_N to restore equality in condition (3), under the 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 $\tilde{S}_N$. This framework delivers two testable predictions. First, jobseekers will direct more search effort to jobs requiring the skill in which they believe they have a comparative advantage to satisfy condition (2), irrespective of their actual comparative advantage. Second, jobseekers whose believed comparative advantage does not match their actual comparative advantage will have worse job search outcomes.

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 we study.

2.2 Context and Target Population

We work in Johannesburg, part of South Africa's commercial and industrial hub of Gauteng, a metropolitan area of 14 million people. 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 (Statistics South Africa, 2022). This matches patterns in large cities in other middle-income African countries (Bandiera et al., 2022a).

Our target population is young, active jobseekers with at least high school education who attended school in low-income areas. We recruit from participant registries at the Harambee Youth Employment Accelerator. Harambee is a social enterprise funded jointly by the South African government and private firms which provides job search assistance to young jobseekers from low-income backgrounds. Since 2013, Harambee has maintained a database of active jobseekers recruited through traditional and social media. Jobseekers sign up online with their national identity number, which Harambee uses to determine that they are aged 18-34 with legal permission to work. They self-report if they are actively searching for work and attended school in a low-income area. Firms receive free access to the database for recruiting. The database captures a sizeable proportion of the population of interest: in Gauteng in 2022 there were 1,078,745 jobseekers on the database and 1,403,064 unemployed people, both restricted to ages 18-34 (Statistics South Africa, 2022). These groups do not perfectly overlap as some jobseekers on the database are employed and some unemployed people are not on the database.

We select this context and population because they are relevant for studying limited information about skills, not because they are nationally or globally representative. These jobseekers are actively engaged in the labor market but face an economic environment that provides limited information about their skill levels relative to other jobseekers, for two reasons. First, jobseekers in low-income schools receive relatively little information about their skills from assessments during school: high school grades and grade progression are weakly correlated with results on independent skill assessments (Lam et al., 2011; Schoer et al., 2010). Such schools rarely have career counselors and few teachers are trained to offer career guidance (Pillay, 2020). Second, unemployment is high: 40.5% of the cohort aged 15-34 in Johannesburg at the time of the tight experiment were unemployed (Statistics South Africa, 2022).[9] Many young jobseekers have no work experience several years after completing education. This limits their scope to learn about their skills through work experience, the main mechanism in models of worker-side learning in other

[9] We use Statistics South Africa's definition of an employed person as someone who did any income-generating activity, for at least one hour, during the reference week. Unemployment rates exclude those in full-time education or not in the labor force.

contexts (Baley et al., 2022; Guvenen et al., 2020). These patterns are not unique to South Africa, as we explain on page 4.

2.3 Sample Description & Skill Assessments

Recruitment: We recruit one sample of 4389 people for our big experiment between September 2016 and April 2017 and one sample of 278 people between July and October 2022 for our tight experiment, using the same sampling frame and method. We contact people from Harambee's database who live within commuting distance of our field location in downtown Johannesburg. We screen out those not actively searching for work and invite the rest to a day-long job search assistance workshop. We state the workshop will include taking assessments that could be used to match them to suitable vacancies; answering questions about their job search; and receiving job search advice. Harambee often runs similar workshops.

Data: In both experiments, participants complete self-administered baseline surveys about their demographics, beliefs about their skills and job search, recent job search activities, current employment, and employment history before treatment. 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 their database of participants to set up this platform in 2019. The platform aggregates job advertisements from all online job boards which jobseekers can access without charges for mobile phone data. Jobseekers in our sample describe it as a key part of their job search strategy. It also 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. We show corresponding statistics for the big experiment in Section 4.1. Jobseekers are young, with interdecile age range 21 - 32. Most, 60%, have only general secondary education, where there is little specialization by subject, limiting their scope to learn about their skills from specialized training. The remainder have either short diplomas or tertiary qualifications. The majority, 67%, are unemployed, a higher rate than for this age group in Johannesburg as a whole as our sample excludes the economically inactive. They have limited, mostly informal work experience so have had limited opportunity to learn about their skills from employment: 33% were employed at baseline but only 13% had a formal written contract and only 25% had *ever* held a long-term wage job.

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 applications, spent 14 hours searching, and spent 22.72 USD on search, including transport to

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. Winsorized variables are winsorized at the 99^{th} percentile. All monetary values are in 2021 USD in purchasing power parity terms.

drop off CVs and attend interviews, mobile phone airtime and data, and printing. This search cost is high: 50% of the average weekly earnings of employed people in our sample, matching other work documenting expensive job search 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. The average jobseeker received only 0.17 job offers in the last 30 days, implying that $<1\%$ of applications yield offers. (We used a longer recall period for offers than applications because offers are rare.)

Skill assessments: We focus on two general-purpose, job-relevant skills: communication and numeracy. The numeracy assessment captures practical arithmetic and pattern recognition. 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 and was developed by an adult education provider (www.mediaworks.co.za). Candidates also complete a measure of "concept formation" (Taylor, 1994), a local measure of fluid intelligence or ability to see underlying commonalities across situations and to use logic in new situations (Raven & Raven, 2003). These assessments are also used in the big experiment, along with four additional assessments described in Section 4.2. We describe the assessment process and the psychometric properties of assessments in Appendix C. There are no ceiling or floor effects (Figure A1).

All our information treatments and measures of skill beliefs use assessment results relative to a benchmark population: roughly 12,000 jobseekers from Harambee's database who have been assessed. We place jobseekers in quintiles relative to this population.

2.4 Jobseekers' Performance Varies Across Different Skill Types

These two assessments differentiate jobseekers horizontally more than vertically: they show which jobseekers are better suited for jobs where communication versus numeracy skills are more important, rather than identifying jobseekers who are likely to be better at most jobs. In the full tight experiment sample, communication and numeracy scores have correlation 0.31 and only 17% of jobseekers are in the top two quintiles for both skills. Communication and numeracy scores are both moderately correlated with concept formation ($\rho \approx 0.25$ for both skills), suggesting neither assessment captures ability better.

In the tight experiment, we restrict our sample to 278 jobseekers with a clear comparative advantage in skills i.e. we drop the 24% of jobseekers in the full sample who have the same quintile for both skills. Our experimental measures of skill-directed job search, described in Section 3.4, can only be sensibly defined for jobseekers with a unique skill comparative advantage. In this sample, communication and numeracy skill quintiles have zero correlation. 62% of our restricted sample have a comparative advantage in communication and 38% have a comparative advantage in numeracy (Table A3).[10]

2.5 Firms Value Different Skills for Different Jobs

We have already shown that the assessments mainly identify jobseekers' relative strengths across different skills. Here we show that firms value the communication and numeracy skills we measure and that there is variation in which firms value which skills. These patterns suggest scope for jobseekers to improve their labor market outcomes by searching for jobs that value their skills where they score highly.

Firms value communication and numeracy skills relative to education qualifications: 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. See Appendix B.2 for details. Over 500 client firms have paid Harambee to screen

[10] The shares of jobseekers in this sample with communication and numeracy comparative advantage do not need to be equal because we compare jobseekers' skills to a separate benchmark population.

roughly 1 million jobseekers using these assessments, which we interpret as revealing a preference for using these skills in hiring. And Carranza et al. (2022) use multiple field experiments to show that jobseekers who can certify their assessment results have better labor market outcomes. We acknowledge employers may also value other skills in this labor market; we simply argue that communication and numeracy skills are valued.

These results are consistent with the school-leaving certificate not providing firms sufficient with information on applicants' skills. Communication and numeracy skill quintiles are positively but weakly associated with jobseekers' self-reported grades on their high school leaving exams in English and mathematics, respectively (Table A10, columns 1-2). The association suggests the Harambee assessments capture meaningful variation in skills; its weak nature highlights that jobseekers and firms both have scope to learn from information on jobseekers' rank 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 labor market outcomes by searching for different types of jobs.

Firms can at least partly observe skills. We conduct a measurement exercise embedded in a firm's hiring process, described in Appendix B.3. We find that the firm's HR team can identify levels of jobseekers' skill based on typical application materials. The HR team is also significantly more likely to invite applicants to interviews for positions with skill requirements that match jobseekers' assessed comparative advantage, compared to positions that do not match their comparative advantage. Together, these patterns suggest that redirecting jobseekers' search towards jobs that match their comparative advantage in skills has the potential to improve their labor market outcomes.

2.6 Jobseekers' Perceived and Measured Comparative Advantage in Skills Differs

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 don't 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 C.1 for details of the measurement and Figure A2 for the ordering of belief measurement, assessments, and treatment.

We do not refer to beliefs about skills as "accurate" or "inaccurate," as any assessment may measure skills with error, even the well-established assessments we use. Instead, we refer to beliefs about skills as "aligned" or "misaligned" with skill assessment results.

Jobseekers have misaligned beliefs about their comparative advantage: A jobseeker has an **aligned comparative advantage belief** about their comparative advantage over numeracy and communication if they have a higher perceived quintile for the skill in which they have a higher assessed quintile. Only 49% of jobseekers have aligned comparative advantage beliefs at baseline (Table 1, panel D).[11]

The measure above captures beliefs about skills *in general*. Jobseekers' beliefs about their comparative advantage *on our specific assessments* are similarly misaligned. We ask jobseekers which quintile they fall in on our assessments, after taking the assessments but before any jobseekers receive information about their skills. Only 54% of jobseekers' comparative advantage beliefs about their assessment results match the actual results.[12]

Jobseekers have misaligned beliefs about their skill levels: We define the **fraction of aligned beliefs** as the average of two indicators, one for communication and one for numeracy, each equal to one if perceived and assessed skill quintiles are equal. The measure can take values of 0, 0.5 and 1. Using this measure, 22% of skill beliefs are aligned (Table 1, panel D).[13] The 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. This pattern is relatively common (Santos-Pinto & de la Rosa, 2020).[14]

Jobseekers learn little about their skills while searching: To show this, we use the big experiment's follow-up survey, 3.5 months after baseline. The share of control group jobseekers whose believed and assessed comparative advantage align hardly changes over 3.5 months of searching, even for those with above-median search effort and the employed (Table A12, columns 1-3). The fraction of aligned beliefs is similarly persistent

[11]If jobseekers were randomly guessing, 40% would have aligned comparative advantage beliefs. To see this, note that there are 5x5 possible beliefs over communication and numeracy skill quintiles. The 5 possible beliefs with tied skill quintiles are misaligned because our sample excludes anyone with tied assessment results. 10 of the other 20 possible skill beliefs lead to misaligned comparative advantage beliefs. Hence random guessing leads to aligned comparative advantage beliefs in (5x0 + 20*0.5)/25 = 40% of cases.

[12]The two belief measures are strongly associated, suggesting jobseekers view the assessments as relevant to their general skills. Regressing general skill beliefs on beliefs about assessment results for the same skill produces coefficients of 0.39-0.52, controlling for assessment results, demographics, and education.

[13]Comparative advantage beliefs and the fraction of aligned beliefs are positively correlated by construction but have some 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 comparative advantage.

[14]Some researchers call this pattern, in which people think they are performing better than others, "overplacement" (Moore & Healy, 2008). We call this pattern overconfidence, except where otherwise noted.

(columns 4-6). This is consistent with high search costs limiting scope for learning from search. Jobseekers also receive little feedback: only 3% of jobseekers report ever receiving feedback about their skills during an unsuccessful job application. Slow learning may also reflect the well-documented difficulty of Bayesian learning of a function with multiple inputs, such as the search outcome-skill relationship (Banerjee & Sequeira, 2023).

Jobseekers seem to draw on both school results and other information in forming skill beliefs: Even with good information from the schooling system, exam scores should not perfectly predict young adults' beliefs about their skills: many jobseekers took school exams multiple years ago, the exam and assessments do not test identical skills, and jobseekers may not perfectly recall their exam scores. Indeed, skill beliefs about communication and numeracy are positively but weakly correlated with jobseekers' self-reports of their results on the high school graduation exams in English and mathematics (Table A10, columns 4-5). Beliefs about comparative advantage 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 (Appendix J, Tables A45 and A46). This matches a recent metastudy showing limited gender heterogeneity in confidence (Bandiera et al., 2022b). This motivates our gender-pooled analysis in this section.

2.7 (Misaligned) Beliefs Predict Skill-Directed Job Search

We have already shown 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' skill beliefs and assessment results might shift their skill-directed job search.

Comparative advantage beliefs predict job search decisions: In the big experiment, we ask candidates what skill is most valuable for the types of jobs they are applying for. These answers are strongly associated with individuals' beliefs about their comparative advantage in skills in the control group. Jobseekers are 9-10pp ($p < 0.01$) more likely to state that they are applying for jobs that value the skill in which they *believe* they have a comparative advantage, compared to jobs valuing other skills (Table A15, rows 1-2). The correlations are robust across skill domains and to controlling for measured comparative advantage and demographics. These results suggest jobseekers try to search for jobs where they have a comparative advantage in skills.

Assessed comparative advantage weakly predicts job search decisions: Jobseekers are only 2-4pp more likely to state that they are applying for jobs that value the skill in which they *scored higher on our assessments*, compared to jobs valuing other skills (Table A15, rows 3-4). The results in this and the preceding paragraph suggest that jobseekers' skill beliefs might direct their search away from jobs that match their assessed skills.

Jobseekers also believe having higher skills improves search outcomes. In the big experiment, we asked jobseekers' 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 PPP more (13% of the mean, $p < 0.001$). We did not ask these questions about communication skills to save survey time. Regressing expected search outcomes on skill self-beliefs produces a similar pattern, although this between-jobseeker comparison might reflect omitted variable bias (Table A14). The result is robust to controlling for assessment results, demographics, education, and work history.

3 Tight Experiment: Effects on Beliefs and Directed Search

The previous section's conceptual framework and descriptive evidence suggests 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 respond to new information about their skills. This experiment collects rich data on beliefs and unique measures of skill-directed job search using jobseekers' choices between jobs with different skill content at the jobseeker$\times$job level. As discussed on page 4, existing datasets do not provide this type of information.

3.1 Experimental Design and Intervention

We run the experiment during day-long job search workshops. We randomize treatment at the workshop level to avoid spillovers between jobseekers. We assign 17 workshops to treatment and 17 to control. 373 jobseekers attend the workshops, of whom 278 have a unique comparative advantage and constitute our main sample. Treatment assignments are balanced on baseline covariates, for both the 373 and the 278 jobseekers (Table A16). Jobseekers know they are participating in a research study but not an experiment, so their behavior and survey responses should not reflect experimenter demand effects. All jobseekers receive reimbursement for transport and the time spent at the workshop.

The timeline of the day is shown in Figure A2. Jobseekers first complete a pre-treatment survey, in which we define communication and numeracy skills and the concept of quintiles. We ask jobseekers their beliefs about which quintile their general communication and numeracy skills are in, relative to the reference group. They then take assessments of their numeracy, communication, and concept formation skills and complete a brief survey about their perceived performance on the assessments.

Treated jobseekers then receive a report describing the assessments and their performance (Figure 1). For each skill, the report shows the quintile in which the jobseeker ranked on each assessment, compared to other jobseekers in the reference group. Treated jobseekers watch a video that explains the skill assessments and how to interpret the report, particularly the quintiles. The video encourages them to think about what jobs will value their skills but does not encourage applying to any specific types of jobs.

Jobseekers in the control group do not receive a report. They watch a control video, which is a strict subset of the treatment video. It contains the parts of the treatment video that explain the assessments to hold constant any effect of having better information about communication or numeracy skills. It also contains the encouragement to jobseekers to think about what jobs will value their skills to hold constant the general idea of skill-directed job search. Both groups take the same assessments and answer the same skill-focused survey questions, so any priming effects about the importance of skill levels and skill match are held constant across treatment groups.

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 report the quintiles they scored for respectively numeracy and communication, and 98% correctly report the skill in which they scored higher. Belief updating does not differ by conceptual formation scores, suggesting general cognitive ability does not limit processing this information.

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.

Appendix D contains a detailed description of the workshops and links to the videos.

3.2 Specification

We estimate effects of receiving information about one's relative ranking on skills:

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

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.[15] We use heteroskedasticity-robust standard errors clustered by

[15] $\mathbf{X}_{id}$ 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 both numeracy and communication skills; a pre-treatment value of the outcome Y_{id} where available; and

Figure 1: Sample Report

REPORT ON CANDIDATE COMPETENCIES
-Personal Copy-

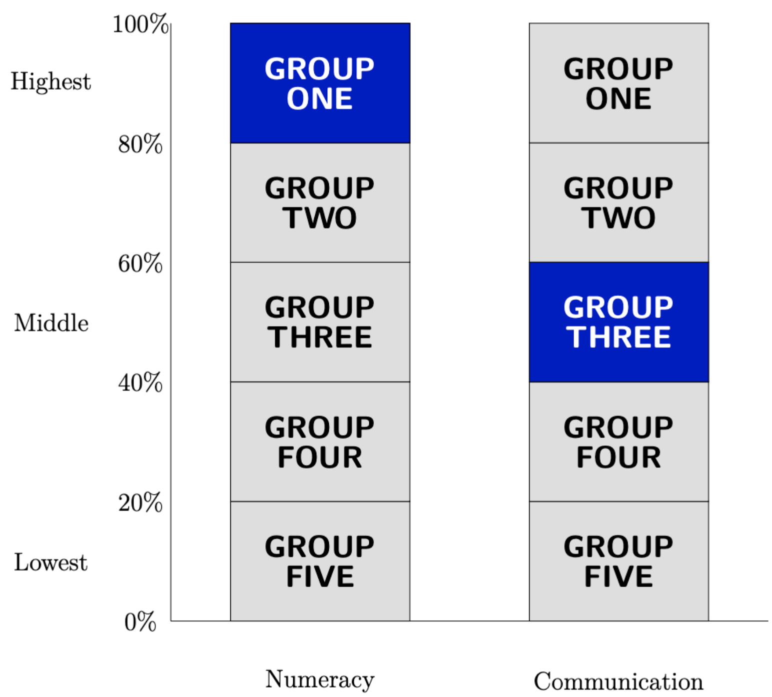
This report contains results from the assessments you took today. These results can help you learn about some of your strengths and weaknesses and inform your job search.

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: **Figure 1 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.

workshop date, the unit of treatment assignment.

We also test if treatment effects are larger for jobseekers whose baseline beliefs are misaligned with their comparative advantage. We estimate models of the form:

$$Y_{id} = T_d \cdot \alpha^{misaligned} + T_d \cdot Aligned_{id} \cdot \alpha^{diff} + Aligned_{id} \cdot \delta + \mathbf{X}_{id} \cdot \Gamma + \epsilon_{id}. \quad (4)$$

$Aligned_{id}$ is an indicator for jobseekers whose pre-treatment beliefs about their comparative advantage on the assessments match their assessment results (measurement details in Appendix C.1). These jobseekers receive less information from treatment. We report the average treatment effect for jobseekers with misaligned baseline comparative advantage beliefs, $\alpha^{misaligned}$; the average treatment effect for jobseekers with aligned baseline comparative advantage beliefs, $\alpha^{misaligned} + \alpha^{diff}$; and the difference between them, α^{diff}.

For both specifications, we focus our analysis on the 278 individuals with a clear comparative advantage in one of the skills, 139 from treatment days and 139 from control days. This is necessary to cleanly define beliefs about comparative advantage and job choices aligned with comparative advantage in skills, as discussed on page 12. However, including these jobseekers in the sample and classifying all their job search choices as misaligned with their comparative advantage produces qualitatively similar treatment effects on skill beliefs (Table A20) and on search direction (Table A21).

As in any heterogeneity analysis, $Aligned_{id}$ may be correlated with other jobseeker-level characteristics, complicating interpretation of α^{diff}. This is a relatively minor concern in this specific analysis because having an aligned comparative advantage belief at baseline is unrelated to gender, age, employment, and work experience. It is related to the communication and numeracy assessment scores. But all these variables jointly explain only 15% of the belief variation (Table A11). Thus $Aligned_{id}$ likely captures heterogeneity by baseline comparative advantage beliefs more than other by other characteristics.

Both estimating equations, all baseline covariates, and most outcome measures are prespecified at https://doi.org/10.1257/rct.10000-1.0. We describe the relationship between the preanalysis plan and our final analysis in Appendix K.

3.3 Information About Skills Aligns Beliefs with Comparative Advantage

In the control group, 47.5% of jobseekers have aligned comparative advantage beliefs. This is measured as a dummy equal to one for those who believe they rank in a higher quintile for the skill in which they have 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, $p = 0.001$). This measure captures their

block fixed effects, to account for the fact that we randomize treatment within blocks of 4 sequential days.

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

	Aligned CA belief		Fraction aligned beliefs	
	(1)	(2)	(3)	(4)
Treatment	0.135***	0.208***	0.078***	0.036
	(0.035)	(0.050)	(0.026)	(0.028)
Treatment × Aligned CA belief (bl)		-0.137		0.080
		(0.082)		(0.050)
Aligned CA belief (bl)		0.586***		-0.050
		(0.079)		(0.046)
Treatment effect:		0.072		0.116***
Aligned CA belief (bl)		(0.058)		(0.041)
Control mean	0.475	0.475	0.183	0.183
Observations	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 indicate different outcomes: a dummy indicating if jobseeker's beliefs about their CA in skills are aligned with the assessed CA (cols. 1-2), and the average absolute deviation of beliefs about skill quintiles and assessed skill quintiles (cols. 3-4). Cols. 2, and 4 show treatment effect heterogeneity by whether jobseekers had aligned CA beliefs at baseline. Control variables are defined in footnote 15. Standard errors clustered at the treatment-day level shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

beliefs about their skills in general, not their results on our specific assessments. Section C.1 describes belief measurement.

Most of the treatment effect is driven by jobseekers with misaligned comparative advantage beliefs at baseline. Treatment increases the proportion of these jobseekers with aligned comparative advantage beliefs 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 1, row 1.

In contrast, treatment has modest effects on jobseekers with aligned beliefs at baseline. The proportion of this group with aligned beliefs increases by only 7.2pp ($p = 0.224$). 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.

Treatment shifts the level of skill beliefs as well as beliefs about skill comparative advantage. In the control group, 18% of jobseekers' skill belief levels match their assessed skill quintiles. Treatment increases this by 7.8pp, a 43% increase (Table 2, column 3, p=0.005). We show heterogeneous effects by baseline comparative advantage beliefs (column 4) but do not focus on them, because baseline comparative advantage beliefs do not reflect individuals' scope to learn about their skill levels.

Updating of comparative advantage beliefs is at least partially explained by learning about the skill distribution skill in the reference population (Appendix C.2). Treatment updates underconfident beliefs more than overconfident beliefs (Table A13), matching findings in other research (e.g. Zimmermann 2020). Belief updating does not differ by gender (Table A47) so we pool genders for the rest of the experimental analysis.

3.4 Job Search with Better Aligned Beliefs about Skills

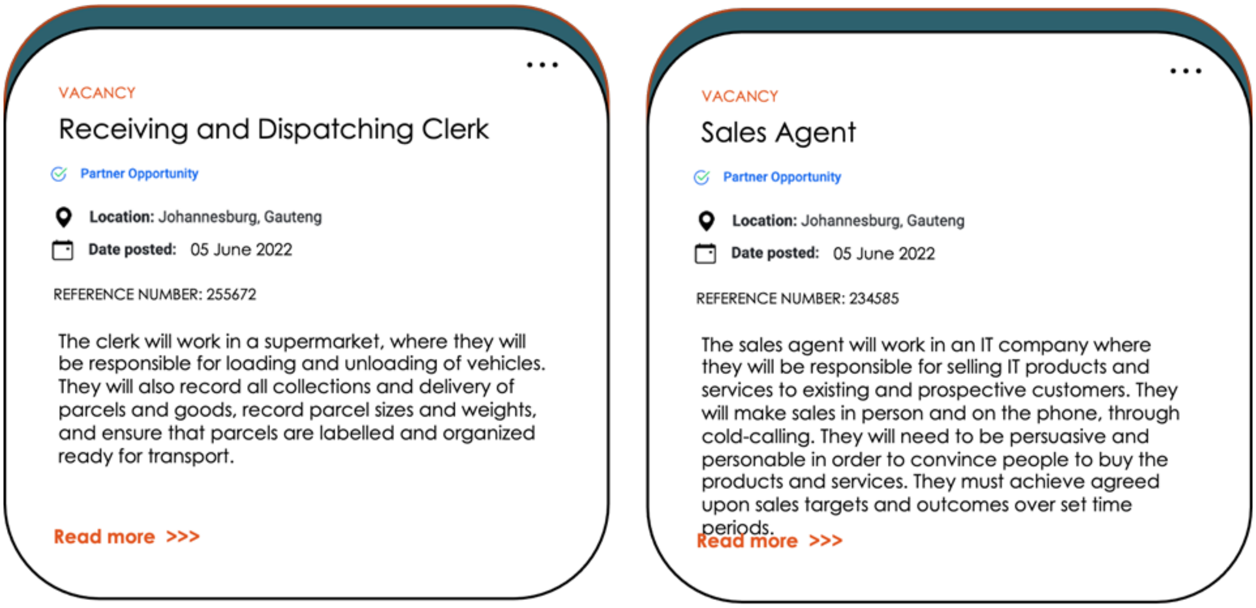
We show here that that treatment shifts jobseekers' search toward jobs that align with their assessed comparative advantage across four measures of skill-directed search.

Job choice task: We design a novel incentive-compatible job search task in which participants make 11 choices between pairs of job advertisements / vacancy postings. 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 advertisements, given time to read them, and asked to select one to apply to. Figure 2 shows an example pair and Appendix Table A17 shows the full list of job titles. Participants viewed the 11 job advertisement pairs in a random order.

The postings in the task are based on real job advertisements on SAYouth.mobi. All jobseekers in the study used the platform, so the choice of which vacancy to apply to represented a real-life choice they often made. To select the specific postings, we focused on entry-level vacancies in Johannesburg on SAYouth.mobi with no specialized education requirements. Among these, we selected a longlist of 28 vacancy postings with a clear numeracy or communication skill requirement, recognizing not all jobs to which jobseekers apply have such requirements. We asked 13 human resources and recruitment professionals with experience hiring for entry-level roles to rate each vacancy on the extent to which it required communication or numeracy skills; the expected wage; its overall and gender-specific desirability; and the transparency of skill requirements. We then created 11 pairs of jobs with similar expected wages and desirability but clearly requiring more of either communication or numeracy. Averaging across all pairs, recruiters scored the job we defined as numeracy-heavy as needing 2.7 standard deviations more numeracy skills but similar on other aspects. For example, the average within-pair difference in expected wages is $< 5\%$ of the mean wage. We removed information on employer name and location and standardized length and format.

The job pairs thus differ only in their skill demands, while treatment gives information only on jobseekers' relative ranks on communication and numeracy skills. This design allows us to measure how jobseekers' skill-directed job search changes in response to

Figure 2: Sample Pair of Jobs from Job Choice Task



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 outcome is the share of the 11 pairs in which the jobseeker chose the job that required the skill aligned with their measured comparative advantage. In the control group, this share was 55%. Treatment increases this share by 3.7pp (Table 3, column 3, $p = 0.29$). This average effect hides important heterogeneity. For jobseekers with misaligned baseline beliefs about their comparative advantage, treatment increases the share of aligned choices by 8.8pp, 16% of the control mean (column 4, $p = 0.024$). Treatment does not change search direction for jobseekers with baseline aligned beliefs.

We see a consistent pattern of heterogeneity for comparative advantage beliefs and search direction: treatment effects on both outcomes are driven by jobseekers with unaligned baseline beliefs about their skill comparative advantage. This pattern is also consistent with the predictions from our conceptual framework.

The search direction results suggest that jobseekers are able to interpret information in postings about the relative skill demands of different jobs. However, they might still face limited information about the jobs' skill requirements. We test this by explicitly revealing relative skill demand for the last two pairs of jobs (recall that order in which jobseekers see the 11 pairs of jobs is randomized). Jobseekers with initially misaligned beliefs align their search for job choices with and without revealed skill requirements, but substantially more when skill requirements are revealed (Table A27). This shows they have limited information about job skill requirements but enough for some skill-directed job search.

Application data from online search platform: With their consent, we link our sample to their profiles on SAYouth.mobi and observe their on-platform job search. In the 30 days after the workshop, the average participant in both the treatment and control groups initiated applications to 15 jobs on the platform, slightly higher than in the preceding 30 days (Table A1). The platform does not consistently record if applications are completed.

We classify vacancies by skill demand where possible. Of the 69,000 vacancies available on SAYouth.mobi in Johannesburg during our study period, we classify 14% as communication-heavy and 13% as numeracy-heavy jobs, suggesting demand among jobs advertised online is similar across skills.[16] SAYouth.mobi automatically scrapes data from all other online job boards so this represents the universe of online job board postings for Johannesburg. Post-treatment, 3.44 (23%) of applications initiated by jobseekers in our sample are to jobs classified as requiring communication or numeracy skills.

Treated jobseekers are more likely to initiate applications to jobs with skill demands that match their measured comparative advantage. To show this, we calculate each jobseeker's number of applications to vacancies coded as requiring the skill aligned with their comparative advantage, 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 submit 6.4pp more aligned applications than non-aligned applications (Table 3, column 5, $p = 0.011$). Again, this effect is driven by jobseekers with misaligned baseline comparative advantage beliefs (8.9pp, $p = 0.074$). We obtained these data directly from the platform after the experiment, so experimenter demand effects are unlikely.

Clicks on links to real jobs: Our third measure captures whether jobseekers can conduct skill-directed job search using only job titles. We send jobseekers three text messages with links to real job opportunities on SAYouth.mobi about a week after the workshop. We send, 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

[16] We classify job adverts that contain one of 20 numeracy or 20 communication-heavy job titles in the advert title or job description as requiring the respective skill. The lists of skill-specific job titles was created by the recruitment staff who classify vacancies for the job choice task.

Table 3: Treatment Effects on Search Direction - Tight Experiment

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

Notes: Table 3 shows that informing jobseekers about their relative comparative advantage in skills aligns their search direction with their assessed comparative advantage 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: an inverse-covariance weighted average of the search alignment measures displayed in columns 3 to 10 following Anderson (2008) (cols. 1-2), the percentage of 11 incentivized job choices that are aligned with the measured CA of the jobseeker (cols. 3-4), the difference between the percentage of aligned and non-aligned applications on the online job search platform SAYouth.mobi (cols. 5-6), the difference in link click rates between aligned and non-aligned jobs sent to job seekers via text message (cols. 7-8), and the difference between aligned and non-aligned planned applications for the 30 days after the workshop (cols. 9-10). Even columns show heterogeneity by whether individuals have aligned CA beliefs at baseline. Control variables are defined in footnote 15. All winsorized variables (w) are winsorized at the 99^{th} percentile. Standard errors clustered at the treatment-day level shown in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

contain a greeting linking the message to the workshop they attended, a note that we found a job opportunity of interest in 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 track whether jobseekers clicked on these links.

We again find treatment increases skill-directed job search. Treatment increases the difference in click rates between jobs that are aligned and misaligned with the jobseeker's comparative advantage by 7pp (Table 3, column 7, $p = 0.25$). The effect is again driven by jobseekers with initially misaligned beliefs. However, estimates are relatively imprecise because this measure uses 3 choices per jobseeker, compared to 11 in the job choice task.

Planned applications to numeracy and communication jobs: After the treatment, before the job choice task, we survey participants about the number of applications they plan 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 comparative advantage, 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 3, column 9, $p = 0.27$). This effect is driven by a 4.8 application effect for jobseekers with misaligned baseline comparative advantage beliefs (Table 3, column 10, $p = 0.011$). The result reflects both a slight increase in planned aligned applications and a reduction in planned misaligned applications. This measure is perhaps more sensitive 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 3, column 1, $p = 0.013$). This effect is entirely driven by jobseekers with misaligned baseline comparative advantage beliefs, for whom we observe a shift in search direction of 0.6 standard deviations (Table 3, column 1, $p < 0.001$). For this index, treatment closes most of the gap between jobseekers with aligned versus misaligned baseline beliefs: treated jobseekers with misaligned baseline comparative advantage beliefs have almost the same level of skill-directed job search (0.60, shown in column 2, row 1) as control jobseekers with aligned baseline comparative advantage beliefs (0.73, shown in column 2, row 3).

3.5 Additional Results and Robustness Tests

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 evaluate this prediction, we measure jobseekers' expectations about the outcomes of applying to communication-and numeracy-heavy jobs, both in the job choice task during the workshop and in their planned search after the workshop. For example, for a jobseeker with numeracy comparative advantage, we construct the expected wage for numeracy-heavy jobs minus the expected wage for communication-heavy jobs.

Treatment increases expected returns to most measures of skill-directed search (Table A30), although some results are imprecisely estimated. Jobseekers' expectations about job-specific wages also predict choices in the job choice task (Table A33). These results are consistent with the model. Appendix F gives full details of the measurement and results.

Robustness 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 following Benjamini et al. (2006) (Table A22). Treatment effect estimates are robust to using more continuous measures of skill belief alignment (Table A23) and to controlling for baseline 'confidence' – the difference between believed and assessed skill quintiles – and its interaction with treatment (Table A26).

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

Having shown that information about skills can shift comparative advantage beliefs and search direction, we now test if this information can also improve individual labor market outcomes in an 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 2016/7, five years before the tight experiment. Recruitment is from the same population: active jobseekers on the database of our partner Harambee, described in Section 2.3. This recruitment is designed to capture active jobseekers with limited access to traditional ways of learning about their skills.

Table A2 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.

There are some differences in demographic characteristics. Jobseekers in the big experiment are slightly younger than in the tight experiment (24 versus 26) and more likely to be male (38 vs 32%), reflecting changes in the demographics of Johannesburg (Statistics South Africa, 2016, 2022). A similar proportion have finished high school, but slightly fewer in the big experiment have university degrees (17 vs 22%) and slightly more have other shorter tertiary qualifications (22 vs 15%), reflecting increased financial aid at universities through time. Only 9% had ever held a permanent or long-term job (versus 23% in the tight experiment), perhaps due to their younger age, longer time in education, or

COVID-related labor market contractions.

But our key results are robust to accounting for these differences. Table A22 shows 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.

4.2 Experimental Design

In the big experiment, we assign 2,114 jobseekers in 27 workshops to the treatment group and 2,274 jobseekers in a different 27 workshops to the control group. To avoid spillovers, treatment is administered at the workshop (day) level. Treatment groups are balanced on covariates for both the full sample and the 96% recontacted for the endline survey (Table A18). Endline response rates are balanced across treatment groups (Table A19).

The big experiment's design is similar to that of the tight experiment, which was designed to mimic it. The big experiment was also run during Harambee workshops where jobseekers receive job search support and sit skills assessments. It also randomly varied whether jobseekers received an unbranded report about their scores on skills assessments relative to similar jobseekers (in the treatment group) or received information about what skills are assessed, but not their own relative scores (in the control). The jobseekers' experience at the workshop, is shown in Figure A3 and was very similar to the tight experiment, including how and when assessments are administered and beliefs were measured. All jobseekers received job search assistance – coaching on drafting a CV and cover letter – although this is at the beginning of the day in the big experiment and the end of the day in the tight experiment. In the experiments, participants know they are participating in a study but not that treatment differs by day, so their behavior and survey responses should not reflect experimenter demand effects.

The experiments differ in four ways. First, the big experiment is designed to measure effects on labor market outcomes, so we study a larger sample and collect post-treatment data later: on average 3.5 months after treatment.[17]

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

In the big experiment, we assess communication, numeracy, and concept formation

[17] We use phone call surveys lasting an average of 25 minutes. We compensate respondents with mobile phone airtime payments. Garlick et al. (2020) show that phone and in-person surveys in this setting deliver similar labor market data.

skills (as in the tight experiment) as well as three non-cognitive measures described in detail in Appendix B: focus, grit, and planning. The reports given to treated jobseekers show results and information on what traits and characteristics all six skill assessments measure (Figure A4), rather than communication and numeracy (as in the tight experiment).[18] The six assessments differentiate jobseekers horizontally more than vertically because assessment results are weakly correlated across skills within candidate and do not allow an easy aggregate ranking of candidates. 12 of the 15 pairwise correlations are below 0.2 (Table A5). Most jobseekers learn they have substantial variation across skills: 85% have at least one top tercile but onlt 1.7% have all six top terciles and 58% have both top and bottom terciles.

Third, we report assessment results in terciles, not quintiles. The coarseness of terciles relative to quintiles and using six rather than two skills means that only 23% of jobseekers have a unique skill comparative advantage, compared to 75% in the tight experiment. We thus keep jobseekers without a unique skill comparative advantage in our sample.[19]

Fourth, there are logistical differences that are unlikely to affect the core mechanism activated by the treatment. In the big experiment, candidates are assessed in bigger groups. The briefing after candidates receive the report is delivered in person instead of by video (but with a script to ensure fidelity in messaging across sessions).

Appendix D contains a detailed description of the workshops and links to the scripts.

4.3 Specification

We estimate average treatment effects using the following specification:[20]

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

Y_{id} is the outcome of interest for jobseeker i assessed on date d and $\mathbf{X}_{id}$ is a vector of pre-specified control variables.[21] The object of interest is the average treatment effect β. As in the tight experiment, we cluster standard errors at the level of treatment (assessment

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

[19]Recall that we focus our analysis in the tight experiment on jobseekers with a unique skill comparative advantage. However, including these jobseekers in the sample and classifying all their belief and job search outcomes as misaligned with their comparative advantage produces qualitatively similar treatment effects on skill beliefs (Table A20) and search direction (Table A21).

[20]Most of the outcome variable definitions, inference methods, and covariates were prespecified, as for Carranza et al. (2022). However, comparative advantage beliefs and aligned search were not prespecified as outcomes. Our analysis of these outcomes should be viewed as ex post in the big experiment but prespecified in the tight experiment. See Appendix K for details.

[21] $\mathbf{X}_{id}$ contains baseline assessment results, self-reported skills, 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.

days) and we combine families of outcomes in inverse covariance-weighted average indices to limit the number of tested hypotheses, following Anderson (2008). Our main results are robust to using a wild cluster bootstrap and to adjusting for multiple hypothesis testing, following Benjamini et al. (2006) (Table A22).

4.4 Information About Skills Aligns Beliefs with Assessed Comparative Advantage

Receiving information about skills substantially increases alignment between measured skills and skill beliefs 3.5 months later. To show this, we ask candidates in which tercile they believe they ranked for each of the communication, numeracy, and concept formation assessments (see Appendix C.1 for measurement details). Jobseekers have aligned comparative advantage beliefs if their believed top skill matches their assessed top skill across the three skills. If jobseekers have tied top skill terciles, we require them to also believe that they have tied top skill terciles.

Only 19% of control group jobseekers have aligned comparative advantage beliefs on this measure. Treatment increases this by 13.9pp, a 72% increase (Table 4, column 1, $p < 0.001$). This is in line with the effects on beliefs observed straight after treatment in the tight experiment.[22] This shows that jobseekers' updated beliefs persist over 3.5 months and that jobseekers are able to process more complex information about assessment results in multiple skills.

Treatment also increases the fraction of skill beliefs that align with assessment results by 14.2pp, a 37% increase from the control mean of 38.8pp (Table 4, column 2, $p < 0.001$). As in the tight experiment, we observe no gender heterogeneity in baseline beliefs or belief updating, so we report all analysis pooling genders (Table A48).

4.5 Job Search with Better-Aligned Comparative Advantage Beliefs

We collect a self-reported measure of whether jobseekers' search direction matches their skill comparative advantage. We ask jobseekers to think about the types of jobs they are applying for and what skills these jobs demand and then to rank the following skills from most to least important: numeracy skills, communication skills, problem-solving skills (the concept formation measure), and soft skills. Aligned search is a dummy indicating that jobseekers' highest ranked skill on this measure (ignoring soft skills) matches the skill in which they have an assessed comparative advantage.[23] Treatment increases this

[22] The effects on aligned comparative advantage 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 comparative advantage to differ. In both experiments, the control group's alignment rate is only slightly higher than random guessing, accounting for the possibility of ties: 19 versus 15% in the big experiment and 48 versus 40% in the tight experiment.

[23] For individuals with a tie for highest skill, our main aligned search measure is set to zero. However, results are robust to recoding this measure to one when the jobseeker says that any one of their tied highest

Table 4: Treatment Effects on Skill Beliefs and Search Direction - Big Experiment

	Beliefs		Search direction
	Aligned CA belief (1)	Fraction aligned beliefs (2)	Aligned search (3)
Treatment	0.139***	0.142***	0.050***
	(0.011)	(0.009)	(0.010)
Control mean	0.196	0.388	0.165
Observations	4118	4195	4205

Notes: Table 4 shows that informing jobseekers about their comparative advantage in skills aligns their skill beliefs with assessed skills and shifts self-reported search direction in the big experiment. 'CA' stands for comparative advantage and 'bl' for baseline. Columns indicate different outcome variables: dummy for aligned CA belief (col. 1), fraction of aligned skill tercile beliefs across communication, numeracy, and concept formation (col. 2), and a dummy indicating if self-reported search direction is aligned with assessed skill CA (col. 3). Control variables are described in footnote 21. Standard errors clustered at the treatment-day level shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

aligned search measure from 17% to 22%, 3.5 months after treatment (Table 4, column 3, $p < 0.001$). This result on self-reported skill-directed job search is qualitatively consistent with the four measures of skill-directed search in the tight experiment, collected in the first month after treatment.[24]

4.6 Improved Labor Market Outcomes

The results on beliefs and search from the two experiments provide strong evidence that giving jobseekers information about their skill assessment results better aligns their search with their assessed skill comparative advantage. We now test the model prediction that the change in search direction should improve jobseekers' labor market outcomes.

Treatment has limited effects on employment quantity: Treatment effects on the number of job offers and probabilities of employment in the first and second months after the workshop and week before the endline survey are positive but mostly small and not statistically significant (Table 5, panel A, columns 2-5). Our prespecified index of employment quantity increases by 0.05 standard deviations (column 1, $p = 0.13$).

Treatment substantially increases employment quality: Treatment increases earnings in the seven days before the endline by 6.52 USD (Table 5, panel B, column 2, $p = 0.019$). This is equivalent to a 26 percent increase from the control group mean or to mov-

skills is most important for the jobs to which they are applying.

[24]The big experiment effect converts to a 0.14 standard deviation increase in directed search. The effect on the directed search index in the tight experiment is 0.27 standard deviations (Table 3, 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.

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

Panel A: Work quantity					
	Index	Job offers (w)	Worked m1	Worked m2	Worked last 7d
	(1)	(2)	(3)	(4)	(5)
Treatment	0.045	0.015	0.023*	0.007	0.009
	(0.032)	(0.017)	(0.013)	(0.015)	(0.013)
Control mean	-0.000	0.182	0.465	0.437	0.309
Observations	4205	4140	4201	4204	4204

Panel B: Work quality				
	Index	Earnings (w)	Hourly wage (w)	Written contract
	(1)	(2)	(3)	(4)
Treatment	0.085**	6.517**	0.295**	0.017*
	(0.035)	(2.712)	(0.130)	(0.010)
Control mean	0.000	25.424	1.267	0.120
Observations	4206	4196	4183	4184

Notes: Table 5 shows that the treatment improves employment quality but not necessarily employment quantity in the big experiment. Panel A shows effects on employment quantity. The columns indicate different outcomes: an inverse covariance-weighted average of the four employment quantity measures (col. 1), the winsorized number of job offers in the last 30 days (col. 2), a dummy indicating any work for pay in month 1 after treatment (col. 3), a dummy indicating any work for pay in month 2 after treatment (col. 4) and a dummy indicating any work for pay in the seven days before the endline survey (col. 5). Panel B shows effects on employment quality. The columns indicate different outcomes: an inverse covariance-weighted average of the three employment quality measures (col. 1), winsorized earnings in the last seven days (col. 2), winsorized hourly wages in the last seven days (col. 3), and a dummy indicating a written contract (col. 4). Employment quality measures are coded as zeroes for the non-employed. Control variables are described in footnote 21. All monetary figures are reported in 2021 USD PPP. All winsorized variables are winsorized at the 99^{th} percentile. Standard errors clustered at the treatment-day level shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

ing from the 75^{th} to 79^{th} percentile of the unconditional earnings distribution.

We code earnings as zero for all non-employed jobseekers, so this variable has a mass point at zero and a continuous distribution above zero. Treatment effects on such measures can be sensitive to scaling choices (Mullahy & Norton, 2022) so we report several robustness checks. First, effects are weakly positive on the entire earnings distribution (Figure A5). Second, treatment raises earnings by 9.4 USD without winsorization, 0.1 standard deviations, 12 log points or 13 inverse hyperbolic sine points, demonstrating robustness across a range of scalings (Table A24, columns 1 and 3-5). Third, restricting the sample to the employed produces qualitatively similar treatment effects relative to the control mean (Table A24, columns 6-7).[25] The earnings effects are within the range of

[25]Restricting analysis to the employed can create a sample selection problem when treatment affects the

effect sizes in related research, although toward the top of that range.[26]

Treatment increases hourly wages by 0.30 USD or 24 percent of the control mean of 1.27 USD ($p = 0.028$, Table 5, panel B, column 3). Treatment also increases the probability of having a written contract by 1.7 percentage points from a 12% mean in the control group (Table 5, panel B, column 4, $p = 0.095$). Written contracts define formal jobs in this context and are valued by workers. Our prespecified index of employment quality increases by 0.085 standard deviations (Table 5, panel B, column 1, $p = 0.019$).

Heterogeneous treatment effects: We report heterogeneous treatment effects by aligned baseline comparative advantage beliefs, as in the tight experiment. However, we interpret these effects with caution because we have a noisier proxy for the "dose" of information delivered by the big than tight experiment. We only observe beliefs about three of the six skills in the big experiment and the additional skills mean there are many more ways for assessed and believed comparative advantage to differ.

Subject to this caveats heterogeneity patterns in this experiment are roughly consistent with larger effects for jobseekers who get more information from treatment. Treatment effects on labor market outcomes are driven by jobseekers with misaligned baseline comparative advantage beliefs (Table A29). Although none of the differences is statistically significant at conventional levels, some are large. For example, the effects on weekly earnings are 7.5 and 2.8 USD for jobseekers with respectively misaligned and aligned baseline comparative advantage beliefs (column 7). Similarly, the treatment effect on skill-directed job search is driven by jobseekers with misaligned baseline beliefs (Table A25, column 3), although this does not hold for the effect on comparative advantage beliefs (column 1).

4.7 How Does Skill-Directed Job Search Improve Labor Market Outcomes?

Here we explore possible economic mechanisms that link skill-directed search to better-paid and higher-quality jobs but not to a higher employment probability. Given the nature of our data, this subsection is less conclusive than the rest of the paper.

We begin by motivating our focus on search-based mechanisms. Treatment effects on earnings and formality are driven by wage jobs that start after treatment, rather than self-employment or wage jobs that start before treatment (Table A28). This suggests that

probability of employment. But, in this case, treatment has little impact on employment, so it is unsurprising that the results are qualitatively robust to different ways of handling the earnings of the nonemployed. Treatment effects for the log and IHS specifications in the sample of employed jobseekers are slightly larger, potentially due to the scale-dependence of these specifications (Mullahy & Norton, 2022).

[26] For example, interventions that combine learning about skills with the ability to signal skills increase earnings by 11-34% (Abebe et al., 2021b; Bassi & Nansamba, 2022; Carranza et al., 2022). Bandiera et al. (2023) find an 11% change 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 increases earnings by 11 percent.

self-employment and bargaining with existing employers are less important mechanisms than search.[27] Given this pattern, we evaluate two non-mutually exclusive mechanisms centered on two standard concepts in the job search literature: **job offer probabilities** and **wage offer distributions**. Returning to our model from Section 2.1, we assume there are multiple job types indexed by j and searching for each job type has an associated wage offer distribution $F_j(w)$ and job offer probability P_j. Each jobseeker's direction of search effort across different job types generates a wage offer distribution $F(w)$ and job offer probability P that are weighted averages across job types. As in standard job search models, jobseekers accept offers with wages above their reservation wages.

In the first potential mechanism, skill-directed search raises P, so some jobseekers receive multiple offers and can choose the higher-wage offers or bargain over wages. Our data are not consistent with this mechanism: treatment effects on the numbers of offers received and refused are both $< 3\%$ of the control group mean (p>0.4) and only 4% of jobseekers receive multiple offers, giving little scope for choice or bargaining.

In the second potential mechanism, skill-directed job search shifts at least part of $F(w)$ to the right, leading to higher wage offers and raising average earnings for the employed. We do not observe the terms of job offers, so we cannot directly test this. But the treatment effect on the CDF of earnings is consistent with a rightward shift in F (Figure A5).[28]

Under the wage offer mechanism, skill-directed job search might generate higher wage offers by raising **job quality** or **match quality**. Higher job quality would occur if some job types pay more to all workers conditional on their skill match, and treated jobseekers applied more to them. 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 pay them more.

We find slightly more evidence consistent with higher match quality than job quality. To show this, we construct a list of all job titles that receive an application from any jobseeker in the tight experiment on the SAYouth.mobi platform. 96.7% of these job titles receive applications from both treated and control jobseekers. This suggests a high overlap in the types of jobs to which the groups apply, making a job quality effect unlikely. Within this overlapping set of job titles, we have already shown that treated jobseekers are more likely to apply for jobs that match their skills. Furthermore, the average jobseek-

[27]We don't distinguish between on- and off-the-job search in this discussion because there are limited differences employed and unemployed participants' search behavior, in both control and treatment groups.

[28]To match the null effect on employment, this mechanism also requires either that reservation wages rise, so some of the higher wage offers are rejected, or that the shift in F occurs only above reservation wages. Treatment increases reservation wages in both experiments but the effects are not statistically significant. This may reflect widespread challenges in measuring reservation wages (Feld et al., 2022).

ers in the tight experiment sends only 26.7% of applications to jobs that prespecify a fixed wage. This suggests some scope for wage offers to vary with skill match, making a match quality mechanism possible.

We conclude that treatment may raise earnings through more skill-directed search, generating higher wage offers, due to higher match quality. However, our evidence for higher wage offers and match quality is relatively indirect, so we view those as more tentative results than the strong evidence for skill-directed search and higher earnings.[29]

4.8 Scale and General Equilibrium Considerations

Job search interventions can generate different results at smaller and larger scales, due to interactions between jobseekers or between jobseekers and firms (Altmann et al., 2022). Here we present a brief framework for thinking about scale and general equilibrium implications of the economic mechanisms we study. We provide empirical evidence about parts of this framework, which suggest that the effects of our intervention on jobseekers' labor market outcomes need not be smaller at larger scales. But our experiments are not designed to study general equilibrium. So this subsection is deliberately brief, tentative, and intended to facilitate future research, rather than providing conclusive evidence.

There are two obvious channels for general equilibrium effects of our intervention. First, jobseekers applying to the same jobs can generate **search congestion**. Search congestion occurs when a treatment-induced shift in search behavior increases competition for some jobs or job types, leading directly to lower offer probabilities and potentially to lower wage offers as firms observe higher application volumes. This is unlikely for the type of information intervention we study. Information about skill comparative advantage 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. For example, in the job choice task, treatment spreads applications marginally more evenly between communication and numeracy jobs (Table A44). And treatment effects on the numbers of on-platform applications to communication and numeracy jobs are $<2\%$ of the total numbers of applications. Less differenti-

[29]In principle, it is also possible that treatment might affect labor market outcomes through a feedback loop between beliefs and labor market experience: treatment $\rightarrow$ beliefs about skill levels or comparative advantage $\rightarrow$ job search $\rightarrow$ labor market outcomes $\rightarrow$ skill beliefs. We find little evidence consistent with such a feedback loop. First, treatment has little effect on employment, so this feedback loop would have to operate through the type of employment, not the quantity. Second, treatment in the tight experiment has large effects on skill beliefs before there's time for any learning from labor market experience. Third, employment in the big experiment control group is uncorrelated with changes in beliefs, as we discuss in Section 2.6. Even if this feedback loop were large, it would not undermine the skill-directed search explanation. It would simply mean that the treatment effects on labor market outcomes reflects both a direct effect and an indirect effect through this feedback loop.

ated interventions may be more prone to search congestion effects, such as those providing non-differentiated information about high-demand sectors or locations, encouraging search effort, or providing job placement services (e.g. Crepon et al. 2013; Johnston & Mas 2018; LaLive et al. 2022).[30]

Second, 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 in both classic matching models (e.g. Mortenson & Pissarides 1994) and recent experiments (e.g. Algan et al. 2022). We show some evidence in the preceding subsection that our intervention raises match quality, although this evidence is indirect. Our intervention also affects earnings conditional on employment, not the employment margin that has been studied by most research on negative spillover effects of job search interventions. It is possible that earnings are more sensitive to match quality effects and employment more sensitive to search congestion effects, at least in the short run before employers respond to search congestion by lowering wage offers.

We tentatively conclude that there is no clear evidence that the type of intervention we study would generate smaller effects at larger scale. However, a more conclusive answer would need research designed to evaluate general equilibrium effects.

We also note that search congestion effects are unlikely to affect the internal validity of our experiments. Our big experiment involves only 4,400 people in a city with roughly 8 million people and 2 million employed workers (Statistics South Africa, 2016); they are sampled from neighborhoods around the city, rather than one small geographic area; treatment is spread over seven months; and participating jobseekers are not encouraged to apply to specific jobs or search for work in specific areas. This suggests that competition between treated and control jobseekers is unlikely to be a large factor.

5 Search Effort and Beliefs About Skill Levels

Potentially, changes in labor market outcomes can be explained by jobseekers being on average overconfident, learning they have lower skills than they thought, and changing search effort. However, we find few treatment effects on search effort in either experiment, suggesting it does not explain the relationship between treatment and labor market outcomes.[31] In the tight experiment, treatment effects are close to zero and not statistically significant on six different measures of search effort: a survey question on

[30] Information about skill comparative advantage 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 on the SAYouth.mobi platform.

[31] 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 (2022).

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 (Tables A36 and A37). In the big experiment, we estimate precise near-zero effects on self-reported number of applications, time spent and money spent on search (Table A38).

The lack of effects on search effort is perhaps surprising. Indeed, we find treatment lowers jobseekers' believed skill levels in both experiments (Table A35, columns 3 and 7). 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 (columns 4 and 8).

However, a generalized version of our conceptual framework clarifies that the level of skill beliefs has a theoretically ambiguous effect on search effort. In this framework, jobseekers endogenously choose their total search effort level based on their expected search outcomes, which fall when treatment lowers their believed skill level. The framework predicts two responses to this effect on skill beliefs. 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 show results and the full conceptual framework in Appendix H.

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.[32]

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 months after treatment (Table A41, columns 1-4), 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 and 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 in the big experiment. 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. (Table A41, columns 5-7).

We find suggestive evidence in the tight experiment that jobseekers might be willing

[32]We prespecified testing for the "human capital investment" mechanism in both experiments and for all three mechanisms in the big experiment.

to invest in skills. Treatment reduces willingness-to-pay (WTP) for a numeracy workbook, mostly for people who learn that they have a comparative advantage in numeracy, suggesting that jobseekers might prefer to invest in skills where they are relatively weak. WTP for a communication workbook is unaffected by treatment (Table A42). Appendix D.1 gives details on the WTP measurement.

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 (Table A43), although we interpret this result cautiously because this analysis conditions on report use, a post-treatment outcome. We have conflicting evidence on how often jobseekers share results with firms: 29% of treated jobseekers in the big experiment report that they ever included a copy of their assessment results with any application in the 3.5 months after treatment but only 0.8% of treated jobseekers receiving this intervention actually included their assessment results in applications to vacancies we created.

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.[33]

7 Conclusion

We provide evidence of a relatively understudied type of job search friction: misdirected search due to limited information about skill comparative advantage. We show that skill-directed job search can be constrained by jobseekers' limited information about their own comparative advantage. Giving jobseekers additional information about their skill comparative advantage can shift their beliefs, redirect search effort toward jobs that better match their comparative advantage, and allow them to get better-quality jobs.

This is important for interpreting established research on firms' limited information about jobseekers' skills. For example, misdirected job search can reduce the effectiveness of firms' investment in screening prospective jobseekers, because it means firms don't see the ideal pool of applicants. It also helps to interpret research on jobseekers' limited information about their labor market prospects. Jobseekers' stated beliefs about their labor

[33] 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.

market prospects implicitly condition on beliefs about their skills and their search strategies given these skills. So directly measuring these can shed light on the co-evolution of beliefs and search behavior. Our findings can inform models of job search and matching with multidimensional skills by providing direct evidence supporting models in which jobseekers' limited information about their relative skills distorts how they direct search.

Our approach highlights the benefits of combining multiple experimental designs with multiple measurement strategies. This allows us to observe shifts in beliefs and search direction in a simple two-skill setting with more precise measurement, and to study effects on labor market outcomes in a larger sample with a longer timeframe.

We end on some deliberately speculative questions for future research. Our results show private gains for jobseekers who acquire more information about their skill comparative advantage. Can this type of information be efficiently provided to jobseekers outside the context of research like this? And by whom? We show in Appendix G that the treatment effect on earnings implies that the average treated jobseeker earns enough extra in the 3.5 months between treatment and endline to cover 1.8 times the average variable cost of the assessment operation. This suggests the possibility of profitable market provision. Assessment allows substantial economies of scale, particularly if implemented by online job search and matching platforms. Many platforms, including our partner's SAYouth.mobi, already offer skill assessments to jobseekers (LinkedIn, 2023). However, 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. A strong education system can in principle provide graduates with reliable information about their comparative advantage, 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 could fill this gap. These sometimes include skill assessments but researchers have not yet studied the role 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 comparative advantage 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 (2017) "Tackling Youth Unemployment: Evidence from a Labor Market Experiment in Uganda," Manuscript, University College London.

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

Algan, Y., B. Cr'epon & 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., 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," *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., 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.

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., A. Elsayed, A. Smurra & C. Zipfel (2022a) "Young Adults and Labor Markets in Africa," *Journal of Economic Perspectives*, 36 (1), 81–100.

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

Banerjee, A. V. & S. Sequeira (2023) "Spatial Mismatches and Imperfect Information in the Job Search," *Journal of Development Economics*, forthcoming.

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

Bazzi, S., L. Cameron, S. G. Schaner & F. Witoelar (2021) "Information, Intermediaries, and Inter-

national Migration," NBER Working Paper 29588.
Beam, E. (2016) "Do Job Fairs Matter? Experimental Evidence from the Philippines," *Journal of Development Economics*, 120, 32–40.
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.
——— (2022b) "How Wage Announcements Affect Job Search – A Field Experiment," *American Economic Journal: Macroeconomics*, 14 (4), 1–67.
Belot, M., P. Kircher & P. Mueller (2022a) "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," Working paper.
Bobba, M. & V. Frisancho (2020) "Self-perceptions about academic achievement: evidence from Mexico City," *Journal of Econometrics*.
Boudreau, L., R. Heath & T. McCormick (2022) "Migrants, Experience, and Working Conditions in Bangladeshi Garment Factories," SSRN Working Paper 4067206.
Card, D., J. Kluve & A. Weber (2018) "What Works? A Meta-Analysis of Recent Active Labor Market Program Evaluations," *Journal of the European Economic Association*, 16 (3), 894–931.
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.
Chen, J. & J. Roth (2022) "Log-like? Identified ATEs Defined with Zero-Valued Outcomes Are (Arbitrarily) Scale-Dependent," Working Paper.
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 Business Majors," *Quarterly Journal of Economics*, forthcoming.
Crepon, B., E. Duflo, M. Gurgand, R. Rathelot & P. Zamora (2013) "Do Labor Market Policies Have Displacement Effects? Evidence from a Clustered Randomized Experiment," *Quarterly Journal of Economics*, 128 (2), 531–580.
Delavande, A. (2022) "Expectations in Development Economics," in Van der Klaauw, W., G. Topa & R. Bachman eds. *Handbook of Economic Expectations*, 261–291: Elsevier.
Diamond, A. (2013) "Executive Functions," *Annual Review of Psychology*, 64, 135–168.
Donovan, K., J. Lu & T. Schoellman (2022) "Labor Market Dynamics and Development," Working Paper, Yale University.
Duckworth, A. (2016) *Grit: The Power of Passion and Perseverance*, New York, NY, US: Scribner/Simon & Schuster.
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.

Field, E., R. Garlick, N. Subramanian & K. Vyborny (2023) “Why Don’t Jobseekers Search More? Barriers and Returns to Search on a Job Matching Platform,” 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.

Gibbons, R., L. Katz, T. Lemieux & D. Parent (2005) “Comparative Advantage, Learning, and Sectoral Wage Determination,” *Journal of Labor Economics*, 23 (4), 681–724.

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

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).

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 (2017) “Are Small Firms Labor Constrained? Experimental Evidence from Ghana,” Working Paper, New York University - Abu Dhabi.

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

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.

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 (2020) “How Do Online Job Portals Affect Employment and Job Search? Evidence from India,” Working Paper.

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

Kessler, J. B., C. Low & C. D. Sullivan (2019) "Incentivized Resume Rating: Eliciting Employer Preferences without Deception," *American Economic Review*, 109 (11), 3713–44.

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.

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.

McKenzie, D. (2017) "How Effective Are Active Labor Market Policies in Developing Countries? A Critical Review of Recent Evidence," *World Bank Research Observer*, 32 (2), 127–154.

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 (2022) "Expectations Data, Labor Market and Job Search," in Bachmann, R., G. Topa & W. van der Klaauw eds. *Handbook of Economic Expectations*, Chap. 22, 677–710, Netherlands: Elsevier Science.

Mueller, A. I., J. Spinnewijn & G. Topa (2021) "Job Seekers' Perceptions and Employment Prospects: Heterogeneity, Duration Dependence and Bias," *Econometrica*, 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.

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.

Sanders, C., C. Taber et al. (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.

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

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

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.

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," *The Review of Economic Studies*, 82 (2), 791–824.

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

Zimmermann, F. (2020) "The Dynamics of Motivated Beliefs," *American Economic Review*, 110 (2), 337–363.

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

This online material contains twelve appendices. Appendix A contains appendix summary statistics tables. Appendix B describes our skill measurements and show that firms not only value the skills we study but are also able to detect them. In Appendix C we describe how the skill beliefs were elicited and provide further descriptive statistics on these beliefs. Appendix D describes the experimental protocols in detail. We collect robustness checks and various heterogenous treatment effect results in Appendix E. Appendix F provides an additional analysis of the role of labor market beliefs in shaping search direction. Appendix G details the cost-benefit calculation. Appendix H contains additional exhibits and results that show that changes in search effort are unlikely to explain our results – as discussed in Section 5. Appendix I contains exhibits supporting the additional mechanism analyses from Section 6. Appendix J shows the results related to gender heterogeneity. Finally, Appendix K describes how our analysis relates to the pre-analysis plans.

A Sample Description

This appendix contains two additional exhibits that describe the two samples respectively. Table A1 shows that jobseekers in the tight experiment are active on the platform 30 days prior to treatment. Table A2 provides summary statistics for jobseekers in the big experiment. Comparing this table to Table 1 confirms that the two sample are very similar.

Table A1: Summary Statistics for Search on SAYouth.mobi Platform - Tight Experiment

	Mean (1)	Median (2)	Min (3)	Max (4)	SD (5)	Obs. (6)
# days active on platform	3.17	2.00	0.00	25.00	4.06	278
# applications clicks (winsorized)	6.53	2.00	0.00	66.00	11.27	278
# applications clicks for numeracy heavy jobs	0.39	0.00	0.00	11.00	1.10	278
# applications clicks for communication heavy jobs	0.82	0.00	0.00	9.00	1.64	278
Fraction of skill coded application clicks	0.12	0.00	0.00	1.00	0.20	278

Notes: **Table A1 shows summary statistics of participants' engagement with the job search platform (SAYouth.mobi) 30 days prior to the intervention in the tight experiment.** Application clicks are defined as initiating an application by clicking on the button "apply here". Winsorized variables are winsorized at the 99^{th} percentile.

Table A2: 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 A2 shows summary statistics for the big experiment.** CA stands for comparative advantage. Winsorized variables are winsorized at the 99^{th} percentile. All monetary values are in 2021 USD purchasing power parity terms.

B Skills

In this section we focus on the measurement of skills, and provide support that firms value the skills we study and are able to detect them. The implementation details of the assessments can be found in Appendix D.

B.1 Measurement

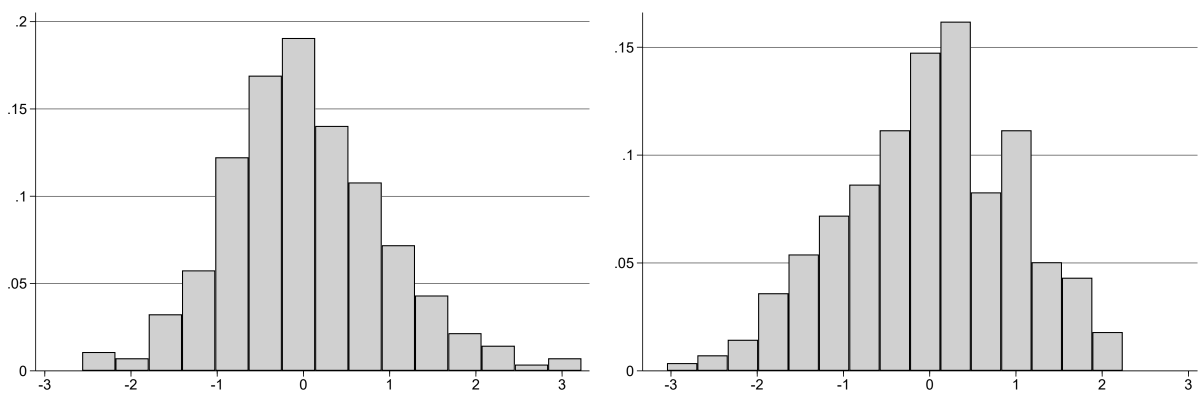
Tight Experiment In the tight experiment, we assess jobseekers' skills on numeracy, communication, and concept formation. The *numeracy* assessment was based on a test developed by a retail chain in South Africa. The chain uses the test to assess candidates' skills needed to become a cashier. The *communication* assessment quantifies jobseekers' listening and reading comprehension. The assessment focuses on high school English proficiency and was developed by a local adult education provider (www.mediaworks.co.za). The *concept formation* assessment was a test similar to Raven's matrices (Raven & Raven, 2003). With this measure, we proxy for jobseekers' fluid intelligence: people's conceptual reasoning and the rate at which people learn.

We find ample variation within skills and across skill quintiles. Figure A1 shows the distribution of measured numeracy and communication scores after standardization in the tight experiment. There is limited scope for flooring or ceiling effects based on the shape of the distributions. Table A3 shows the joint distribution of numeracy skill quintiles and communication skill quintiles. Table A4 shows the correlations between skill measures in the tight experiment. We find moderate positive correlations between all three skill measures.

Figure A1: Distribution of Standardized Assessment Scores - Tight Experiment

Panel A: Numeracy

Panel B: Communication



Notes: **Figure A1 shows the distribution of standardized assessment scores in the tight experiment.** Panel A displays the fraction of standardized numeracy scores in bins. Panel B displays the same for communication scores.

Big Experiment In the big experiment we measured six skills: communication, numeracy, concept formation, focus, grit, and planning. Further details on the implementation are in Appendix D.

Table A3: Joint Distribution of Skill Quintiles - Tight Experiment

Numeracy quintile	Communication quintile				
	Bottom	Lower middle	Middle	Upper middle	Top
Bottom	0.00%	7.19%	6.47%	4.68%	4.68%
Lower middle	11.51%	0.00%	7.19%	7.91%	9.35%
Middle	3.96%	2.52%	0.00%	4.68%	2.88%
Upper middle	2.52%	4.32%	3.96%	0.00%	7.19%
Top	0.72%	2.16%	0.72%	5.40%	0.00%

Notes: **Table A3 shows that there is a lot of variation in quintiles across numeracy and communication skills in the tight experiment.** Sample is restricted to the 278 jobseekers with a unique comparative advantage in skills, hence the main diagonal is zero.

Table A4: Correlation Between Skill Quintiles - Tight Experiment

	Numeracy (1)	Communication (2)	Concept formation (3)
Panel A: Restricted sample			
Numeracy	1.000	-0.004	0.284
Communication		1.000	0.216
Concept formation			1.000
Panel B: Full sample			
Numeracy	1.000	0.306	0.375
Communication		1.000	0.326
Concept formation			1.000

Notes: **Table A4 shows positive but moderate Pearson correlation coefficients between the skill quintiles in the tight experiment.** Panel A shows results for the sample with a clear comparative advantage between numeracy and communication. Therefore, the correlation between those two skills is zero in this subsample. Panel B shows results for the full sample.

The *communication* assessment is equivalent to the assessment used in the tight experiment. It measured listening and reading comprehension skills. The *numeracy* assessment measures practical arithmetic skills and pattern recognition. The first part of the assessment was the same as the numeracy assessment in tight experiment. This part assess jobseekers' skills needed for a cashier position. The second part of the numeracy assessment was developed by an adult education provider (www.mediaworks.co.za) and measures high school mathematics skills including calculations involving money, time, areas and quantities. For measuring *concept formation* skills, we used a very similar test to Raven's matrices (Raven & Raven, 2003). The *focus* measure is a computerized, color based Stroop task (Stroop, 1935). It evaluates jobseekers' inhibitory control, controlling one's attention and guiding thought and action to achieve a goal (Diamond, 2013; Posner & DiGirolamo, 1998). We rely on Duckworth's self-reported 8-item scale to assess jobseekers' level of *grit* (Duckworth, 2016). It rates jobseekers' willingness to work on difficult tasks and persevere to achieve long-term goals. Finally, *planning* measures how jobseekers are able to search for relevant information and anticipate the consequences of actions. The assessment adapts the Hit 15 task in Gneezy et al. (2010), in which the computer and the participant take turns to add one, two, or three point to a point basket. The party wins, whose action leads to a score of 15 in the point basket.

Appendix Table A5 shows the correlation matrix between different skill terciles in the big experiment. Scores are weakly correlated across assessments, with pairwise correlations between 0.09 and 0.44. Hence, the assessments horizontally differentiate candidates based on their relative skills rather than only ranking or vertically differentiating them in a single dimension of skills.

Table A5: Correlation Between Skill Terciles - Big Experiment

	Concept formation (1)	Communication (2)	Numeracy (3)	Grit (4)	Planning (5)	Focus (6)
Concept formation	1.000	0.298	0.435	0.098	0.214	0.189
Communication		1.000	0.331	0.095	0.213	0.173
Numeracy			1.000	0.139	0.266	0.159
Grit				1.000	0.090	0.047
Planning					1.000	0.173
Focus						1.000

Notes: **Table A5 shows positive but moderate Pearson correlation coefficients between the skill terciles used in the big experiment.**

B.2 Employers Value These Skills

In this section we provide support for the importance of skills we measured in the labor market.

First, a subset of the assessments (communication, concept formation, and numeracy) has been used by our partner to screen jobseekers in the past. Our partner had been contracted to screen roughly 160,000 prospective workers using these assessments by 2016 by firms in South Africa. Based on this, we presume that the information content of the assessments is valuable for the firms. This, however, does not mean that assessment results are the only information firms use in their hiring decisions. Additionally, we do not assume that firms use the information at their disposal optimally, and thus, we do not claim that these tests are the best predictors for jobseekers' productivity.

Second, we use an incentivized choice experiment to show that firms vary in their valuation of communication and numeracy and value both highly relative to some forms of education. For this data collection, we recruited 67 firms soon after the big experiment by going door-to-door in areas of Johannesburg where most of the jobseekers in the big experiment lived. 81% of firms are in the retail or hospitality sectors, where many jobseekers in both experiments applied for jobs. Recruited firms have a mean size of 15 workers, half of whom are in entry-level roles, and planned to hire an average of 4 new entry-level workers in the next year. Carranza et al. (2022) describe the sample and data collection in detail and use this data collection to show that there is substantial variation in firms' preferences over skills.

Importantly for our current argument, we measured the preferences of these firms over the six skills used in the big experiment relative to each other and to additional education. Each firm was asked to rank multiple jobseeker profiles with different levels of skills and with or without a post-secondary diploma, all with completed secondary school. To incentivize the choices, firms' rankings were used to match them with jobseekers with specific skill profiles from Harambee's database, in a similar spirit to Kessler et al. (2019).

Table A6 shows the average ranking of numeracy, communication and education over the 67 firms. There are six different possible rankings of these three elements, each shown in a row. The shares of firms in these bins are shown in Column 4. Column 6 collapses these shares based on the most important skill. In this column we see that 57% of firms prefer a candidate with top-tercile numeracy skills, 34% prefer a candidate with top-tercile communication skills and only 9% firms prefer a candidate with a relatively better educational achievement, that is a candidate with a diploma but with only middle-tercile communication and numeracy skills.

Table A6: Firms' Preference Ranking Over Communication Skills, Numeracy Skills, and Formal Education

	Top (1)	Middle (2)	Bottom (3)	Share (%) (4)	Most Important Skill (5)	Share (%) (6)
1	Num	Comm	Educ	52.24	Numeracy	56.72
2	Num	Educ	Comm	4.48		
3	Comm	Num	Educ	28.36	Communication	34.33
4	Comm	Educ	Num	5.97		
5	Educ	Num	Comm	1.49	Education	8.96
6	Educ	Comm	Num	7.46		

Notes: **Table A6 shows that firms, on average, value applicants' numeracy (Num) and communication skills (Com) more than a one-year post-secondary certificate (Educ) and vary in their relative ranking of communication and numeracy skills.** The results are based on an incentivized choice experiment with 67 small and medium sized businesses in Johannesburg. The rows represent all the possible rankings. Column 4 shows the share of firms who chose the respective rankings. Cols. 5 and 6 do the same but collapses rows according to the most important skill.

B.3 Observability of Skills and Firm's Value of Applicant Skill Match

As part of the tight experiment, we conducted a measurement exercise to show that firms partially observe assessed skills and value applicants whose skill profile matches their job requirements. During the job search workshop, we asked jobseekers to choose between applying for a real communication or numeracy job at a firm that hires for a range of entry-level roles, including call center and data capture jobs. (See Appendix D.1 for details about the task.) Jobseekers prepared a resume and a cover letter during the workshop, both designed for general use rather than tailored to these specific jobs. Two mem-

Table A7: Employer Evaluation of Job Applicants Based on Skills

	Mean			Difference			Obs.
	Aligned	Non-aligned	SD	Δ	Δ/SD	p($\Delta = 0$)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Skill levels							
Skill (1-5)	2.93	2.79	0.66	0.15	0.22	0.01	277
Panel B: Job-related evaluation							
Overall score (1-5)	3.00	2.84	0.86	0.16	0.18	0.05	277
Interview invitation (dummy)	0.43	0.34	0.49	0.09	0.18	0.07	277

Notes: **Table A7 shows that the HR team of an employer can observe jobseekers' skills and evaluates applicants more highly if their assessed comparative advantage in skills matches the job's requirements.** One pair of the job choice task advertisements was from a firm that hires for a range of entry-level roles. We submitted the jobseekers' coverletter and resume to the firm based on which two members of the firm's HR team evaluated every applicant for both jobs. Evaluators rated the jobseekers' skill levels (Panel A) as well as their general suitability for the job (on a scale from 1 to 5) and whether they would invite the candidate for an interview (Panel B). We show the mean outcomes across evaluators for the skill/ job that is aligned with the jobseekers' comparative advantage in col. 1; and the outcomes for the misaligned skill / job in col. 2. The pooled standard deviation of the measures in col. 1 and 2 are in col. 3. Col. 4 shows the difference between cols. 1 and 2. Col. 5 shows this difference in terms of standard deviations. Col. 6 shows the p-value associated with a test of equality across cols. 1 and 2. Col. 7 shows the number of observations.

bers of the firm's HR team evaluated every applicant for both jobs based on their CV and generic cover letter. Evaluators were blind to which applicant applied for which job and were not shown applicants' skill assessment results. We received data on the evaluators' assessment of each applicant's communication skills, numeracy skills, and suitability for each type of job, as well as whether the jobseekers were recommended being interviewed for each job.

This measurement exercise shows that the firm's HR team could identify levels of jobseekers' skills. Table A7, panel A, column 5 shows that the HR team's assessments of skills are positively but not perfectly correlated with our measures of skill: they assigned a 0.22 standard deviation higher score to the skill we assess as higher ($p = 0.01$). The HR team's applicant ratings are also correlated with our measures of skill: they rate the applicant as 0.18 standard deviations more suitable for the job aligned with that applicant's comparative advantage (panel B, row 1, column 6). HR managers were also 9 percentage points more likely to recommend interviewing the candidate for the job aligned with that applicant's comparative advantage (panel B, row 2, column 6). This is a 26% increase relative to a 34% interview recommendation rate in the non-aligned job. These patterns show that skills are at least partly observable to the firm even when jobseekers could not tailor their resumes or cover letters to the specific role, suggesting the possibility of greater observability in natural job search where jobseekers can tailor their applications.

The evidence that jobseekers have different skills, firms value these skills, but differ in which skill they value more, and that firms can at least partly observe skills, suggests that redirecting jobseekers' search towards jobs that match their comparative advantage in skills has the potential to improve their labor market outcomes.

C Skill Beliefs

C.1 Skill Beliefs Measurement

This appendix serves the purpose of describing and justifying our choices of our main belief measures and of detailing the measurement of beliefs for the interested reader. Overall, our empirical results are extremely robust to the choice of belief measure, however, there are slight conceptual differences in measurement across and within experiment that we clarify in this appendix.

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 measure beliefs about three skill domains: numeracy, communication, and concept formation skills (see Appendix B for a description of how we assess those skills). Further, we distinguish two types of skill beliefs: beliefs about *assessment results* and beliefs about *general* skills. Table A8 contains the exact wording of our skill belief elicitation questions for both. In practice, these two belief measures are strongly positively correlated within skill, suggesting that jobseekers view the assessments as relevant to their general skills.

Before eliciting beliefs, we define the skills in the following way: "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 beliefs about general skills as beliefs about one's skills relative to the reference group in a specific domain, abstracting from specific assessment results. We see these skill beliefs as being *most relevant for search decisions* as they capture general, labor market relevant skills that are not affected by the idiosyncrasy of the performance of our assessments. Put differently, what matters to employers is not how well one does on a specific assessment but rather how well one is able to use a skill consistently at work relative to others. Thus, we use these beliefs to define our preferred measure of aligned comparative advantage beliefs.

In the tight experiment, we measure beliefs about general skills before and after treatment. (See Figure A2 that summarizes the experimental design.) We use general skill beliefs in the tight experiment for the descriptive statistics in the summary Table 1. We use general skill beliefs to define our main belief outcome measures in the tight experiment (Table 2) and the corresponding appendix tables (Tables A9, A13, A20, A26, A35 (columns 1 to 4), and A47).

In the big experiment, we only measure beliefs about general skills for a random subsample of participants after treatment. As a robustness check, we estimate treatment effects using the CA beliefs defined using general skills. We find that treatment increases aligned CA beliefs by 7.6 percentage points (pp) ($p = 0.003$) and 10.8pp ($p < 0.001$) - smaller than the effects on beliefs about assessment results but qualitatively similar and still highly significant.

Beliefs about Assessment Results: We measure beliefs about the relative placement in the assessments candidates completed in both experiments. We consider these beliefs as a *proxy for the information content of the interventions* for each individual because the treatment provides information about relative assessment results (following Haaland et al., 2023). Jobseekers who have inaccurate beliefs about their relative assessment results after taking the assessment will learn that their actual performance differed from their beliefs. Conversely, jobseekers with initially accurate beliefs should not update their beliefs about their assessment results (though they might still become more certain about their beliefs). We hypothesize that individuals with initially accurate beliefs about their comparative advantage in the assessment should exhibit stronger treatment effects. Hence, we estimate heterogeneity using a dummy variable indicating accurate baseline beliefs about jobseekers' comparative advantage in the assessments throughout both the tight and big experiment (equation 4). On average, 46% of jobseekers have aligned comparative advantage beliefs using this measure. The results are similar when we use the measure of general skills for heterogeneity instead of assessment beliefs because the two measures of comparative advantage beliefs are highly correlated ($\rho = 0.68$). Similarly, regressing domain specific general skill beliefs on beliefs about assessment results produces coefficients of 0.39-0.52 across the two skills, with or without controls for assessment results and demographic characteristics.

In the tight experiment, we asked about beliefs about assessment results after the assessment but before the treatment for the whole sample (see Figure A2). We use these

Table A8: Measurement of beliefs about comparative advantage

Description	Survey question
Panel A: Tight experiment	
General skills belief, pre-treatment (most likely quintile)	Think about 100 people who are job-seekers 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?
General skills belief, post treatment (most likely quintile)	Think about 100 people who are job-seekers 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?
Assessment results belief, pre- and post-treatment (most likely quintile)	Think about 100 people who are job-seekers 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	
General skills belief, 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?
Assessment results belief, pre-treatment (most likely tercile)	Now think about all the people who are in the room with you. They are all job-seekers 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?
Assessment results belief, 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 A8 displays the exact wording of our skill belief measures.**

measures for all heterogeneity analysis by baseline beliefs in the tight experiment main Tables 2 and 3 and all the appendix tables that report heterogeneity by skill beliefs in the tight experiment. We also 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. We did not ask the control group again to avoid asking the same question twice in a short amount of time without providing additional information.

In the big experiment, we ask both the control and treatment group twice about their beliefs about their assessment results. (See Figure A3.) First, we ask them at baseline after they took the assessments but before they could receive the treatment. Second, we ask them again at endline about three months after treatment. Given that we only measured general skills for a subsample of jobseekers after treatment, we use assessment specific beliefs both as outcomes and as heterogeneity variables for the big experiment throughout the paper and appendix.

C.2 Skill Belief Descriptive Statistics and Treatment Effects

In this section we provide descriptive statistics about jobseekers' skill beliefs and their relationship with high school exam results, expectations for job search outcomes, and job search activities. We then present addition treatment effects on skill beliefs.

Descriptive statistics: Table A9 displays correlations between prior beliefs about skills and assessment results in the tight experiment. We find that numeracy beliefs are strongly correlated with assessment results but communication beliefs are not significantly correlated with the assessment results. Table A10 displays correlations between self-reported grades in English and mathematics in the high school leaving exam (matric) and assessment results and beliefs. Columns 1-3 indicate that the matric results correlate with our assessments and the score differences between the respective subjects on those exams positively correlate with the measured comparative advantage in the tight experiment.

We also find that the high school leaving exam results predict jobseekers' beliefs (Table A10, Columns 4-7). However, we do not find that any baseline demographic variables, including labor market exposure, meaningfully predict having aligned comparative advantage beliefs at baseline in the tight experiment (Table A11). Table A12 shows that beliefs do not get updated over time in the control group in the big experiment. Overall, this set of results suggests that jobseekers have limited capacity to learn about their comparative advantage in this labor market.

Finally, we show additional evidence that jobseekers' beliefs about skills are related to perceptions of the labor market and search direction in the big experiment. Table A14 shows the correlation between jobseekers' beliefs about skills and beliefs about the returns to search using the control group in the big experiment. Table A15 shows that jobseekers' beliefs about their comparative advantage correlate positively and significantly with search direction in the control group of the big experiment.

Additional treatment effects: Underconfident skills are more likely to update than overconfident skills. Table A13 displays treatment effects on dummies indicating the fraction of under and overconfident skill beliefs. We find that the relative reduction in underconfident beliefs (30.6% and 28.2% of the control mean in the tight and big experiment respectively) is significantly larger than the relative reduction in overconfident beliefs (3.5% and 21.9% of the control mean in the tight and big experiment respectively). This aligns with lab evidence on asymmetric updating (Zimmermann, 2020).

Jobseekers might update their skill comparative advantage beliefs through two channels: because they update 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 got right on each assessment, after taking the assessments but before any treatment. We construct a dummy for **accurate score ranking**, equal to one if and only if the jobseeker correctly identifies the assessment where they get more questions correct. We then identify jobseekers who have an accurate score ranking and a misaligned comparative advantage belief. These jobseekers can only update their comparative advantage beliefs due to channel (2); channel (1) is largely shut down for them. This admittedly small subset of jobseekers substantially update their comparative advantage beliefs ($p = 0.001$). This suggests that learning about the distribution of other jobseekers' skills at least partially explains the updating of comparative advantage beliefs.

Table A9: Association Between Assessed and Believed Skill Quintiles - Tight Experiment

	Skill quintile beliefs	
	Numeracy	Communication
	(1)	(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 A9 shows that assessed numeracy quintiles correlate with numeracy quintile beliefs but assessed communication skills do not correlate with communication beliefs.** No control variables are included. Robust standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A10: Association of High School Graduation Exam Results with Assessed and Believed Skills - Tight Experiment

	Asessed skills			Beliefs about skills			
	Skill quintile		Comp. adv.	Skill quintile		Comp. adv.	
	Num.	Com.	Num.	Num.	Com.	Num.	Com.
	(1)	(2)	(3)	(4)	(5)	(6)	(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 A10 shows that self-reported grades in English and mathematics in the secondary school leaving exam (matric) correlate positively with jobseekers' assessed skills and their baseline beliefs about skills.** Cols. 1, 2, 4, and 5 correlate math and English matric scores with assessed skill quintiles (cols. 1 and 2) and beliefs about skill quintiles (cols. 4 and 5). Matric grades are rescaled to range from 0 to 1. So the coefficient in column 4 row 1 shows that moving from the lowest to highest possible matric grade in math is associated with a 1.78 higher believed numeracy quintile. No further control variables are included. Robust standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A11: Association Between Baseline Aligned Comparative Advantage Belief and Other Baseline Characteristics - Tight Experiment

	Aligned CA belief	
	(1)	(2)
Age	-0.012	-0.004
	(0.009)	(0.008)
Female	0.048	0.024
	(0.064)	(0.059)
Has completed secondary education only	-0.008	-0.024
	(0.114)	(0.109)
Has a post secondary certificate	-0.059	-0.071
	(0.133)	(0.126)
Has a post secondary diploma	0.095	0.058
	(0.140)	(0.135)
Has a post secondary degree	-0.208	-0.208
	(0.140)	(0.140)
Employed in any form at baseline	0.030	0.026
	(0.068)	(0.063)
Total work experience at baseline (years)	0.017	0.013
	(0.013)	(0.012)
Numeracy assessment score (%)		-0.007**
		(0.003)
Communication assessment score (%)		0.012***
		(0.002)
Constant	0.739**	0.076
	(0.244)	(0.283)
# jobseekers	278	278
R2 (not adjusted)	0.031	0.177
p: all coefficients = 0	0.233	0.000

Notes: **Table A11 shows that baseline comparative advantage beliefs are uncorrelated with demographic characteristics.** It displays coefficients from regressions with baseline data of an indicator for aligned comparative advantage belief on age, gender, education categories (omitting less than completed high school), employment, and total work experience and, in the second column only, communication and numeracy assessment scores. Robust standard errors shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A12: 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 A12 shows that the misalignment of skill beliefs persists over time in the control group of the big experiment.** Estimation is at the survey round times jobseeker level and is restricted to the control group. Cols. 1 to 3 show results for aligned beliefs about CA. Cols. 4 to 6 show results for the fraction of aligned skill beliefs. There is no heterogeneity by whether individuals exerted above median search effort as measured on the search effort index at endline described in Table A38 (cols. 2, 5). Similarly, there is no heterogeneity by whether an individual worked in the last seven days elicited at endline (cols. 3, 6). Robust standard errors are in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A13: Treatment Effects on Over- and Underconfident Beliefs - Both Experiments

	Tight experiment (quintiles)		Big experiment (terciles)	
	Underconfident (1)	Overconfident (2)	Underconfident (3)	Overconfident (4)
Treatment	-0.057***	-0.022	-0.043***	-0.101***
	(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	0.187	0.629	0.150	0.461
Observations	278	278	4205	4205

Notes: **Table A13 shows that underconfident beliefs are more likely to be updated than overconfident beliefs.** It displays treatment effects on dummies indicating if the skill beliefs of the workseeker are underconfident (cols. 1 and 3) or overconfident (cols. 2 and 4). Cols. 1 to 2 show results for the tight experiment. Cols. 3 and 4 show results for the big experiment. The effect sizes relative to the mean suggest that underconfident beliefs are more likely to update. All specifications include randomization block fixed effects. Controls further include prespecified baseline covariates described in footnotes 15 and 21. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A14: 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 A14 shows that jobseekers' beliefs about skills correlate positively with 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 purchasing power parity terms. Robust standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A15: Association Between Skill Beliefs and Search Direction - Big Experiment Control Group

	Target numeracy jobs		Target communication jobs	
	(1)	(2)	(3)	(4)
Believed numeracy CA	0.100***	0.091***	-0.094***	-0.084***
	(0.018)	(0.018)	(0.021)	(0.022)
Believed communication CA	-0.079***	-0.067***	0.087***	0.088***
	(0.021)	(0.020)	(0.024)	(0.023)
Numeracy CA		0.021		-0.012
		(0.018)		(0.022)
Communication CA		-0.060***		0.045**
		(0.019)		(0.023)
Control mean	0.222	0.222	0.471	0.471
Observations	2183	2179	2183	2179
Controls	No	Yes	No	Yes

Notes: **Table A15 shows that jobseekers' beliefs about comparative advantage correlate positively with skill-directed search in the control group of the big experiment.** Dependent variables are dummies indicating that jobseekers rate the respective skill as most important for the jobs that they applied to in the last 30 days, measured at endline. Independent variables are dummies for beliefs indicating that individuals have a clear comparative advantage in each skill. Even columns additionally control for age, gender, having worked in a wage job, as well as dummies for three education categories, and measured comparative advantage. Robust standard errors are in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D Protocol and Intervention Details

This appendix summarizes the protocol and intervention details for the tight and the big experiment. Other relevant materials are in Appendix B (skill measurement), and in Appendix C (skill belief measurement).

D.1 Tight Experiment

The data collection for the tight experiment ran between August and October in 2022 in downtown Johannesburg. We recruited 373 participants using the user database of our implementation partner. We contacted users who were active on SAYouth.mobi in the past month, said they were residing in Gauteng province, completed at least a high school leaving exam and were weakly less than 35 years old. Using this contact list, surveyors called potential participants, and after a short screening[34], invited them for a daylong job search workshop in the city center. We offered 150 Rand (approximately 21 USD) airtime for compensation. The structure of data collection is depicted on Figure A2 which we will detail below.

When participants arrived at the venue, participants 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 to the participants 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, and also measured risk and time preferences. This survey was followed by three assessments; participants completed the communication, numeracy, and cognitive performance tests in order. Participants had 30 minutes for each of the communication assessment and numeracy assessment and 15 minutes to complete the cognitive performance test[35]. 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 1) 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 any questions from the surveyor. On control days, participants still viewed a minimally mod-

[34] The screening questions confirmed if the user is currently looking for a job, not a full-time student, and has not participated in the tight experiment before (even during the piloting).

[35] Section B.1 describes the skills and tests in more detail.

ified version of the video that omitted the explanation of the results. The scripts and the video was 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 A16).

Table A16: Balance Table - Tight Experiment

	Restricted sample					Full sample				
	Control (1)	Treatment (2)	Δ (3)	$p(\Delta=0)$ (4)	N (5)	Control (6)	Treatment (7)	Δ (8)	$p(\Delta=0)$ (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 A16 shows that covariates are balanced across treatment groups in the tight experiment.** CA stands for comparative advantage in skills. Cols. 1 to 5 show results for the sample of individuals with a clear CA in skills. Cols. 6 to 10 show results for the full sample of individuals including those without unique CA in skills. P-values are obtained conditional on randomization block fixed effects. Winsorized variables (w) are winsorized at the 99^{th} percentile. All monetary values are in 2021 USD purchasing power parity terms.

After the treatment and a lunch break, the surveyors elicited 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. Each job pair contained jobs with opposite skill demand (one communication- and one numeracy-heavy job) based on 13 recruiters' prior evaluation. The job titles 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. The job titles were all entry-level jobs that did not require certifications or equipment to ensure that all participants could reasonably apply for them (see Table A17 for the job title pairs). The job descriptions and

Table A17: 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

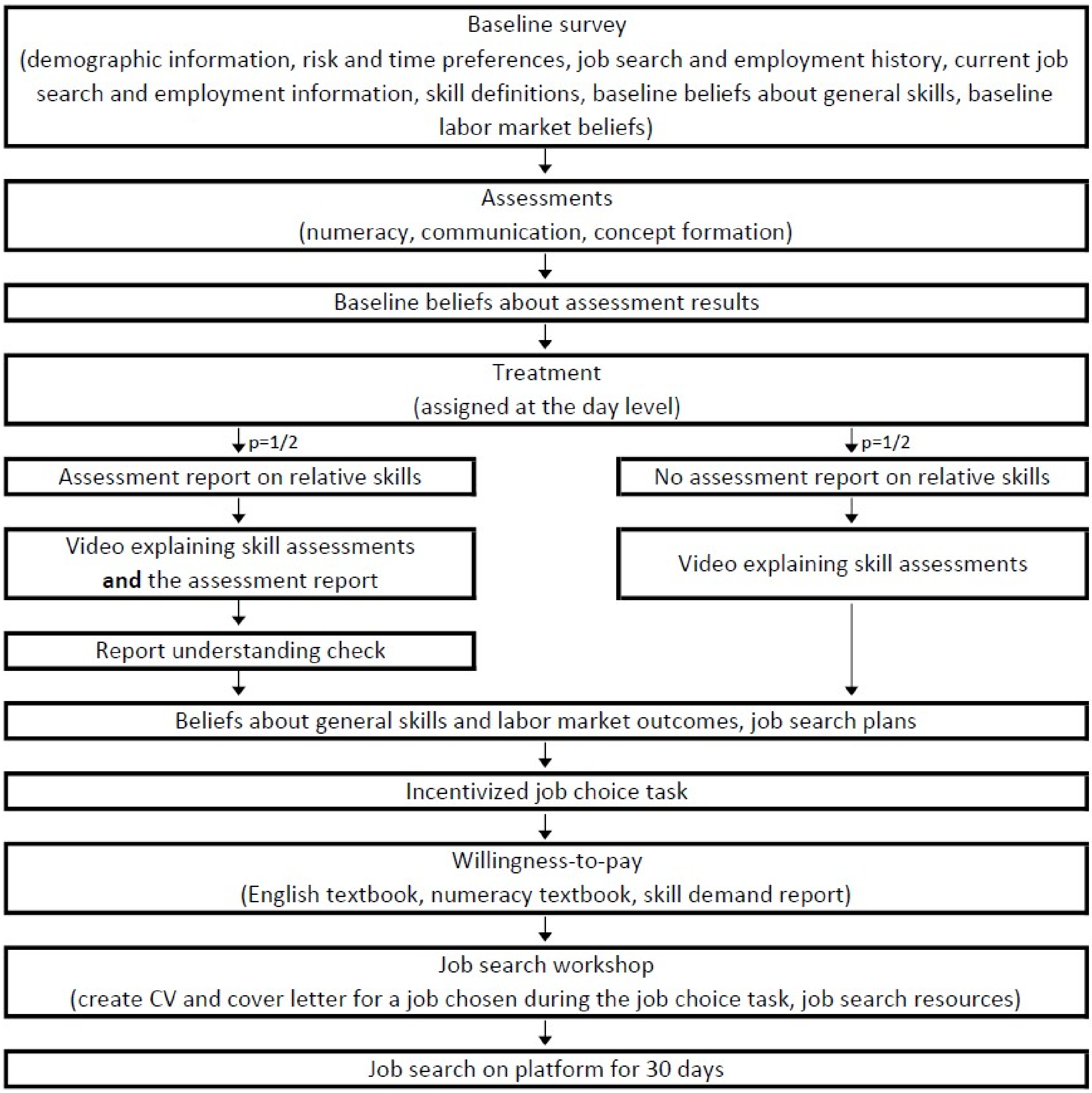
Notes: **Table A17 shows the job titles for numeracy and communication-heavy jobs in each job pair in the job choice task in the tight experiment.** Each job pair contains one relatively numeracy-heavy and one communication-heavy job title. The pairs were created by matching jobs with different skill 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.

their layout followed the design and style of the SAYouth.mobi platform. The job descriptions were presented in a printed booklet side-by-sideto allow for easy comparisons (see Figure 2). 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 (numeracy or communication) that the job required. When participants completed their choices, we elicited their beliefs and experience related to each job in 5 out of 11 pairs.

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, at the end of the workshop participants received a list of job titles 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 an incentivized multiple price lists for three products: for self-study materials to improve participants' communication skill and numeracy skill respectively, and a document that reveals the skill demand of a set of common jobs as per the rating of HR experts. Par-

Figure A2: Tight Experiment Design



Notes: **Figure A2 shows the experimental design of the tight experiment.**

ticipants complete a practice round before the elicitation and the elicitation is simplified: we are asking a sequence of pairwise choice between a monetary value and the product itself. Further details of the protocol and the results are available at the following link: https://bit.ly/3E34oYH.

After the willingness-to-pay module, we collected information typically found in CVs. Using this information we prepared a CV for each participant and delivered them to the participants after the workshop. We also submitted participants' CVs to the job that they preferred among the real jobs in the job 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 application email to the employer of the real job in the job choice task that they preferred. The message, along with participants' CV that we created, were both delivered to the respective employers. After this part, participants completed the check out procedure and received their compensation.

D.2 Big Experiment

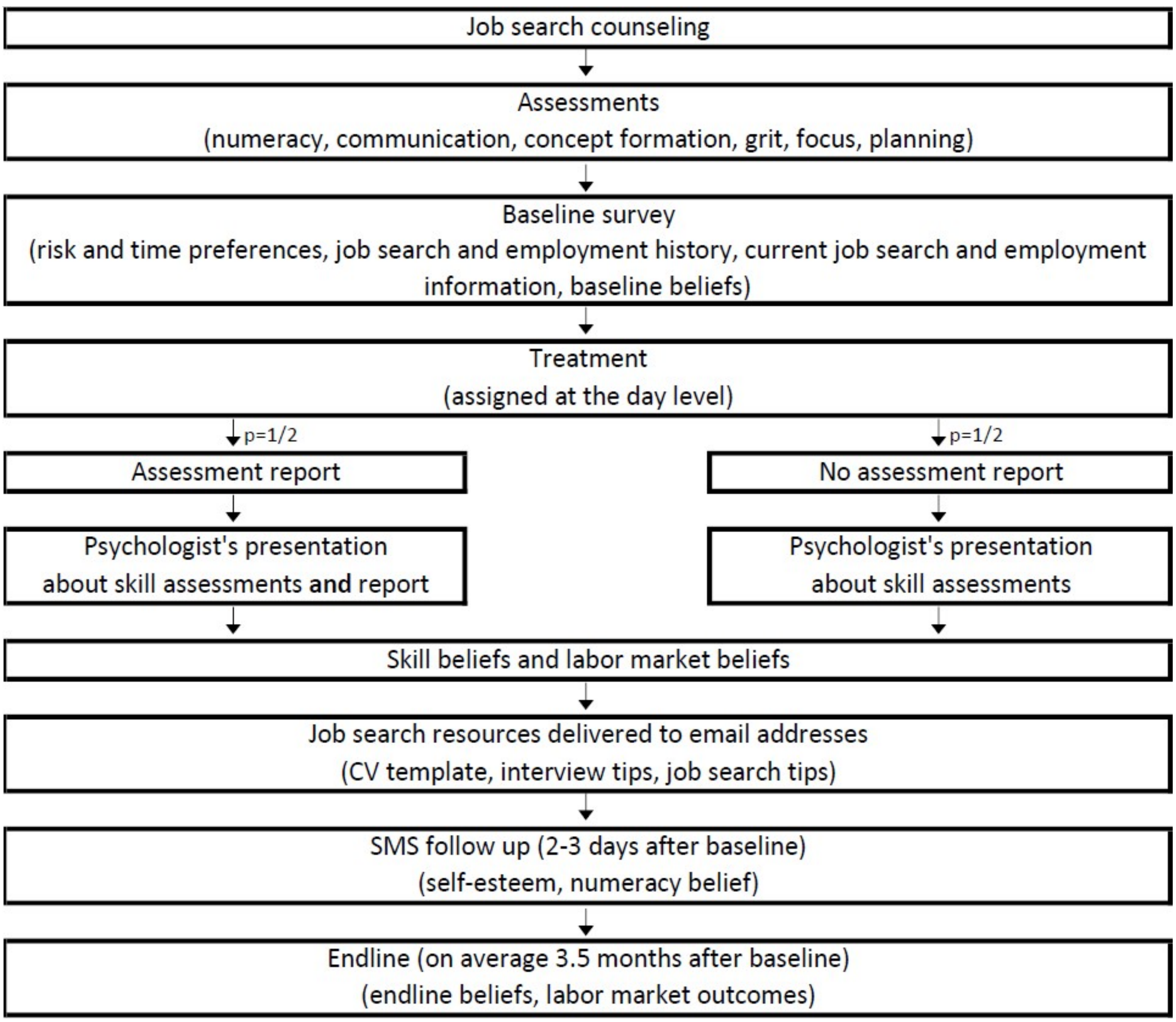
The recruitment, data collection procedure and intervention of the big experiment is very similar of the tight experiment. The specific details are described in Carranza et al. (2022) and also shown inn Figure A3, in this section we highlight the main differences between the big and tight experiment.

The big experiment took place in the same labor market, but earlier in time, in 2016/17. We again used the contact list of the Harambee Youth Employment Accelerator. Since the job search platform was not yet developed, we recruited young jobseekers (aged 18–29) who have completed secondary school, have at most 12 months of formal work experience and were from disadvantaged backgrounds. The venue of the data collection was the Harambee office at the time in downtown Johannesburg.

Similar to the tight experiment, jobseekers completed surveys and assessments. Treatment was randomized at the day level and the intervention consisted of the receiving a report. The report (Figure A4) showed the performance results for six skills (as opposed to only numeracy and communication in the tight experiment) and the results were reported in terciles and not quintiles. The report was handed to participants privately in an envelope and was followed by a briefing from industrial psychologists. The content of the briefing is akin to that in the tight experiment video. The materials used for the briefing are available at https://bit.ly/44oOp1v and the script is available at https://bit.ly/3YLaDK6. Baseline covariates are balanced across treatment arms (Table A18).

Two follow-up surveys were conducted. 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 sample was successfully interviewed at endline and we do not find differential attrition based on treatment (Table A19).

Figure A3: Big Experiment Design



Notes: **Figure A2 shows the experimental design of the big experiment.**

Figure A4: 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.

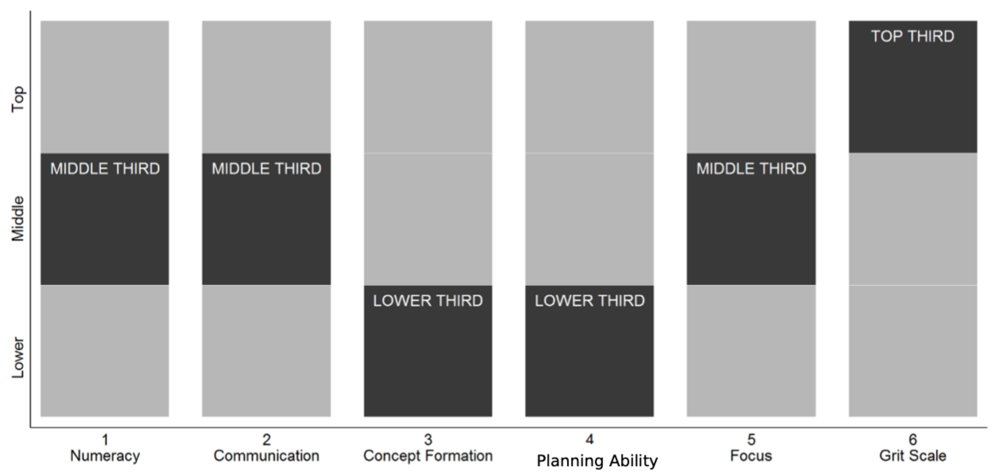
1. The Numeracy tests measure various maths abilities. Your score is the average of the two maths tests you did today at Harambee.
2. The Communication test measures English language ability through listening, reading and comprehension.
3. The Concept Formation test measures the ability to understand 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 A4 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.

Table A18: 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 A18 shows that covariates are balanced across treatment arms for the big experiment.** CA stands for comparative advantage. Cols. 1 to 5 show statistics for the baseline sample. Cols. 6 to 10 show statistics for the sample of individuals reached at endline. P-values are obtained conditional on randomization block fixed effects. Winsorized variables (w) are winsorized at the 99^{th} percentile. All monetary values are in 2021 USD purchasing power parity terms.

Table A19: Treatment Effects on Completing Endline Survey - Big Experiment

	Interviewed at endline	
	(1)	(2)
Treatment	-0.005	-0.003
	(0.006)	(0.006)
Control mean	0.960	0.960
Observations	4389	4389
Controls	No	Yes

Notes: **Table A19 shows that attrition is low and balanced across treatment groups in the big experiment.** Dependent variable is a dummy indicating whether an individual was successfully recontacted for the endline survey. Column 2 includes pre-specified baseline covariates described in footnote 21. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

E Additional Treatment Effects

E.1 Robustness Checks

In this section we first evaluate robustness for sample restrictions. In the tight experiment we restrict the sample to jobseekers who have a clear comparative advantage. Our treatment effect results about skill belief (Table A20) and search direction (Table A21) are robust for using the full sample of the tight experiment. In Table A22 we present our main results for the tight and big experiment side-by-side, showing consist results across the two experiments.

Next, we analyze robustness for different variable definitions. Using a continuous measure for belief alignment does not change our results (Table A23). Likewise, different transformations of earnings leave our conclusions intact (Table A24), which is supported by Figure A5: the treatment effects on conditional and unconditional earnings are not driven by outliers.

Table A20: Treatment Effects on Skill Beliefs - Tight Experiment Including Jobseekers Without Unique Comparative Advantage

	Aligned CA belief		Fraction aligned beliefs	
	(1)	(2)	(3)	(4)
Treatment	0.120***	0.180***	0.089***	0.064***
	(0.037)	(0.055)	(0.020)	(0.018)
Treatment × Aligned CA belief (bl)		-0.130		0.057
		(0.089)		(0.041)
Aligned CA belief (bl)		0.447***		-0.039
		(0.081)		(0.037)
Treatment effect:		0.050		0.121***
Aligned CA belief (bl)		(0.061)		(0.036)
Control mean	0.446	0.446	0.171	0.171
Observations	368	368	368	368

Notes: **Table A20 shows that treatment effects on beliefs about skills are robust to including people without a clear comparative advantage in skills.** CA stands for comparative advantage. Columns indicate different outcome variables: a dummy indicating whether the jobseeker's beliefs about their CA in skills are aligned with the assessment results (cols. 1-2), and the fraction of skill quintile beliefs that are aligned with the assessment results (cols. 3-4). Cols. 2 and 4 show treatment effect heterogeneity by whether individuals had aligned CA beliefs at baseline. Control variables include variables described in footnote 15 and a dummy variable if the jobseeker has no comparative advantage. This last dummy variable is a prespecified control variable but cannot be used when we restrict the sample to jobseekers with clear comparative advantage. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A21: Treatment Effects on Aligned Search Direction - Tight Experiment Including Jobseekers Without Unique Comparative Advantage

	Aligned search index		% aligned (job choice)		Δ % aligned platform apps		Δ SMS click rate		Δ planned apps (w)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment	0.240***	0.436***	0.032	0.062**	0.052***	0.062**	0.071	0.122*	1.121	3.363***
	(0.079)	(0.094)	(0.025)	(0.026)	(0.017)	(0.030)	(0.046)	(0.063)	(0.878)	(1.140)
Treatment × Aligned CA belief (bl)		-0.446***		-0.070**		-0.023		-0.113		-5.190**
		(0.156)		(0.029)		(0.060)		(0.098)		(2.162)
Aligned CA belief (bl)		0.506***		0.111***		0.011		0.070		6.113***
		(0.121)		(0.026)		(0.035)		(0.076)		(1.872)
Treatment effect:		-0.011		-0.008		0.040		0.009		-1.827
Aligned CA belief (bl)		(0.118)		(0.029)		(0.039)		(0.070)		(1.613)
Control mean	-0.000	-0.000	0.416	0.416	0.005	0.005	-0.024	-0.024	3.272	3.272
Observations	368	368	368	368	368	368	368	368	368	368

Notes: **Table A21 shows that treatment effects on aligned search direction are robust to including people without a clear comparative advantage in skills.** The table includes the full sample of workshop participants. 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: an index constructed as the variance-covariance weighted average of the search alignment measures displayed in cols. 3 to 10 (Anderson, 2008) (cols. 1-2), the percentage of 11 incentivized job choices that are aligned with the measured CA of the jobseeker (cols. 3-4), the difference between the percentage of aligned and non-aligned applications on the online job search platform SAYouth.mobi (cols. 5-6), the difference in link click rates between aligned and non-aligned jobs sent to job seekers via SMS (cols. 7-8), and the difference between aligned and non-aligned planned applications for the 30 days after the workshop (cols. 9-10). Even columns show heterogeneity by whether individuals have aligned CA beliefs at baseline. Control variables include variables described in footnote 15 and a dummy variable if the jobseeker has no comparative advantage. This last dummy variable is a prespecified control variable but cannot be used when we restrict the sample to jobseekers with clear comparative advantage. Winsorized variables (w) are winsorized at the 99th percentile. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A22: Robustness Checks for Main Outcomes - Comparing Tight and Big Experiments

	Tight experiment		Big experiment			
	Aligned CA belief	Aligned search index	Aligned CA belief	Aligned search direction	LM quantity index	LM quality index
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Robustness for inference						
Treatment	0.135	0.254	0.139	0.050	0.045	0.085
clustered p-value	(0.001)	(0.017)	(0.000)	(0.000)	(0.165)	(0.017)
wild-bootstrapped p-value	[0.005]	[0.009]	[0.000]	[0.000]	[0.395]	[0.005]
q-value	{0.002}	{0.011}	{0.001}	{0.001}	{0.029}	{0.011}
Panel B: Reweighting to match other sample						
Treatment	0.149***	0.254**	0.144***	0.054***	0.056*	0.086**
	(0.044)	(0.101)	(0.011)	(0.010)	(0.032)	(0.037)
Control mean	0.475	0.000	0.196	0.165	-0.000	0.000
Observations	278	278	4118	4205	4205	4206
Number of clusters	34	34	54	54	54	54

Notes: **Table A22 shows that the main results are robust to multiple hypothesis testing, using wild-bootstraps, and reweighting results to match the demographics across experiments. Panel A** shows p-values based on cluster-robust standard errors, p-values based on wild cluster bootstrap, and sharpened q-values that control the false discovery rate across tests following (Benjamini et al., 2006) for the treatment effect estimated across all outcomes. Wild-bootstrapped p-values are obtained using 10,000 draws. **Panel B** shows the same treatment effects using weights to match the observational characteristics of the other experimental sample. The covariates used to estimate weights are all variables displayed in the summary Tables 1 and A2 that are consistently measured in both experiments: dummies for being black, male, having only a high school leaving exam, having a post-secondary degree, or having a post-secondary certificate, having worked in the last seven days, and for having ever worked in a wage job, as well as earnings, the number of job applications, job search hours and expenditure, and job offers in the last seven days. Columns 1 and 2 show effects on the main outcomes of the tight experiment. Columns 3 to 6 show effects on earnings conditional on doing any work. Control variables for the tight experiment include variables described in footnote 15. Control variables for the big experiment are described in footnote 21. Standard errors clustered at the treatment-day level in parenthesis in panel B. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ in Panel B.

Table A23: Heterogeneous Treatment Effects by Non-binary 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

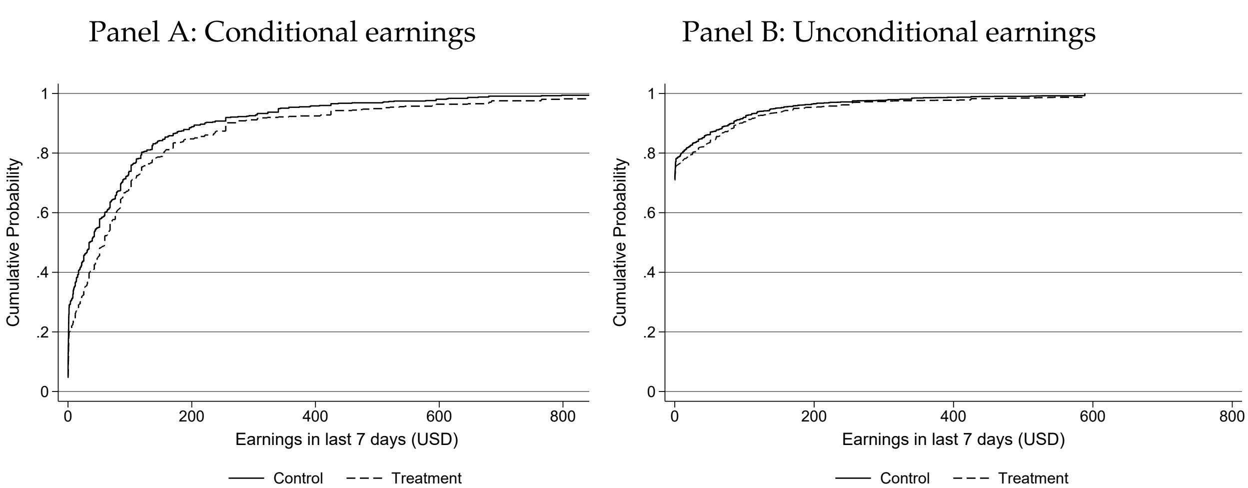
Notes: **Table A23 shows that treatment effect heterogeneity and treatment effects on beliefs are robust to using a non-binary measure of aligned comparative advantage beliefs in the tight experiment.** CA stands for comparative advantage in skills. The non-binary measure of comparative advantage beliefs captures the degree to which comparative advantage beliefs are aligned and ranges from 0 to 1. If comparative advantage beliefs are aligned it is set to 1. If beliefs are misaligned, the measure is defined as $1 - abs(\Delta bel\_skill_{i,j} + 1)/5$, where $\Delta bel\_skill_{i,j}$ is the difference between numeracy and communication beliefs (in quintiles). Hence $abs(\Delta bel\_skill_{i,j} + 1)$ shows the minimum number of quintiles the jobseeker needs to adjust their skill beliefs to align them with their measured comparative advantage beliefs provided they are misaligned. The second term of the measure is 1 for those who need to adjust their beliefs by 5 quintiles to have aligned comparative advantage (worst possible case), which leads to an overall alignment measure of 0. The mean of this measure is 0.83 and the standard deviation is 0.22. Columns indicate different outcomes: the aligned search direction index described in table 3 (cols. 1-2), a dummy indicating aligned comparative advantage beliefs (cols. 3-4), the non-binary measure of aligned comparative advantage beliefs at endline (cols. 5-6), and the overall alignment of skill beliefs with measured skills (cols. 7-8). This last measure is defined as one minus the average difference between beliefs about skill quintiles and assessed quintiles relative to maximum possible difference. A value of 1 indicates perfectly aligned beliefs while a value of zero indicates maximally misaligned beliefs. Even columns show heterogeneity by the baseline value of the non-binary measure of comparative advantage beliefs. Control variables are described in footnote 15. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A24: 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 A24 shows that treatment effects on earnings in the last seven days are robust to a range of functional forms.** Earnings are measured in 2021 USD PPP. Columns show effects on transformed unconditional earnings variable (cols. 1-5) and on transformed earnings variable conditional on doing any work (cols. 6-10). We display the effects on the raw variable (cols. 1, 6) and on its transformed values: the winsorized earnings at the 99^{th} percentile (cols. 2, 7), the standardized earnigns using the control group mean and standard deviation (cols. 3, 8), the natural logarithm of earnings plus one (cols. 4, 9), and the inverse hyperbolic sine transformation of earnings (cols. 5, 10). Control variables are described in footnote 21. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure A5: Cumulative Distributions of Earnings by Treatment Group - Big Experiment



Notes: **Figure A5 shows that the earnings effects in the big experiment are not driven by outliers.** It shows the cumulative distribution function of earnings in the last seven days in the big experiment. Earnings are measured in 2021 USD PPP and winsorized at the 99^{th} percentile. Panel A shows earnings conditional on having worked in the last seven days. Panel B shows unconditional earnings.

E.2 Heterogenous Treatment Effects

This section collects various heterogeneity analyses. Table A25 shows that heterogeneous treatment effects on skill beliefs and search direction by baseline aligned comparative advantage belief in the big experiment are consistent with the results of tight experiment. Table A26 shows treatment effect heterogeneity by baseline comparative advantage beliefs and confidence simultaneously. Table A27 displays treatment effects on search direction in the tight experiment by whether skill requirements were revealed. Table A28 explores treatment effects on work quality outcomes (earnings and written contract) multiplied by potential mediators, such as tenure and start date. Finally, Table A29 shows that the heterogeneous treatment effects on labor market outcomes by aligned baseline comparative advantage beliefs in the big experiment are also consistent with the results of the tight experiment.

Table A25: Heterogeneous Treatment Effects on Skill Beliefs and Search Direction by Aligned Baseline Comparative Advantage Belief - Big Experiment

	Beliefs		Search direction
	Aligned CA belief (1)	Fraction aligned beliefs (2)	Aligned search (3)
Treatment	0.142***	0.134***	0.060***
	(0.011)	(0.009)	(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	0.388	0.165
Observations	4118	4131	4131

Notes: **Table A25 shows that treatment effects by baseline aligned comparative advantage beliefs on skill beliefs and search direction in the big experiment are broadly in line with the results of the tight experiment.** Columns indicate different outcome variables: a dummy of aligned comparative advantage beliefs (col. 1), the fraction of aligned belief domains (col. 2), a dummy indicating if search direction is aligned with comparative advantage in skills (self-reported) (col. 3). Control variables are described in footnote 21. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A26: Heterogeneous Treatment Effects on Main Outcomes by 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 A26 shows that the main results are robust to interacting the treatment with both baseline comparative advantage beliefs and baseline confidence in skills.** The magnitude of the interaction term in row 2 changes little relative to Tables 2 and 3, and the magnitude of the three-way interaction term (row 6) is small and insignificant. Baseline confidence levels are defined as the average baseline deviation of quintile skill beliefs from measured quintile (across numeracy and communication) standardized to have control group standard deviation of one. Positive values indicate overconfidence and negative values indicate underconfidence. Columns indicate different outcome variables: a dummy indicating aligned comparative advantage beliefs (col. 1), the aligned search index (col. 2), and the search effort index (col. 3). Control variables are described in footnote 15. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A27: 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 A27 shows that, for jobseekers with initially misaligned comparative advantage beliefs, treatment effects on search direction in the job choice tasks are stronger when skill requirements are revealed.** The outcome variable of the regressions is a dummy variable indicating if job choices in the job choice task are aligned with jobseekers' assessed comparative advantage. The effects are estimated at the job-choice individual level for jobseekers with initially misaligned beliefs (col. 1) and for jobseekers with initially aligned beliefs (col. 2). Controls include randomization block fixed effects, job pair and job pair order fixed effects, and pre-specified baseline covariates described in footnote 15. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A28: Treatment Effects on Combinations of Labor Market Outcome Measures - Big Experiment

	Earnings in the last seven days (w)					Written contract	
	Start before treatment	Start after treatment	Tenure	Wage emp.	Self emp.	Start before treatment	Start after treatment
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treatment	1.985*	4.607**	10.303	6.923***	-0.405	0.003	0.014
	(1.072)	(2.237)	(8.297)	(2.562)	(0.694)	(0.004)	(0.010)
Control mean	4.164	21.095	51.369	18.463	4.184	0.018	0.102
Observations	4176	4176	4173	4174	4174	4184	4184

Notes: **Table A28 shows that treatment effects on labor market outcomes are driven by wage jobs started after the treatment.** It displays treatment effects on selected work quality outcomes (earnings and written contract) multiplied by potential mediators (row 2 headers). Control variables are described in footnote 21. Winsorized variables (w) are winsorized at the 99^{th} percentile. All monetary figures are reported in 2021 USD PPP. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A29: Heterogeneous Treatment Effects on Labor Market Outcomes by Aligned Baseline Comparative Advantage Beliefs - Big Experiment

	Work quantity					Work quality			
	Index	# job offers (w)	Worked month 1	Worked month 2	Worked last 7 days	Index	Earnings (w)	Hourly wage (w)	Written contract
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment	0.065*	0.036	0.037**	0.009	0.013	0.097**	7.467**	0.355**	0.018
	(0.035)	(0.024)	(0.015)	(0.018)	(0.014)	(0.041)	(3.230)	(0.149)	(0.011)
Treatment × Aligned CA belief (bl)	-0.079	-0.023	-0.057	-0.015	-0.017	-0.054	-4.697	-0.304	-0.002
	(0.075)	(0.044)	(0.042)	(0.034)	(0.031)	(0.062)	(5.261)	(0.251)	(0.022)
Aligned CA belief (bl)	0.053	0.009	0.045	0.018	-0.014	-0.004	2.796	0.244	-0.023
	(0.055)	(0.032)	(0.028)	(0.024)	(0.019)	(0.044)	(3.577)	(0.178)	(0.015)
Treatment effect:	-0.015	0.013	-0.019	-0.006	-0.004	0.043	2.770	0.051	0.016
Aligned CA belief (bl)	(0.069)	(0.043)	(0.038)	(0.029)	(0.027)	(0.048)	(4.140)	(0.219)	(0.018)
Control mean	-0.000	0.195	0.465	0.437	0.309	0.000	25.424	1.267	0.120
Observations	4131	4071	4127	4130	4130	4132	4122	4111	4111

Notes: **Table A29 shows that treatment effects on labor market outcomes in the big experiment are driven by jobseekers with initially misaligned comparative advantage beliefs in line with the tight experiment.** Columns indicate different outcome variables: an Anderson (2008) index of the four employment quantity measures (col. 1), the number of job offers in the last 30 days (winsorized) (col. 2), a dummy indicating any work for pay 1 month after treatment (col. 3), a dummy indicating any work for pay 2 months after treatment (col. 4), a dummy indicating any work for pay in the last seven days (col. 5), an Anderson (2008) index of the three employment quality measures (col. 6), earnings in the last seven days (winsorized) (col. 7), hourly wages in the last seven days (winsorized) (col. 8), and a dummy indicating a written contract (col. 9). Control variables are described in footnote 21. All monetary figures are reported in 2021 USD PPP. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

F Beliefs About Wages and Job Offer Probabilities

The model predicts a four-stage reaction to treatment: beliefs about skill comparative advantage $\rightarrow$ expected return to skill-directed search $\rightarrow$ search direction $\rightarrow$ search outcomes. We show evidence consistent with the first, third, and fourth stages in the paper. We assess the second stage in this appendix, using measures of jobseekers' expectations about their labor market outcomes.

We find treatment effects on beliefs about search outcomes are broadly consistent with the model. But we view them as less important than the skill belief, search direction, and search outcome results we include in the main paper. The questions about expected search outcomes rely on complex forecasts by jobseekers, as discussed in recent reviews (Delavande, 2022; Mueller & Spinnewijn, 2022). For example, we ask jobseekers about their expected search duration. This requires jobseekers to forecast both the expected number of offers and attributes of those offers (wages, hours, travel costs, working conditions, etc.), as the attributes determine whether they will accept offers. In contrast, our tight experiment is optimized to measure search direction in multiple different ways, including direct measures of behavior rather than survey questions. And the survey questions about skill beliefs and current labor market outcomes are simpler to answer.

We only use data from the tight experiment in this section. In the tight experiment, we collect beliefs about search outcomes on the same day as treatment. This means that treatment effects on beliefs reflect only information acquired during the workshops. In the big experiment, the endline survey occurs months after treatment. This means that any treatment effects on beliefs about search outcomes would reflect both the direct effect of treatment and any indirect effects arising from treatment-induced changes in search actions and their outcomes. The indirect causal channel is interesting but not a key part of the argument we make in this paper.

Treatment effects on beliefs about returns to search direction: The model predicts that changes in search direction will be driven by changes in beliefs about the relative returns to searching for different types of jobs, i.e., beliefs about the returns to skill-directed job search. To evaluate this prediction, we collect two types of data.

First, we survey jobseekers about their expected outcomes from applying to each type of job and estimate treatment effects on these measures. We ask for the expected number of job offers in the next 30 days, time to employment, and wage when employed. We ask these questions after treatment and just after asking their planned number of applications in the next 30 days, and we ask the expected offers and time to employment questions conditional on their planned number of applications. We ask all questions sep-

Table A30: Treatment Effects on Beliefs About Returns to Skill-Directed Job Search - Tight Experiment

	Index		ΔE [offers per app] (w)		- (Δ [sear. dur.]) (w)		ΔE [wage] (w)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.095	0.260*	0.043**	0.053**	-0.007	-0.079	0.073	23.981**
	(0.100)	(0.130)	(0.018)	(0.026)	(0.157)	(0.185)	(6.605)	(9.383)
Treatment × Aligned CA belief (bl)		-0.341*		-0.023		0.114		-45.738***
		(0.180)		(0.034)		(0.223)		(13.960)
Aligned CA belief (bl)		0.689***		0.107***		0.432**		35.845***
		(0.140)		(0.030)		(0.197)		(12.069)
Treatment effect:		-0.081		0.030		0.035		-21.757**
Aligned CA belief (bl)		(0.124)		(0.022)		(0.191)		(9.164)
Control mean	-0.000	-0.000	0.027	0.027	0.244	0.244	6.007	6.007
Observations	278.000	278.000	273.000	273.000	272.000	272.000	278.000	278.000

Notes: **Table A30 shows that providing jobseekers with information about their comparative advantage in skills increases beliefs about the returns to directed search.** All outcomes are coded so that positive values mean beliefs about the returns to skill-directed search are higher. Columns indicate different outcome variables: an index of beliefs about the returns to directed search (cols. 1-2), the believed return of directed search in expected offers per application (cols. 3-4), the believed return of directed search on expected search duration (winsorized) (cols. 5-6), and the return of directed search on expected weekly wages measured in 2021 PPP USD (winsorized) (cols. 7-8). Returns are calculated as the arithmetic difference between beliefs about jobs that are aligned and are not aligned with the jobseeker's assessed comparative advantage. Even columns show heterogeneity by whether individuals have aligned comparative advantage beliefs at baseline. CA stands for comparative advantage in skills. Control variables are described in footnote 15. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

arately about all jobs, communication-heavy jobs, and numeracy-heavy jobs. We define the expected return to skill-directed job search as the expected outcome for jobs aligned to comparative advantage minus the expected outcome for jobs nonaligned to comparative advantage. For example, for a jobseeker with numeracy comparative advantage, the expected number of offers for numeracy jobs minus the expected number of offers for communication jobs. We also construct an inverse covariance-weighted average of the three measures.

Treatment effects on these survey measures of expected returns to skill-directed job search are mostly consistent with our model. Treated jobseekers expect 0.043 more job offers from skill-directed job search (165% of the control group mean with $p = 0.023$) and this effect, like most in the paper, is driven by jobseekers with baseline misaligned comparative advantage beliefs (Table A30, columns 3-4). Treated jobseekers with baseline misaligned comparative advantage beliefs expect weekly earnings 24 USD higher from skill-directed job search: roughly four times the control group mean with $p < x0.001$

(column 8). Treatment has negligible effects on the expected length of time until getting a job from skill-directed job search (columns 5-6). We return to this result on page 81.

Treatment effects on an index combining these survey measures of expected returns to skill-directed job search are consistent with our model, although slightly imprecisely estimated. Treatment increases the expected return to skill-directed search by 0.10 standard deviations ($p = 0.35$) for the average jobseeker (column 1). This result is driven by the same heterogeneity we see elsewhere in the paper: jobseekers with initially misaligned comparative advantage beliefs increase their expected returns by 0.26 standard deviations ($p = 0.054$) while jobseekers with initially aligned comparative advantage beliefs do not increase their expected returns (column 2).

Reassuringly, we see a 'sensible' relationship between baseline comparative advantage beliefs and expected returns to skill-directed job search. Jobseekers whose baseline comparative advantage belief matches their assessed comparative advantage belief have a 0.69 standard deviation higher expected return to skill-directed job search (column 2). The relationship is also positive for all three components of the index (columns 4, 6, 8).

We find similar but much less precisely estimated results using belief measures collected during the job choice task (Table A31). For 5 of the 11 job pairs, we ask jobseekers about the probability of getting an offer if they apply, the expected starting wage, and the general desirability of the job. We estimate treatment effects on these measures with a prespecified regression of the belief measure on treatment, a dummy for job alignment with jobseeker comparative advantage, their interaction, job fixed effects, and prespecified controls. Within each pair of jobs, treatment increases the expected offer probability and wage for the job aligned with the jobseeker's comparative advantage (Table A32). But the effects are small – roughly 2% of the control group mean and not statistically significant at conventional levels. So we do not view this as strong evidence supporting the model. These results might be less precise than results using the survey measures of expected returns to skill-directed search discussed above because the questions during the job choice task only ask about five specific pairs of jobs, rather than allowing jobseekers to implicitly average over many skill-directed job application choices.

Relationship between search direction and expected return to skill-directed job search: Beliefs about job attributes predict jobseekers' choices in the job choice task, consistent with a role for belief-based job search. For 5 of the 11 pairs of jobs, we asked jobseekers for the probability they would get an offer if they applied to each job and their expected salary if offered a job. We regress the job choice on the job offer probability times expected monthly wage using a logit regression model, following Wiswall & Zafar (2015). This is not an experimental analysis because we regress post-treatment choices

Table A31: Treatment Effects on Beliefs About Jobs in the Job Choice Task - Tight Experiment

	Desirability (sd)	Expected earnings (w)	Job offer probability
	(1)	(2)	(3)
Treatment	-0.038	-3.240	-0.022
	(0.032)	(5.921)	(0.022)
Treatment× Aligned skill req.	0.008	3.873	0.011
	(0.029)	(4.006)	(0.013)
Aligned skill req.	0.030	-3.297	0.017
	(0.022)	(2.724)	(0.011)
	-0.000	193.588	0.544
Control mean	2770	2770	2770

Notes: **Table A31 analyzes whether the treatment shifted beliefs about jobs differentially by whether skill requirements align with assessed comparative advantage.** Beliefs were elicited for 10 jobs in five job pairs for each jobseeker. Columns indicate different outcome variables: a standardized measure of desirability of jobs measured on a 0 to 10 Likert-scale (col. 1), expected weekly earnings winsorized at the 99^{th} percentile (col. 2), and perceived job offer probability (col. 3). Analysis is at the job-individual level and includes pre-specified control variables are described in footnote 15 and additionally include having clear comparative advantage in numeracy and communication, job choice order and job fixed effects. Standard errors clustered at the treatment-day level are in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

on post-treatment beliefs. Column 1 of Table A33 shows that a 100 USD increase in the expected weekly wage scaled by the offer probability is associated with a 6.7 percentage point increase in the probability of choosing that job ($p = 0.009$). This relationship is robust to adding both jobseeker and job pair fixed effects (columns 2-4). Treatment does not significantly change the slope of the belief-choice relationship, suggesting that at least part of the treatment effect on choices operates through beliefs.

Treatment effects on beliefs about search outcomes: The preceding analysis focuses on jobseekers' beliefs about the returns to skill-directed job search, which we define as the difference in jobseekers' expected outcomes from searching for jobs that do and do not align with their comparative advantage in skill assessments. We also construct measures of jobseekers' expected outcomes from searching for any type of job. We estimate treatment effects on these beliefs using the specification in equation (4).

Treatment has a positive but imprecisely estimated effect on expected wages, driven by jobseekers with baseline misaligned comparative advantage beliefs. Columns 3-4 of table A32 show that treatment increases expected weekly wage by 9.6 USD for the average treated jobseeker and 22.1 USD for the average treated jobseeker with misaligned baseline comparative advantage belief. Both effects are relatively imprecisely estimated (standard errors 7.4 and 14.9 respectively) perhaps reflecting the difficult of forecasting wage offers for respondents with limited work experience. Effects on jobseekers' reservation wages, beliefs about the minimum and maximum wages they might earn if employed, and an index combining all of these wage beliefs follow the same qualitative pattern (columns 1-2 and 5-10).[36] These positive effects on wage beliefs in the tight experiment are consistent with the positive effects on actual wages in the big experiment, although we cannot compare the magnitudes to evaluate forecast accuracy because the estimates come from two different experiments.

Treatment has a positive effect on the expected probability of formal employment, again driven by jobseekers with baseline misaligned comparative advantage beliefs. Columns 9-12 of Table A34 show that treatment increases the probability of employment in 1-3 months by 3-4 percentage points for the average treated jobseeker and 8-9 percentage points for the average treated jobseeker with misaligned baseline comparative advantage belief. Effects on other, less direct proxies for employment probability – callbacks and offers per application and search duration – are closer to zero (columns 3-8). These results might differ because the callback, offer, and search duration questions all explicitly condition on the jobseeker's planned number of applications, while the probability of employment questions are asked later in the survey and do not include this explicit conditioning.[37] The explicit conditioning might mean jobseekers put more mental weight on the role of search effort relative to search direction when answering these questions, but this is a speculative suggestion that future work could better evaluate. These results from the tight experiment are qualitatively consistent with the big experiment's positive effect on employment with a written contract. But the magnitudes are substantially different, perhaps in part because the control group's expectations are much higher than realized outcomes.

[36] In standard job search models, the reservation wage is both a decision rule and a feature of the wage distribution. This is not inconsistent with our interpretation of reservation wages as another proxy for wage expectations.

[37] For example, we ask "How many job applications do you plan to submit in the next 30 days?" and then "If you submit X job applications in the next 30 days, how many months starting from today do you think it will take you to find a formal job, with an employment contract where you are paid a regular salary?

Table A32: Treatment Effects on Beliefs About Wages - Tight Experiment

	Index		Wage expectations (w)		Minimum expected wage (w)		Maximum expected wage (w)		Reservation wage (w)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment	0.099	0.319*	9.566	22.119	8.934	19.180*	0.650	3.115	1.318	3.805
	(0.104)	(0.179)	(7.740)	(14.852)	(6.090)	(10.607)	(11.984)	(24.259)	(3.628)	(5.948)
Treatment × Aligned CA belief (bl)		-0.461*		-26.041		-21.065		-6.425		-5.366
		(0.258)		(21.034)		(17.531)		(38.832)		(6.578)
Aligned CA belief (bl)		0.247		10.018		4.604		20.878		5.287
		(0.208)		(17.725)		(15.279)		(29.176)		(4.039)
Control mean	-0.000	-0.000	212.045	212.045	131.392	131.392	324.076	324.076	109.567	109.567
Observations	278	278	278	278	278	278	278	278	277	277

Notes: **Table A32 shows that there are positive but insignificant average treatment effects on beliefs about wages in the tight experiment.** This effect is positive and significant for people with initially misaligned comparative advantage beliefs. CA stands for comparative advantage in skills. Columns indicate different outcome variables: an Anderson (2008) wage expectation index (cols. 1-2), the expected wage (cols. 3-4), the minimum expected wage (cols. 5-6), the maximum expected wage (cols. 7-8), and the reservation wage (cols. 9-10). Even columns show heterogeneity by whether jobseekers had aligned comparative advantage beliefs at baseline. Control variables are described in footnote 15. Winsorized variables (w) are winsorized at the 99^{th} percentile. All monetary figures are reported in 2021 USD PPP per week. Standard errors clustered at the treatment-day level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A33: Association Between Job Choices, Offer Probabilities, and Expected Wages - Tight Experiment

	Marginal effects on choice of numeracy job (logit estimate)							
	Control group				Control and treatment groups			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$E[Wage^{num}] - E[Wage^{com}]$ (w)	0.092***	0.097***	0.119***	0.127***	0.092**	0.099**	0.120***	0.124***
	(0.028)	(0.030)	(0.038)	(0.040)	(0.038)	(0.039)	(0.038)	(0.039)
Treatment $\times$ $E[Wage^{num}] - E[Wage^{com}]$ (w)					0.023	0.022	-0.076	-0.078
					(0.052)	(0.051)	(0.055)	(0.055)
Observations	695	695	600	600	1385	1385	1195	1195
Job pair fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
Individual fixed effects	No	No	Yes	Yes	No	No	Yes	Yes

Notes: **Table A33 shows that expected earnings are predictive of job choice and that treatment does not affect this relationship in the tight experiment.** It displays the marginal effects of the winsorized difference in expected returns to applying for a numeracy vs. a communication job using a logit estimator. Expected returns are defined as expected weekly wages (in 100s USD PPP in 2021) times the perceived likelihood of receiving a job offer. The winsorization is at the 1% and 99% level. The outcome is a dummy indicating choosing the numeracy job in a job pair in the job choice task. Cols. 1 to 4 are limited to the control group. Cols. 5 to 8 also interact the expected returns with a treatment indicator (cols. 5 and 6 also include a treatment dummy alone). Sample sizes drop from columns 1-2 and 5-6 to columns 3-4 and 7-8 because some jobseekers choose the numeracy-heavy job in all pairs, so their fixed effects are perfect predictors. Standard errors are robust (cols. 1-2), bootstrapped with 500 repetitions and stratification based on treatment status (cols. 3-4, cols. 7-8) and clustered at the treatment-day level (cols. 5-6). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A34: Treatment Effects on Beliefs About Offers and Search Duration - Tight Experiment

	Index		Callbacks / apps (w)		Offers / apps (w)		Month to job (w)		p(employed in 1 months)		p(employed in 3 months)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treatment	0.095	0.264	-0.001	0.039	-0.006	0.016	0.011	0.228	3.493	8.674**	4.235*	8.431**
	(0.122)	(0.194)	(0.023)	(0.036)	(0.021)	(0.031)	(0.247)	(0.392)	(2.238)	(3.748)	(2.315)	(4.042)
Treatment × Aligned CA belief (bl)		-0.360		-0.086		-0.049		-0.444		-10.821*		-8.678*
		(0.247)		(0.057)		(0.042)		(0.470)		(6.003)		(5.126)
Aligned CA belief (bl)		0.234		0.061		0.049		0.069		4.789		2.647
		(0.228)		(0.043)		(0.036)		(0.386)		(4.416)		(4.186)
Control mean	0.000	0.000	0.391	0.391	0.257	0.257	2.466	2.466	54.094	54.094	68.230	68.230
Observations	278	278	276	276	274	274	275	275	278	278	278	278

Notes: **Table A34 shows that treatment effects on beliefs about job offers and search duration are positive but mostly insignificant in the tight experiment.** Columns indicate different outcome variables: an Anderson (2008) index of labor market expectations (cols. 1-2), expected callbacks (cols. 3-4) and offers per application (cols. 5-6), expected time to find a full-time job (enters the index negatively) (cols. 7-8), perceived probability of being employed one month (cols. 9-10) and three months after baseline (cols. 11-12). Winsorized variables (w) are winsorized at the 99^{th} percentile. Even columns show heterogeneity by whether jobseekers had aligned CA beliefs at baseline. Control variables are described in footnote 15. Standard errors clustered at the treatment-day level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

G Benefit-Cost Comparison

This appendix reports the variable costs of the assessment operation and compares these to earnings gains experienced by treated jobseekers in the big experiment, showing that the gains likely exceed the benefits.

The average treatment effect on weekly earnings is 6.52 USD at the time of the follow-up survey, which occurs an average of 14.5 weeks after treatment. Using this to forecast the average lifetime earnings gain for treated jobseekers requires very strong modeling assumptions. Instead, we take a conservative approach and assume that the treatment effect is constant from treatment to endline and zero thereafter. This implies an average treatment effect of 94.50 USD on earnings.

We calculate that the average variable cost of the assessment operation is 45.74 USD using data from Harambee and J-PAL Africa's accounting records. This consists of 12.17 for rent and utilities for the assessment center; 10.14 for depreciation of the computers used in assessment; 0.35 on assessment and software licenses; 8.11 on airtime and data, mainly for internet access for the assessment computers and for contacting jobseekers; 3.82 on Harambee salaries for assessment support staff, psychologists delivering results briefings, and administrative support; 1.40 on J-PAL Africa salaries for research staff who helped to run the assessments and results briefings; and 9.74 on transport money for jobseekers to attend the assessments.

This implies that average benefit / average variable cost is **94.50 / 45.74 = 1.82**.

We view this benefit-cost calculation as suggestive rather than conclusive because, like all such calculations, it requires some simplifying assumptions. Most obviously, we use the conservative approach to estimating lifetime benefits described above, we omit average fixed costs because these are very dependent on the scale of the assessment service, and we do not consider general equilibrium effects or the benefits that might be accrued from alternative ways of spending this money. However, we do note that our measure of average variable costs is relatively broad, as it includes semi-fixed costs like facility and equipment rental, most staff costs, and assessment and software licenses. The only costs we exclude are those of actually creating the organization, senior management time, and general functions such as accounting.

Finally, we note that there is scope to run similar interventions far more cheaply through job search and matching platforms that incorporate assessments and personalized, automated feedback on assessment results. This approach would reduce or eliminate most of the large components of the average variable of in-person assessment: facility rental (27% of average variable cost), equipment rental (22%), and transport (21%).

H Search Effort Appendix

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, 63% of skill beliefs are above assessed skills and only 19% are below; these shares are 46% and 15% in the big experiment (see control means in Table A13). Thus, jobseekers receive on average negative news about their skill levels. If jobseekers react to this information by changing their search effort, this could also affect labor market outcomes.

Treatment effects on beliefs about skill levels: We first document that, on average, jobseekers update their skill beliefs negatively. Treated jobseekers in the tight experiment reduce their believed skill level by an average of 0.08 quintiles or 0.15 standard deviations over the two skills (Table A35, column 1, $p = 0.025$). Treated jobseekers in the big experiment reduce their believed skill level by an average of 0.11 terciles or 0.3 standard deviations over the three skills (Table A35, column 3, $p < 0.001$).

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

$$Y_{id} = T_d \cdot \beta_1 + T_d \cdot confidence_i \cdot \beta_2 + confidence_i \cdot \beta_3 + \mathbf{X}_{id} \cdot \Gamma + \varepsilon_i \tag{6}$$

Where $confidence_i$ is believed skill level minus assessed skill level.[38]

We find that $\hat{\beta}_2 < 0$, so treatment effects on belief levels are more negative for jobseekers with higher levels of baseline confidence (Table A35, row 2). One standard deviation higher baseline confidence is associated with a 0.26 quintile more negative treatment effect in the tight experiment (column 2, $p < 0.001$) and a 0.125 tercile more negative treatment in the big experiment (column 4, $p < 0.001$).

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 and numeracy jobs, E_C and E_N. This gives a utility function with three arguments: the expected outcome of search for communication jobs, $V_C(S_C, S_N, E_C)$, the expected outcome of search for numeracy 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 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,

[38] In the tight experiment, this uses quintiles and averages over two skills; in the big experiment this uses terciles and averages over three skills.

or an intertemporal budget constraint. The jobseeker's problem becomes

$$\max_{E_C, E_N} U\left(V_C\left(S_C, S_N, E_C\right), V_N\left(S_C, S_N, E_N\right), Y - P \cdot E_C - P \cdot E_N\right). \tag{7}$$

As in Section 2.1, we assume that utility is an increasing concave function of all three arguments, that expected search outcomes are increasing concave functions of skill and search effort, and that 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 product of search for communication 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 jobs and toward both search for numeracy jobs and alternative activities. Second, an income effect: it lowers the expected outcome from any given level of search effort, so the jobseeker has to increase search for both communication and numeracy jobs to maintain the same expected income. The net effect is a increase in search for numeracy jobs and an ambiguous effect on search for communication jobs, and hence an ambiguous effect on total search effort.[39]

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 10% of the control group mean – and none is statistically significant.[40] Treatment lowers an index of these search effort measures by 0.08 standard deviations (Table A36, column 1, $p = 0.47$) and a prespecified index of search effort on the SAYouth.mobi platform by 0.1 SDs (Table A37, column 1, $p = 0.29$). In the big experiment, treatment has a tiny effect of 0.003 SDs on an index of search effort measures (Table A38, column 1, $p = 0.92$). Treatment effects on the three components of this index – applications submitted, hours and money spent searching – are positive but tiny ($< 3\%$ of the control group mean) and none is statistically significant.

[39]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.

[40]Our six search effort measures are planned applications from surveys during the job search 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 after the workshop: days active, jobs viewed, and applications submitted. The planned applications, text messages, job applications are described in Section 3.4. The cover letter is a task-based measure of real search effort: the time jobseekers choose to spend drafting a cover letter for a real job application at the end of the workshop, on a laptop we provided, rather than collecting their incentive and leaving early.

Treatment effects also do not vary substantially by jobseekers' baseline confidence levels in either experiment. We estimate equation (6) and show heterogeneous treatment effects by baseline skill belief confidence in even-numbered columns in Tables A36, A37, and A38. 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 A36, 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 key search direction results are also robust to including the interaction with baseline confidence and baseline comparative advantage beliefs in the same regression (Table A26). The platform-based measure has a somewhat larger interaction effect of 0.14 standard deviations but it is still not statistically significant (Table A36, column 2, $p = 0.25$).

Similarly, treatment effects in the big experiment do not vary substantially by jobseekers' baseline confidence levels. For the search effort index, the interaction term is a tiny 0.03 standard deviations (Table A38, column 2, $p = 0.41$). The interaction effects on the index components are positive but small and not statistically significant.

In addition to adjusting search effort, jobseekers could also redirect their search effort to less lucrative jobs after receiving negative news about their skills. However, we find no evidence that jobseekers choose jobs with different salary levels in the job choice task in the tight experiment (Table A39). Treatment effects on the assessed salaries of chosen jobs are small (less than 1% of the control mean), far from significant, and do not significantly vary by baseline confidence about skills. This holds when using expert assessments of salaries (columns 1 and 2) and average beliefs of control group jobseekers (columns 3 and 4).

Finally, we show that treatment effects on labor market outcomes in the big experiment do not vary substantially by baseline confidence about skills (Table A40) The interaction effects are 0.001 on the work quantity index (column 1, $p = 0.97$) and -0.014 on the work quality index (column 6, $p = 0.70$). The effects on all index components are 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. See Abebe et al. (2022), Bandiera et al. (2023), and Banerjee & Sequeira (2023), and Mueller

Table A35: Heterogeneous Treatment Effects on Skill Beliefs by Baseline Confidence - Both Experiments

	Tight experiment		Big experiment	
	Average skill quintile beliefs (0-4)		Average skill tercile beliefs (0-2)	
	(1)	(2)	(3)	(4)
Treatment	-0.082**	0.157*	-0.108***	-0.007
	(0.035)	(0.083)	(0.010)	(0.011)
Treatment × Average confidence (bl)		-0.263***		-0.125***
		(0.060)		(0.012)
Average confidence (bl)		0.512***		-0.064***
		(0.071)		(0.016)
Control mean	2.693	2.693	1.446	1.446
Observations	278	278	4195	4131

Notes: **Table A35 shows that beliefs about relative skill level update downwards and differentially depending on jobseekers' baseline confidence about their relative skills.** Baseline confidence is defined in footnote 38. Columns indicate different outcome variables: average belief quintiles (cols. 1-2), and average belief tercile (cols. 3 and 4). Even numbered columns show treatment effects by baseline confidence. Controls include randomization block fixed effects, and prespecified baseline covariates described in footnotes 15 and 21. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

& Spinnewijn (2022) for mixed results about the sign of the relationship between search effort and beliefs about labor market prospects.

Table A36: Heterogeneous Treatment Effects on Search Effort - 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	-0.077	-0.101	-3.854	-3.574	-0.530	-0.377	0.003	-0.055	-1.446	-2.068
	(0.103)	(0.144)	(2.555)	(4.124)	(0.591)	(0.925)	(0.032)	(0.053)	(1.544)	(2.375)
Treatment × Avg. confidence (bl)		0.018		-0.429		-0.205		0.062		0.603
		(0.111)		(2.936)		(0.771)		(0.045)		(2.058)
Avg. confidence (bl)		0.197		4.032		1.280		-0.059		1.251
		(0.124)		(2.675)		(0.900)		(0.046)		(1.907)
Control mean	0.000	0.000	37.878	37.878	8.828	8.828	0.635	0.635	15.194	15.194
Observations	278	278	278	278	267	267	278	278	278	278

Notes: **Table A36 shows that treatment effects on search effort in the tight experiment do not vary significantly by baseline confidence levels.** Baseline confidence is defined in footnote 38. Columns indicate different outcome variables: an Anderson (2008) index of search effort measures (cols. 1-2), the number of planned applications in the next 30 days (cols. 3-4), the time individuals spent drafting a cover letter during the job-search workshop in minutes (cols. 5-6), the click rate for three SMS with links to job adverts sent to individuals (cols. 7-8), and the number of applications sent on the job-search platform in the 30 days following the treatment (cols. 9-10). Winsorized varaibles (w) are winsorized at the 99^{th} percentile. Control variables include randomization block fixed effects and controls described in footnote 15. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A37: Heterogeneous Treatment Effects on Platform Search Effort - Tight Experiment

	Platform search effort index		Days active		Adverts clicked (w)		Observed apps (w)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	-0.101	-0.233	-0.723	-1.451**	-0.183	-2.692	-1.446	-2.068
	(0.093)	(0.158)	(0.477)	(0.598)	(1.530)	(3.157)	(1.544)	(2.375)
Treatment × Average confidence (bl)		0.138		0.759		2.648		0.603
		(0.122)		(0.449)		(2.602)		(2.058)
Average confidence (bl)		-0.061		-0.278		-2.123		1.251
		(0.112)		(0.530)		(2.304)		(1.907)
Control mean	-0.000	-0.000	6.014	6.014	9.050	9.050	15.194	15.194
Observations	278	278	278	278	278	278	278	278

Notes: **Table A37 shows that treatment effects on prespecified search effort measures on the job-search platform SAYouth.mobi do not vary significantly by baseline confidence levels in the tight experiment.** Baseline confidence is defined in footnote 38. Columns indicate different outcomes: Anderson (2008) index of the three search effort measures (cols. 1-2), the number of days jobseekers were active on the platform in the 30 days following the treatment (cols. 3-4), the number of job adverts jobseekers clicked on on the platform in the 30 days following the treatment (cols. 5-6) and the number of applications sent on the job-search platform in the 30 days following the treatment (cols. 7-8). Winsorized variables (w) are winsorized at the 99^{th} percentile. Control variables are described in footnote 15. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A38: Heterogeneous Treatment Effects on Search Effort by Baseline Confidence - Big Experiment

	Index		Applications (w)		Hours spent searching (w)		Search expenditure (w)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.003	-0.017	0.065	0.021	-0.319	-0.600	0.227	-0.101
	(0.032)	(0.038)	(0.487)	(0.571)	(0.319)	(0.501)	(0.819)	(0.857)
Treatment × Average confidence (bl)		0.030		0.137		0.355		0.504
		(0.036)		(0.464)		(0.491)		(0.677)
Average confidence (bl)		-0.034		-0.069		-0.603		-0.428
		(0.049)		(0.677)		(0.657)		(0.807)
Control mean	-0.000	-0.000	11.716	11.716	11.083	11.083	20.878	20.878
Observations	4205	4131	4184	4111	4198	4124	4196	4122

Notes: **Table A38 shows that treatment effects on search effort do not vary significantly by baseline confidence levels in the big experiment .** Baseline confidence is defined in footnote 38. Columns show different outcomes: an Anderson (2008) index of search effort (cols. 1-2), the number of applications in the last 30 days (cols. 3-4), the number of hours spent searching for jobs in the last 30 days (cols. 5-6), and job search expenditure in the last 30 days (cols 7-8). Winsorized variables (w) are winsorized at the 99^{th} percentile. Control variables are described in footnote 21. All monetary figures are in 2021 USD PPP. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A39: Heterogeneous Treatment Effects on Vertical Search Direction - Tight Experiment

	Average monthly salary of job chosen in job choice task			
	Experts' assessment		Control group's assessment	
	(1)	(2)	(3)	(4)
Treatment	2.583	3.655	-0.931	-0.976
	(2.228)	(3.251)	(3.791)	(5.968)
Treatment × Average confidence (bl)		-1.008		-0.068
		(2.242)		(4.804)
Average confidence (bl)		-2.495		3.226
		(2.248)		(5.789)
Control mean	661.434	661.434	845.472	845.472
Observations	278	278	278	278

Notes: **Table A39 shows that jobseekers' negative updating about skill levels does not lead them to apply to jobs with lower (expected) wages.** The outcome in all columns is the average monthly salaries in 2021 USD PPP of the jobs chosen in the job choice task in the tight experiment. The salaries are assessed by the hiring experts (cols. 1-2) or by jobseekers in the control group (cols. 3-4). Baseline confidence is defined in footnote 38. Control variables are described in footnote 15. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A40: Heterogeneous Treatment Effects on Labor Market Outcomes by Confidence - Big Experiment

	Work quantity					Work quality			
	Index	Job offers (w)	Worked month 1	Worked month 2	Worked last 7 days	Index	Earnings (w)	Hourly wage (w)	Written contract
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment	0.048	0.025	0.009	0.015	0.010	0.097**	6.461*	0.290*	0.023*
	(0.036)	(0.022)	(0.019)	(0.019)	(0.017)	(0.045)	(3.456)	(0.172)	(0.013)
Treatment × Average confidence	0.001	-0.009	0.022	-0.011	-0.000	-0.014	0.100	0.008	-0.007
	(0.025)	(0.018)	(0.016)	(0.016)	(0.014)	(0.036)	(2.755)	(0.132)	(0.011)
Average confidence	-0.004	0.019	-0.006	-0.005	-0.019	-0.051	-4.044	-0.016	-0.024
	(0.038)	(0.023)	(0.018)	(0.020)	(0.017)	(0.046)	(3.640)	(0.177)	(0.015)
Control mean	-0.000	0.182	0.465	0.437	0.309	0.000	25.424	1.267	0.120
Observations	4131	4071	4127	4130	4130	4132	4122	4111	4111

Notes: **Table A40 shows that treatment effects on labor market outcomes in the big experiment do not vary by baseline confidence levels.** Baseline confidence is defined in footnote 38. Columns indicate different outcome variables: an Anderson (2008) index of the employment quantity measures (col. 1), the number of job offers in the last 30 days (col. 2), a dummy indicating any work for pay in month 1 after treatment (col. 3), a dummy indicating any work for pay in month 2 after treatment (col. 4), a dummy indicating any work for pay in the last seven days (col. 5), an Anderson (2008) index of the employment quality measures (col. 6), earnings in the last seven days (col. 7), hourly wages in the last seven days (col. 8), and a dummy indicating a written contract (col. 9). Winsorized variables (w) are winsorized at the 99^{th} percentile. Control variables are described in footnote 21. All monetary figures are reported in 2021 USD PPP. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

I Additional Mechanism Tables

This appendix contains tables with an analysis of potential mechanisms other than comparative advantage beliefs or beliefs about skill levels discussed in Section 6. Table A41 shows evidence against two potential mechanisms in the big experiment: self-esteem and education investments from the big experiment. Table A42 shows treatment effects on the three different willingness-to-pay (WTP) measures in the tight experiment. We find negative treatment effects on jobseekers' WTP for the numeracy workbook and no average treatment effects on the valuation for the communication workbook. Overall these results indicate that educational investment is not a likely mechanism in our setting. Further details of the WTP protocol are available at: https://bit.ly/3E34oYH

Table A43 shows that average treatment effects on labor market outcomes are driven by jobseekers who attached their skill report with applications. In Table A42 we do not find treatment effect on jobseekers' willingness-to-pay for the skill demand report. This suggests that jobseekers are unlikely to seek information about jobs' skill demand based on their treatment status, and hence this behavior is not responsible for the labor market effects. Table A44 shows that the treatment did not induce congestion effects.

Table A41: Treatment Effects on Additional Mechanisms - Big Experiment

	Self-esteem				Education investment		
	SMS (z) (1)	SMS above med. (2)	Endline (z) (3)	Endline above med. (4)	Any (5)	Apprenticeship (6)	Formal (7)
Treatment	0.003	0.014	0.007	0.014	0.011	0.006	0.009
	(0.008)	(0.014)	(0.021)	(0.015)	(0.011)	(0.005)	(0.011)
Control mean	-0.000	0.483	0.000	0.471	0.224	0.036	0.185
Observations	3334	3334	4206	4206	4205	4205	4205

Notes: **Table A41 shows that there are no average treatment effects on self-esteem or education investments in the big experiment.** Columns 1 and 2 show effects on self-esteem in the SMS survey 2-3 days after treatment on one item from the Rosenberg (1965) scale, "Sometimes I think I am no good". Jobseekers texted back one number "do you 4) strongly agree 5) agree a little 6) neutral 7) disagree a little 8) strongly disagree?". Columns 3 and 4 show effects on self-esteem measured in the endline survey 2-4 months after treatment on a five-item scale adapted from the Rosenberg (1965) scale, administered by fieldworkers as part of the phone survey. Each item uses a Likert scale and measures the extent to which jobseekers agree with a statement indicating high self-esteem. Columns 1 and 3 show effects on standardized measures and columns 2 and 4 show results for above median dummies. Column 5 shows the effect on a dummy indicating any newly started education. Column 6 shows the effect on a dummy indicating any newly started apprenticeship. Column 7 shows the effect on a dummy indicating any other newly started formal post-secondary education. Control variables are described in footnote 21. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A42: Treatment Effects on Willingness-to-pay - Tight Experiment

	Info on skill requirements						Numeracy materials						Communication materials					
	Pooled		Num. CA		Comm. CA		Pooled		Num. CA		Comm. CA		Pooled		Num. CA		Comm. CA	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Treatment	-0.719	-6.851	-12.823	-14.035	1.708	-8.179	-13.410***	-15.309**	-14.527*	-20.583**	-10.046**	-3.217	-0.822	-0.247	5.252	-2.571	-3.821	3.862
	(5.766)	(11.255)	(13.609)	(15.888)	(8.396)	(17.726)	(3.539)	(6.775)	(8.049)	(8.944)	(3.898)	(11.201)	(3.790)	(6.747)	(9.785)	(12.162)	(5.882)	(11.719)
Treatment × Aligned CA belief (bl)		12.287		8.804		14.077		5.019		36.732**		-9.825		-0.847		45.821**		-11.987
		(18.773)		(38.367)		(22.913)		(11.915)		(15.806)		(16.400)		(10.474)		(21.394)		(14.833)
Aligned CA belief (bl)		0.174		-12.737		7.989		-14.744		-38.444***		-4.144		-3.753		-43.991***		7.831
		(16.202)		(28.007)		(16.687)		(8.948)		(11.399)		(11.074)		(9.163)		(11.637)		(10.853)
Treatment effect:		5.436		-5.231		5.898		-10.290		16.149		-13.042*		-1.094		43.251**		-8.125
Aligned CA belief (bl)		(10.847)		(33.252)		(11.314)		(7.268)		(14.781)		(7.247)		(6.130)		(16.512)		(7.682)
Control mean	78.867	78.867	79.337	79.337	78.611	78.611	54.964	54.964	55.408	55.408	54.722	54.722	48.327	48.327	49.184	49.184	47.861	47.861
Observations	277	277	105	105	172	172	277	277	105	105	172	172	277	277	105	105	172	172

Notes: **Table A42 analyzes how treated jobseekers in the tight experiment change their willingness-to-pay for different goods.** Columns 1 to 6 show effects on the willingness-to-pay for a document with the expert-assessed skill requirements for the 11 job pairs in the job choice task. Columns 7 to 12 show impacts on willingness to pay for numeracy course materials. Columns 13 to 18 show impacts on communication course materials. Columns 1, 2, 7, 8, 13, and 14 show results for the full sample. Columns 3, 4, 9, 10, 15, and 16 show results for individuals with a comparative advantage in numeracy. Columns 5, 6, 11, 12, 17, and 18 show results for individuals with a comparative advantage in communication. Control variables are described in footnote 15. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A43: Heterogeneous Treatment Effects on Labor Market Outcomes by Skill Report Attachment to Applications - Big Experiment

	Work quantity					Work quality			
	Index	Job offers (w)	Worked month 1	Worked month 2	Worked last 7 days	Index	Earnings (w)	Hourly wage (w)	Written contract
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treatment	0.068*	0.040	0.033**	0.018	0.014	0.110***	8.572***	0.359**	0.023**
	(0.039)	(0.027)	(0.016)	(0.018)	(0.016)	(0.039)	(3.020)	(0.143)	(0.011)
Treatment × Attached report w. application	-0.066	-0.038	-0.021	-0.039	-0.013	-0.071	-5.108	-0.179	-0.019
	(0.049)	(0.038)	(0.022)	(0.024)	(0.024)	(0.058)	(4.229)	(0.211)	(0.017)
Control mean	-0.000	0.195	0.465	0.437	0.309	0.000	25.424	1.267	0.120
Observations	3992	3933	3988	3991	3991	3993	3983	3971	3971

Notes: **Table A43 shows that treatment effects on labor market outcomes are not driven by jobseekers that report attaching their skill reports to at least one application in the big experiment.** The main effect of *Attached report w. application* cannot be estimated because no control individual received a report. Columns indicate different outcome variables: an Anderson (2008) index of the four employment quantity measures (col. 1), the number of job offers in the last 30 days (col. 2), a dummy indicating any work for pay in month 1 after treatment (col. 3), a dummy indicating any work for pay in month 2 after treatment (col. 4), a dummy indicating any work for pay in the last seven days (col. 5), an Anderson (2008) index of the three employment quality measures (col. 6), earnings in the last seven days (col. 7), hourly wages in the last seven days (col. 8), and a dummy indicating a written contract (col. 9). Winsorized variables (w) are winsorized at the 99^{th} percentile. Control variables are described in footnote 21. All monetary figures are reported in 2021 USD PPP. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A44: Treatment Effects on Concentration of Job Search - Tight Experiment

	Degree of job search concentration		
	Individual level measure		Job-pair level measure
	(1)	(2)	(3)
Treatment	-0.007	0.009	-0.029**
	(0.017)	(0.019)	(0.013)
Treatment × Aligned CA belief (bl)		-0.032	
		(0.028)	
Aligned CA belief (bl)		0.050**	
		(0.024)	
Treatment effect:		-0.023	
Aligned CA belief (bl)		(0.024)	
Control mean	0.181	0.181	0.077
Observations	278	278	22

Notes: **Table A44 shows that the treatment (weakly) decreases the concentration of jobseekers' applications in the job choice task.** Columns 1 and 2 show effects on the concentration of job choices in the job choice task at the jobseeker-level. We measure concentration of job choices as the absolute deviation of the fraction of chosen numeracy jobs from 0.5 averaged across job pairs at the jobseeker level. Column 3 shows effects on the concentration of job choices in the job choice task at the job-pair level. We construct our measure concentration of job choices at the job-pair level in two steps. First, we calculate the fraction of jobseekers choosing the numeracy job in each job-pair – treatment group combination. Second, we calculate the absolute deviation of this measure from 0.5. Higher numbers indicate a higher degree of concentration of job choices for both measures. We do not consider heterogeneity for the job-pair level analysis because the outcome of interest is at the labor market level averaged across jobseekers. Controls used in columns 1 and 2 are described in footnote 15. Standard errors are clustered at the treatment-day level in columns 1 and 2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

J Gender

In this appendix we present findings related to gender. Table A45 displays unadjusted and adjusted gender differences in baseline skill beliefs in the tight experiment and Table A46 shows the same for the big experiment. We find that the gender differences in beliefs if anything are small. Table A47 and Table A48 displays treatment effects on skill beliefs by gender in the tight and big experiment respectively. The treatment effects on skill beliefs do not differ by gender.

Table A45: Gender Differences in Skill Beliefs - Tight Experiment

	Female	Male	Δ	p($\Delta = 0$)	$\Delta(adjusted)$	p($\Delta(adjusted) = 0$)	N
Aligned CA belief	0.50	0.46	-0.04	0.53	-0.01	0.83	278
Fraction aligned beliefs	0.23	0.21	-0.02	0.46	-0.05	0.15	278
Fraction overcofident beliefs	0.62	0.59	-0.03	0.58	0.04	0.11	278
Fraction underercofident beliefs	0.15	0.20	0.05	0.12	0.01	0.79	278

Notes: **Table A45 shows that gender differences in beliefs are small in the tight experiment.** Adjusted gender differences control for age, education level, skill quintiles, and randomization block fixed effects.P-values are calculated using robust standard errors.

Table A46: Gender Differences in Skill Beliefs - Big Experiment

	Female	Male	Δ	p($\Delta = 0$)	$\Delta(adjusted)$	p($\Delta(adjusted) = 0$)	N
Aligned CA belief	0.19	0.23	0.03	0.01	0.03	0.02	4312
Fraction aligned beliefs	0.35	0.42	0.07	0.00	0.01	0.19	4378
Fraction overconfident beliefs	0.53	0.45	-0.08	0.00	0.00	0.73	4378
Fraction underconfident beliefs	0.11	0.13	0.01	0.06	-0.01	0.03	4378

Notes: **Table A46 shows that gender differences in baseline beliefs in the big experiment are small**. Adjusted differences control for pre-specified covariates described in footnote 21 (except the baseline value of the outcome variable). CA stands for comparative advantage in skills. P-values are calculated using robust standard errors.

Table A47: Heterogeneous Treatment Effects on Skill Beliefs by Gender - Tight Experiment

	Aligned CA belief		Fraction aligned beliefs	
	(1)	(2)	(3)	(4)
Treatment	0.181*	0.178**	0.087	0.063
	(0.091)	(0.079)	(0.079)	(0.051)
Treatment × Female	-0.074	-0.062	0.059	0.022
	(0.107)	(0.103)	(0.095)	(0.062)
Female	0.082	0.019	-0.038	-0.033
	(0.066)	(0.063)	(0.060)	(0.042)
Treatment effect:	0.107	0.116**	0.146***	0.085**
Female	(0.065)	(0.047)	(0.049)	(0.032)
Control mean	0.475	0.475	0.183	0.183
Observations	278.000	278.000	278.000	278.000
Controls	No	Yes	No	Yes

Notes: **Table A47 shows that treatment effects on skill beliefs do not differ by gender in the tight experiment.** Columns 1 and 2 show effects on a dummy indicating beliefs about respondents' comparative advantage in skills that are aligned with the assessment results. Columns 3 and 4 show treatment effects on the fraction of skill beliefs that align with measured skill quintiles. Control variables are described in footnote 15. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A48: Heterogeneous Treatment Effects on Beliefs about Skills by Gender - Big Experiment

	Aligned CA belief		Fraction aligned beliefs	
	(1)	(2)	(3)	(4)
Treatment	0.156***	0.155***	0.150***	0.153***
	(0.021)	(0.020)	(0.019)	(0.015)
Treatment × Female	-0.032	-0.026	-0.021	-0.016
	(0.027)	(0.026)	(0.020)	(0.017)
Female	-0.029	-0.011	-0.051***	0.003
	(0.018)	(0.016)	(0.015)	(0.012)
Treatment effect:	0.124***	0.130***	0.128***	0.136***
Female	(0.013)	(0.015)	(0.013)	(0.010)
Control mean	0.196	0.196	0.388	0.388
Observations	4191	4118	4205	4195
Controls	No	Yes	No	Yes

Notes: **Table A48 shows that treatment effects on skill beliefs do not differ by gender in the big experiment.** Columns 1 and 2 show impacts on a dummy indicating comparative advantage beliefs that align with assessment results. Columns 3 and 4 show impacts on the fraction of skill beliefs that align with measured skill terciles across three skill domains. Control variables are described in footnote 21. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

K Preregistration Appendix

The paper relies on two pre-registrations and analysis plans (AEARCTR-0001631 for the big experiment and AEARCTR-0010000 for the tight experiment).

The pre-analysis plan for the big experiment includes the experimental design, the treatment effect analysis, and the main labor market outcomes but it does not include outcomes on comparative advantage beliefs and aligned search or the heterogeneity analysis by baseline aligned comparative advantage beliefs.

The tight experiment was set up to test the research question that arose from the exploratory analysis of the big experiment and thus the results from this experiment are confirmatory. We deviate from the pre-analysis plan in the following ways:

- We restrict the sample to people who have a clear comparative advantage. When we use this restricted sample, we are not including a prespecified control variable, the dummy variable about whether the jobseeker has a clear comparative advantage, because this control variable have no variation in the restricted sample. In Appendix Tables A20 and A21 we show the main results of the paper for the full sample adding the dummy control variable as a robustness check. The interpretation of results remains unchanged.

- We added further outcomes on search alignment in the tight experiment. These were SMS and platform outcomes (and their index). We were able to get access to these measures from our partner after the time of the pre-registration. We correct for the inclusion of these additional measures by creating a summary search alignment index.

- To align the search alignment and search effort measures, we add the SMS click rate and the number of observed application clicks to the main search effort table and, again, construct a summary measure. We show results for the pre-specified platform search index in Table A37.

- Following recent methodological critiques (Chen & Roth, 2022; Mullahy & Norton, 2022) we use winsorization instead of the pre-specified inverse hyperbolic sine transformation to handle outliers. The key results are all robust to using the inverse hyperbolic sine instead of winsorization.