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

Andrea Kiss ⓡ Robert Garlick ⓡ Kate Orkin ⓡ Lukas Hensel

July 24, 2023

Abstract

Worker sorting into tasks and occupations based on their skills plays a potentially important role in aggregate labor productivity. This sorting may be inefficient if jobseekers have inaccurate beliefs about their skills and do not apply to jobs that match their skills. In South Africa, we assess young jobseekers' numeracy and communication skills and find jobseekers' self-perceived and measured comparative advantage persistently differ. In a field experiment, jobseekers given their skill assessment results have more accurate beliefs about their comparative advantage and redirect search toward jobs whose skill demand matches their comparative advantage in multiple search measures: an incentivized job choice task where workers choose between jobs with different skill requirements, usage data from an online job search platform, and self-reported search plans. In a larger field experiment, treatment has similar effects on self-reported measures of beliefs and search direction and substantially raises earnings and job quality, although not employment rates, after three months. These patterns are consistent with models of endogenously directed job search, where jobseekers' beliefs about their skills influence where they direct job applications and hence the wages they are offered.

*Robert Garlick, Duke University, robert.garlick@duke.edu; Lukas Hensel, Guanghua School of Management, Peking University, lukas.hensel@gsm.pku.edu.cn; Andrea Kiss, Carnegie Mellon University, akiss@andrew.cmu.edu; Kate Orkin, Blavatnik School of Government, University of Oxford, kate.orkin@bsg.ox.ac.uk. Order of authors is randomized and all authors contributed equally. For helpful comments, we thank Johannes Abeler, Peter Arcidiacono, Michele Belot, Clement Imbert, David Huffman, Jeremy Magruder, 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, 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. 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, and Upjohn Institute for Employment Research.

1 Introduction

Worker sorting into tasks and occupations has long been recognized as an important feature of the labor market (Roy, 1951). Efficiently matching workers with the tasks where their skills are most productive offers the prospect of substantial gains in output (Lise & Postel-Vinay, 2020). The efficiency of sorting depends crucially on the job search process, where jobseekers choose how to direct their search effort to different types of jobs. However, jobseekers may have imperfect information about their productivity in different types of jobs, causing them to direct search effort to jobs that poorly match their skills.

This paper studies how jobseekers' beliefs about their own skills influence their job search behavior and outcomes. We emphasize how beliefs about relative skill ranking in different dimensions, which we call **comparative advantage beliefs**, influence search across jobs with different skill demands, which we call **skill-directed job search**. We run two experiments with complementary designs. In both, we assess jobseekers' skills in multiple domains, measure their beliefs about these skills, experimentally shift these beliefs by providing assessment results, and measure their post-treatment search activities and outcomes. We find that giving jobseekers their assessment results moves their comparative advantage beliefs toward their measured comparative advantage, redirects search to jobs that match their comparative advantage in skills, and leads to substantial increases in their earnings and wages but not their employment rate. These results suggest jobseekers face information frictions that limit their ability to match to jobs where they will be most productive. Existing studies cannot answer this question because they lack the necessary combination of data: jobseeker-level data on multidimensional skills, beliefs about these skills, and job search activities and outcomes; application-level data about skill demands; and exogenous variation in beliefs about skills.[1]

We begin by describing skills, beliefs, and job search of young jobseekers from disadvantaged backgrounds in Johannesburg, South Africa. We assess their skills in communication and numeracy – two general-purpose, job-relevant skills that we verify firms in this context value in prospective workers. We then measure jobseekers' beliefs about their own skills and document substantial differences between measured and perceived skills, both expressed as ordinal rankings. Measured and perceived comparative advantage differ for almost half the jobseekers.[2] These differences persist through time, even for those

[1]Survey data typically record aggregated search data such as total applications submitted, not application-level data including skill demand. Job search platform data seldom measure skills, beliefs, or labor market outcomes. Administrative data from unemployment insurance systems do not measure beliefs and seldom measure skills or search.

[2]We define each jobseeker's comparative advantage as the skill dimension in which she ranks highest , following some of the few existing job search studies using multidimensional skill measures (e.g. Guvenen

who find work or search intensively, perhaps because jobseekers receive limited feedback on their skills from job applications. Jobseekers' skill beliefs also predict their job search choices and jobseekers believe their search outcomes would improve if their skills were higher.

We propose a simple model to capture these patterns and show how these skill beliefs can influence job search. In the model, jobseekers choose the levels of search effort directed to jobs with different skill requirements. Inaccurate skill beliefs distort the perceived return to searching for each type of job and hence distort search decisions. If jobseekers apply to jobs that poorly match their skills, this can reduce the number of job offers received or wages offered.

We then run two field experiments to identify the relationship between comparative advantage beliefs, skill-directed job search, and search outcomes. Our first experiment (N = 278) focuses on skill-directed job search and is embedded in day-long job search workshops run with a job search assistance agency that include taking the skill assessments. We randomly assign half the workshops to treatment: giving participants their skill assessment results. Treated participants' skill beliefs are 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 different measures: applications in a novel, incentivized task in which participants choose between applying to jobs with different skill requirements; planned applications in an end-of-workshop survey; and actual applications on a job search platform in the month after the workshop. These results show the link between comparative advantage beliefs and skill-aligned job search using unusually rich data but do not show if this affects search outcomes.

Our second experiment (N = 4,389) focuses on labor market outcomes. This experiment is also embedded in day-long job search workshops that include skill assessments. We again randomly assign half of the workshops to give participants their skill assessment results. Survey data collected on average 3.5 months after the workshops show that treatment improves labor market outcomes: treated jobseekers' weekly earnings are 6.5 USD higher and hourly wages are 0.3 USD higher than control group jobseekers (23% and 26% of control group mean, respectively) and they are more likely to have formal job contracts, although we find at most weak evidence for a higher employment rate. The average earnings gain between treatment and the follow-up survey is roughly 1.8

et al. 2020). This definition has both advantages and disadvantages relative to some other approaches in labor economics, as we discuss later in the paper. In particular, it avoids the need for any information about labor demand and avoids the challenging task of estimating the relative prices of different skills in the labor market.

times the average variable cost of running the assessment system, suggesting a potentially cost-effective role for these types of interventions in active labor market policies. Treatment effects on jobseekers' comparative advantage beliefs and skill-directed job search are qualitatively similar to the first experiment, although measurement differences make quantitative comparisons difficult.

The two field experiments and the descriptive data show a consistent picture of this labor market: shifting jobseekers' perceived comparative advantage toward their measured comparative advantage redirects their search toward jobs aligned with their measured comparative advantage and leads to better-quality employment. They provide two complementary types of evidence. The first experiment provides unusually rich survey, task, and administrative data on beliefs and search, including application-level data. The second experiment provides survey data on beliefs, search, and labor market outcomes for a larger sample over a longer time period. Given the two experiments' different strengths, we call them the "tight" and "big" experiments respectively throughout the paper.[3] The similar belief and search results across two experiments show some generalizability for the economic process we study because the experiments are conducted six years apart in the same location with similar but not identical implementation and similar but not identical samples.

We find limited evidence for several other theory-informed ways that jobseekers might respond to new information about their skills. On average, jobseekers in our sample believe at baseline that their skills are higher than their measured skills, so treated jobseekers receive on average negative news about their skill levels. The effect of lower perceived skill level on search effort and hence labor market outcomes is theoretically ambiguous in our model. In practice, we find negligible effects on search effort for jobseekers on average and for those who receive positive or negative news, suggesting that changes in overall search effort are unlikely to explain the observed effects on labor market outcomes. Furthermore, we show treatment does not affect jobseekers' generalized self-esteem, just their beliefs about their skills; does not affect their educational investment, at least over the time horizon we observe; and does not provide information directly to firms about jobseekers' skills.

More generally, what labor market conditions might lead to inaccurate comparative advantage beliefs and hence to poorly skill-directed job search? Jobseekers might en-

[3]We recognize this is non-standard terminology and hope readers will take this in the lighthearted spirit we intend. Standard terminology is not well-suited to distinguish the two experiments. 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.

ter the labor market with inaccurate comparative advantage beliefs in the many countries where education systems give students inaccurate feedback on their performance (Pritchett, 2013). High unemployment might limit jobseekers' opportunities to learn about their skills through work experience, a standard learning mechanism in many models (e.g. Guvenen et al. 2020). And 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. This form of limited information might be particularly costly when screening many mismatched job applicants is costly to firms and when job search is expensive for jobseekers (Abebe et al., 2021a; Banerjee & Sequeira, 2020), particularly if well-matched jobseekers face systematically higher application costs (Abebe et al., 2021b).

This paper makes two primary contributions. First, we provide the first direct evidence that jobseekers' beliefs about their comparative advantage in skills influence how they direct search to different job types and hence influence their search outcomes. Sorting on skill comparative advantage is central to both classic and contemporary labor market models but is less studied empirically (Roy, 1951; Sanders & Taber, 2012). There is a growing literature studying the relationship between jobseekers' beliefs, search activities, and search outcomes, reviewed by Mueller & Spinnewijn (2022). This includes work on the co-evolution of beliefs and search in panel data (Conlon et al., 2018; Mueller et al., 2021; Spinnewijn, 2015); experiments that provide information about labor market conditions (Altmann et al., 2018; Jones & Santos, 2022); and work on search subsidies, matching services, or mentoring programs that influence jobseekers' beliefs about their labor market prospects (Abebe et al., 2022; Alfonsi et al., 2022; Bandiera et al., 2021; Banerjee & Sequeira, 2020; Kelley et al., 2020).[4] This work focuses on jobseekers' beliefs about the level of their labor market prospects, captured by job offer arrival rates or wage offer distributions. But this does not study beliefs about comparative advantage, perhaps because very few datasets combine information on jobseekers' skills and vacancies' skill demands. We complement this literature by collecting unique data to study beliefs about comparative advantage and how search is directed toward jobs with different skill demands. Comparative advantage beliefs are likely to influence the beliefs about labor market prospects that are more widely studied. Methodologically, our work is closest to Altmann et al. (2022) and Belot et al. (2019, 2022a), who show that encouraging jobseekers to apply to different occupations can improve some labor market outcomes, consistent with jobseekers having imperfect information about their job offer probabilities in some occupations.

[4]This relates to research into jobseekers' beliefs about attributes of specific jobs (e.g. Bazzi et al. 2021; Heath et al. 2022; Sockin & Sojourner 2020; Subramanian 2022).

Like these papers, we study how jobseekers direct search to different job types using jobseeker×application-level data, rather than relying only on more common jobseeker-level measures of search such as the total number of job applications they submit. Unlike these papers, we study the role of comparative advantage and measure jobseekers' skills, beliefs, and a broader range of search behavior and outcomes.

Second, our work complements a growing literature on the role of job search in worker-firm matching with multidimensional skills (e.g. Lise & Postel-Vinay 2020). Recent work uses dynamic models to estimate large lifetime earnings losses when jobseekers enter the labor market with imperfect information about their skills and hence their productivity in different jobs (e.g. Guvenen et al. 2020). These papers use rich longitudinal data on employment, guided by dynamic models, to show long-term consequences of the belief-induced search frictions that we document. But they do not observe beliefs or search. Our measures of skills, beliefs about skills, search, and short-term labor market outcomes, combined with experimental variation in beliefs, allow us to directly assess the assumptions of these models. Within this literature, our findings are most consistent with the modeling framework of Baley et al. (2022), who study the long-term consequences of jobseekers directing search based on inaccurate beliefs about their multidimensional skills.

A parallel literature focuses on firm-side limited information about job applicants' skills.[5] We differ by focusing on jobseekers' assessments of their own skills. This also matters for firms, because the self-assessments determine how jobseekers direct search, which in turn determines the pool of applicants to each job. Carranza et al. (2022) study the relative roles of firm- and jobseeker-side learning about jobseekers' skills using some data from our "big" field experiment. This paper focuses more deeply on jobseeker-side learning and differs from Carranza et al. (2022) in several ways. This paper adds the "tight"framed field experiment, uses more detailed measures of jobseekers' beliefs and search activity from the shared field experiment, introduces the idea of comparative advantage beliefs, and substantially expands the idea of skill-directed job search, both empirically and theoretically. In contrast, Carranza et al. (2022) use additional firm-facing experiments to show how firms respond to new information about jobseekers' skills, which this paper does not study.

[5] For example, providing firms with additional information about prospective workers' skills can increase employment and earnings (Abebe et al., 2021a; Abel et al., 2020; Bassi & Nansamba, 2022; Pallais, 2014), some wage-tenure relationships are consistent with firms learning about current workers' productivity over time (Altonji & Pierret, 2001; Arcidiacono et al., 2010; Kahn & Lange, 2014), and workers do better on some labor market outcomes when they have formal educational qualifications, conditional on measured skills (Alfonsi et al., 2017; MacLeod et al., 2017).

Our work also relates to research into the formation and consequences of beliefs about skills in contexts other than job search. Some papers study how skill beliefs influence decisions in the workplace, typically focusing on the role of overconfidence rather than comparative advantage beliefs (e.g. Hoffman & Burks 2020; Huffman et al. 2022; Malmendier & Tate 2015). A growing literature in education shows that potentially noisy self-assessments of skill levels and comparative advantage matter for enrollment and subject choices, even though education might offer scope for more structured feedback on skills than job search (e.g. Arcidiacono et al. 2016; Stinebrickner & Stinebrickner 2014a,b). Finally, a large literature studies the impact of beliefs on decisions in lab settings, including a few papers studying the role of beliefs about skills in lab tasks that mimic job search (e.g. Falk et al. 2006).

Methodologically, we add to a small, recent literature showing the value of combining multiple types of experiments to understand both a job search process and the outcomes of that process (Carranza et al., 2022; Cortés et al., 2021; Field et al., 2023). The job search task in our tight experiment provides a novel way to use jobseekers' choices over job applications to measure their valuation of skill match relative to other vacancy characteristics, building on work using choice experiments to study job or education choices (Mas & Pallais, 2017; Wiswall & Zafar, 2015).

The paper is organized into four sections. Section 2 describes the context and sample, patterns of measured and perceived skills in the sample, and a simple model. In Section 3 we show the relationship between comparative advantage beliefs and skill-directed job search in the tight field experiment. In Section 4 we show the relationship between comparative advantage beliefs, skill-directed job search, and labor market outcomes in the big field experiment. We show that search effort is unlikely to be the mechanism behind our results in Section 5. Section 6 discusses the role of additional mechanisms.

2 Economic Environment

This section documents five theoretically relevant stylized facts about our setting. Jobseekers have multidimensional skills: skills in more than one dimension which are not perfectly correlated. Firms can partly observe skills and value different skills for different jobs. Jobseekers' beliefs about their comparative advantage in skills often differ from their comparative advantage measured on assessments commonly used by firms. Jobseekers learn little about their skills while searching: the gap between perceived and measured skills remains wide. Beliefs predict job search targeting: how jobseekers direct search effort across different job types.

We begin with a simple conceptual framework to define a comparative advantage in skills and link jobseekers' beliefs about their comparative advantage to their search effort and labor market outcomes. We then describe the samples of jobseekers and how we assess their skills and provide descriptive evidence of the stylized facts in our setting.

2.1 Conceptual Framework

Our conceptual framework is directly designed to show how the skill belief patterns we later document can influence job search activities and outcomes, in a similar spirit to Mahoney (2022). It 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, but is more stylized and not designed for estimation (Lentz et al., 2022; Wu, 2021).

Informed by the stylized facts in our data, and in particular limited evidence of learning about skills during search, we use a static framework and do not model dynamic belief updating in response to search outcomes. Given the types of data we observe, we use a partial equilibrium framework focusing on jobseekers' search and beliefs but treating labor demand and wage posting as fixed.

Comparative advantage in skills: In this framework, each jobseeker is endowed with communication skill level S_C and numeracy skill level S_N.[6] These can be assessed by firms, schools or job centers . Based on assessments, jobseekers' skills can be ranked relative to a reference population of other jobseekers from similar backgrounds with similar education levels in the same area, and indeed many education systems generate such rankings. We classify jobseekers into skill quintiles for both skills relative to a reference group.[7]

We classify a jobseeker as having a **comparative advantage in communication** if they score in a higher quintile for communication than numeracy. This captures the spirit of standard definitions of comparative advantage: for a given time endowment, each jobseeker has a comparative advantage in the skill that they can supply more of to the market than other "suppliers", who are other jobseekers from a similar background who are likely to be applying for the same jobs. This approach is relatively rare in labor economics because multidimensional skill measures are available in few datasets, but it follows existing work that uses similar data (e.g. Guvenen et al. 2020).

This approach does not capture the jobseeker's rank in the skill distribution of applicants for any specific job, which would require very different data sourced from a large

[6]We assume that the two skills have the same distributions, so absolute and comparative advantage are aligned. This simplifies the algebra and matches our data, which measures skills in both domain using rankings not levels.

[7]We use ordinal measures of skill and skill beliefs throughout the paper. We prefer these to cardinal skill measures (e.g. 80%) because cardinal measures are dependent on assessment-specific scales.

number of firms. Nor does it capture the jobseeker's relative rank in the skill distribution for the entire country, which would require nationally representative data. But our approach does take advantage of the direct measures of multidimensional skills we have available, compares people to the relevant reference group (unlike the nationally representative samples) and avoids the assumptions used in some other definitions of comparative advantage. For example, some labor economists estimate occupation-specific wages for either a single skill or for education level and use this to define comparative advantage in occupations based on levels of occupation or skill (e.g. Acemoglu & Autor 2011; Gibbons et al. 2005). These approaches have the advantage of pricing 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 occupations. These are sensitive to a Lucas-style critique that wages might be different under a different type of sorting, which is precisely the type of mechanism we study.

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. Given skills and search effort, jobseekers receive an offer for job type j with probability $P_j(S_C, S_N, E_j)$ with associated "wage" $W_j(S_C, S_N, E_j)$, which is a reduced-form expression for the expected present value of the offer. We define $V_j(S_C, S_N, E_j) = P_j(S_C, S_N, E_j) \times W_j(S_C, S_N, E_j)$ and assume V_j is increasing and concave in all three arguments and that

$$\partial V_j / \partial S_j > \partial V_j / \partial S_i > 0. \tag{1}$$

for $j \neq i$. This assumption allows both types of jobs to value both skills, but each type of job to value one skill more than the other. We also assume that skill and search effort are technical complements and are 'more complementary' within than across dimensions. Intuitively, this means that a jobseeker with high numeracy skills will get a higher return to directing marginal search effort to numeracy jobs than communication jobs and vice versa. Formally,

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

for $j \neq i$. Finally, we assume that the gross utility from job search $U(V_C, V_N)$ is increasing and concave in both arguments. We simply assume that jobseekers derive utility from offers, rather than modeling the offer acceptance decision or reservation wage.

The jobseeker directs search effort to equalize the marginal utility of searching for each job type:

$$\frac{\partial U}{\partial V_C} \times \frac{\partial V_C}{\partial E_C} = \frac{\partial U}{\partial V_N} \times \frac{\partial V_N}{\partial E_N}, \tag{3}$$

where $\frac{\partial U}{\partial V_j}$ captures the jobseeker's preferences over non-pecuniary aspects of the job types. 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 assumptions (1) and (2) , $\frac{\partial V_C}{\partial E_C}$ is more steeply increasing in communication skill than $\frac{\partial V_N}{\partial E_N}$. So, to satisfy the first order condition in (3), a jobseeker with higher communication skills, i.e., with comparative advantage in communication will direct more effort to searching for communication than for numeracy jobs.

Jobseekers' beliefs about their skills: We assume each jobseeker has beliefs about their skill levels $\tilde{S}_C$ and $\tilde{S}_N$. In this framework, jobseekers direct search based on their skill beliefs but their search outcomes depend on their actual skills. Or, in a slightly more general framework, search outcomes could depend on noisy signals of skills received by firms.

Predictions: This framework delivers two testable predictions that we evaluate in our experiments:

- P1 Jobseekers will direct more search effort to jobs requiring the skill in which they have a perceived comparative advantage to satisfy condition (3), irrespective of their actual comparative advantage.
- P2 Jobseekers whose perceived comparative advantage does not match their actual comparative advantage will receive fewer offers, lower wages conditional on offers, or both, due to assumption (2).

We use multiple measures of search effort and outcomes, defined below, to test these predictions.

2.2 Sample

We study young, active jobseekers with high school education, but few further qualifications, from low-income backgrounds in Johannesburg, South Africa. We recruit our samples from participant registries at the Harambee Youth Employment Accelerator, a social enterprise funded jointly by the South African government and private firms to provide job search assistance services to young jobseekers from low-income backgrounds. Since 2013, Harambee has maintained a database of active jobseekers recruited through advertising in traditional and social media and provided firms with access to the database to recruit entry-level workers. Harambee often contacts jobseekers about training opportunities, job search workshops, or invitations to take assessments, either for practice or to determine eligibility for specific jobs.

Jobseekers sign up to Harambee's services with their national identity number, which Harambee uses to screen out those not aged 18-34 or without legal permission to work in South Africa. They self-report actively searching for work and having attended a township or rural high school .[8]

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. In both cases, we contact a random sample of people who live within commuting distance of our field location in downtown Johannesburg. We invite those who actively search for work to a day-long job search assistance workshop. The workshop is advertised to be similar in format to other workshops run by Harambee; including taking standardized skill assessments, answering questions about their job search, and receiving job search advice. For both experiments, jobseekers are told the assessments could be used to match them to suitable vacancies.

In 2016, any matching between firms and jobseekers was entirely administered by Harambee. In 2019, Harambee set up an online job search and matching platform SAYouth.mobi. The platform allows firms to post advertisements directly, aggregates job advertisements from alternative sources and offers data-free access, so jobseekers in our sample describe it as a key part of their job search strategy. In 2022, there were 3,739,861 jobseekers aged 18-34 on the platform in South Africa of whom 1,078,745 were in Johannesburg.

Survey data: In both experiments, participants take skill assessments and complete self-administered surveys about their demographics, beliefs about their skills, beliefs about their job search, recent job search activities, current employment, and employment history before treatment. We leave discussion of post-treatment measurement to sections 3 and 4.

Sample characteristics: We focus here on descriptive statistics for the tight experiment sample, shown in Table 1. These jobseekers are actively engaged in the labor market but likely to face limited information about their skills in the labor market. They are young and so have limited time to learn about their skills from search or employment (interdecile age range 21 - 32). Only 23% had completed any post-secondary education, limiting their scope to learn from specialized training. They have limited, mostly informal

[8] "Townships" are urban areas which were zoned as being for black African residents only under apartheid and schools remain of poor quality.

work experience: 33% were employed at baseline but only 13% had a formal written contract and only 25% had *ever* held a long-term wage job.[9]

Their job search effort was high but met with limited success. 96% were actively searching and the average jobseeker submitted 10 job applications, spent 14 hours searching for work, and spent 22.72 USD on search expenses in the week before the baseline. This search cost is high – equal to 50% of the average weekly earnings of employed people in our sample. 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 due to their rarity.)

2.3 Jobseekers Have Multidimensional Skills

For simplicity, in the tight experiment, we focus on two general-purpose, job-relevant skills that we verify firms in this context value in prospective workers: communication and numeracy. The numeracy assessment focuses on practical arithmetic and pattern recognition. It was developed by a large retail chain and used in their applicant screening process, as they believe it identifies some of the skills needed by cashiers. The communication assessment captures English language listening, reading and comprehension skills. The assessment was developed by a South African adult education provider (www.mediaworks.co.za) and is designed to assess English proficiency for high school students.[10] Candidates also complete a measure of "concept formation" (Taylor, 1994), similar to a Raven's test (Raven & Raven, 2003). This is a non-verbal measure of fluid intelligence, people's ability to see underlying commonalities across situations and to use logic in new situations.

We describe the assessments, their psychometric properties, the assessment process, and the assessment result distributions in Appendix C. The distributions show a wide range of assessment scores and little evidence of ceiling or floor effects (Figure A1). These assessments are also used in the big experiment, along with four additional assessments that we describe in Section 4.2.

As discussed in Section 2.1, we focus on ordinal measures of skills relative to a benchmark population. Our reference group consists of roughly 12,000 jobseekers assessed by

[9]Throughout the paper, we define an employed person as someone who did any income-generating activity during the survey reference period, following Statistics South Africa's definition. Unemployment rates exclude those in full-time education or not in the labor force.

[10]For jobseekers, we describe numeracy skills as "working with numbers. It includes using addition, subtraction, multiplication and division to solve real problems involving money, time, and quantities". Communication skills are "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."

Table 1: Summary Statistics

	(1) Mean	(2) Median	(3) Min	(4) Max	(5) SD	(6) Obs.
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	0.69	1.00	0.00	1.00	0.46	278
University degree / diploma	0.09	0.00	0.00	1.00	0.28	278
Other post-secondary education	0.14	0.00	0.00	1.00	0.35	278
Panel B: Labor market background						
Any work in the 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)	44.95	0.00	0.00	702.75	102.07	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	3.00	0.52	278
Panel D: Skills beliefs						
Aligned belief about comparative advantage	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.** Winsorized variables are winsorized at the 99^{th} percentile. All monetary values are in 2021 USD purchasing power parity terms.

Harambee in Johannesburg using the same assessments randomly sampled from Harambee's database who meet the same criteria: they are aged 18-34 and completed high school in townships around Johannesburg or rural areas. We place jobseekers in quintiles relative to this population.

In the full sample for the tight experiment, communication and numeracy scores have correlation 0.31 and only 17% of the jobseekers are in the top two quintiles for both skills. Thus these two assessments horizontally rather than vertically differentiate jobseekers: they show which jobseekers are better suited for specific types of jobs, rather than identifying jobseekers who are likely to be better at all or most job types. Consistent with this, communication and numeracy quintiles are both moderately correlated with concept formation quintiles ($\rho \approx 0.25$ for both skills).

In the tight experiment, we restrict our sample to 278 jobseekers with a clear comparative advantage in skills. This means that we drop the 24% of jobseekers who have the same quintile for both skills in the full sample. This sample restriction is important because our experimental measures of skill-directed job search, described in Section 3.4, can only be defined for jobseekers with a unique skill comparative advantage. Table A4 shows that 62% of our restricted sample has a comparative advantage in communication and 38% has a comparative advantage in numeracy. Note that the shares of jobseekers in this sample with communication and numeracy comparative advantage are not equal because we compare jobseekers' skills against a separate benchmark population of 12,000 jobseekers. In this sample, communication and numeracy skill quintiles have zero correlation.

2.4 Firms Value Different Skills for Different Jobs

Searching based on the skill comparative advantage we measure may impact jobseekers' labor market outcomes under three conditions: 1) firms value communication and numeracy skills; 2) firms' relative valuation of the two skills varies, so it is optimal for different jobseekers to apply to different firms; and 3) firms can at least partly observe jobseekers' skills. Here we show these conditions hold in this context. This is not, of course, evidence that skill-directed job search improves labor market outcomes; the experiments are designed to do that.

Firms value communication and numeracy skills: Three patterns show that numeracy and communication skills are valued by firms in this labor market, suggesting that jobseekers with skills more suitable to a vacancy may be more likely to be hired or offered higher wages.[11] First, in an incentivized resume-ranking experiment, described in

[11] We do not claim that these are the only or necessarily even the most important skills in this labor

Appendix B.2, 91% of many firms preferred job applicants with top tercile higher communication or numeracy assessment results to job applicants with middle tercile skills and an additional six-month post-secondary training qualification. Second, over 500 client firms have paid Harambee to screen roughly 500,000 prospective workers using the communication, numeracy, and concept formation assessments, which we interpret as revealing a preference for using these skills in hiring. Finally, Carranza et al. (2022) use multiple field experiments to show that firms are willing to pay to learn jobseekers' results on these skill assessments and that jobseekers who can certify their assessment results for communication, numeracy, and a few other skills have higher employment rates.

Firms' relative valuation of skills varies: In the same incentivized ranking experiment mentioned above (and detailed in Appendix B.2), 58% of the firms in our sample ranked applicants with high numeracy skills ahead of the profile with high communication skills and 42% of firms had the opposite ranking. This cross-firm variation, combined with the fact that our assessments horizontally differentiate jobseekers on communication versus numeracy skills, creates scope for jobseekers to improve their labor market outcomes by searching for different types of jobs.

Firms' 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 the comparative advantage. This suggests that redirecting jobseekers' search towards jobs that match their comparative advantage in skills has the potential to improve their labor market outcomes.

2.5 Jobseekers' Perceived and Measured Comparative Advantage in Skills Differs

We measure jobseekers' beliefs about their communication and numeracy skill quintiles before they take assessments (see Appendix C.1 for measurement details and Figure A3 for the precise ordering of measurement relative to assessments and treatment). We define communication skills and numeracy skills, explain the concept of quintiles, define the reference group and ask jobseekers which quintile their communication and numeracy skills are in, relative to the reference group.

Misaligned beliefs about comparative advantage: Our main measure of the relationship between a jobseekers' perceived and measured relative rank on skills is their **aligned**

market. Employers may also value other skills: we simply argue that communication and numeracy skills are, on average, valued by firms in this labor market.

comparative advantage belief: an indicator equal to one if the jobseeker has a higher perceived quintile for the skill in which they have a higher assessed quintile. Before taking assessments, only 49% of jobseekers' beliefs about their comparative advantage in general skills align with their comparative advantage on our skill assessments (Table 1, panel D). This is slightly better than the 40% of aligned beliefs that would arise from random guessing of skill quintiles.[12]

Jobseekers' beliefs about their results on our specific assessments are similarly misaligned. Straight after taking assessments but before treatment, we ask jobseekers which quintile they fall in on our assessments. Only 54% percent of jobseekers assess their relative assessment performance correctly.[13]

Misaligned beliefs about skill levels: We can also examine whether workseekers have aligned skill level beliefs for communication and numeracy. To create one simple measure, we refer to the **fraction aligned beliefs**: 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 fraction of aligned beliefs can take values of 0, 0.5 and 1. Averaging across all jobseekers, only 22% of believed general skill quintiles equal assessed skill quintiles and only 17% of believed assessment result quintiles equal assessed skill quintiles (Table 1, panel D).[14] This pattern is driven more by overconfidence than underconfidence: 60% of jobseekers' perceived general skills are above their assessed skills and only 17% are below, in line with patterns documented in many other settings (Santos-Pinto & de la Rosa, 2020).

(Lack of) differences in beliefs by gender: We focus on gender-pooled results throughout the paper. We find no gender differences in baseline skill beliefs (Tables A48 and A49) and we observe very limited gender heterogeneity in the relationship between measured and perceived skills conditional on demographic differences in both the tight and big experiments (Tables A48 and A49).

2.6 Jobseekers Learn Little about their Skills While Searching

To analyze the persistence of misaligned beliefs over time, we use the big experiment's follow-up survey, which occurred on average 3.5 month after baseline. The share of con-

[12]The random benchmark is below 50% because our sample has by definition no ties in their skill quintiles. This implies that jobseekers with tied beliefs are always coded as misaligned.

[13]These two beliefs are strongly positively correlated within skill, suggesting that jobseekers view the assessments as relevant to their general skills. Regressing general skill belief on belief about assessment results produces coefficients of 0.39-0.52 across the two skills, with or without controls for assessment results and demographic characteristics.

[14]Comparative advantage beliefs and the fraction of aligned beliefs are positively correlated by construction but have some separate variation, generating a correlation coefficient of 0.21.

trol group jobseekers whose believed and assessed comparative advantage align changes little from baseline to the endline, even for those with above-median search effort and those who are employed (Table A13, columns 1-3). The same persistence holds for the fraction of aligned beliefs (columns 4-6).

The strong persistence of misaligned beliefs emphasizes that this labor market offers limited scope for jobseekers in our sample to learn about their skills, at least for the time horizon we study. It suggests that a dynamic conceptual framework incorporating belief updating through search or work is not needed in this setting.

What labor market conditions might lead to persistent inaccurate beliefs about one's relative rank in skills? First, the education system provides noisy signals of jobseekers' skills. Most of our sample have secondary education as their highest level of qualification and, due to Harambee's eligibility criteria, all attended school in a low-income neighborhood. In such schools, grades and grade progression are weakly correlated with results in independent skill assessments (Lam et al., 2011; Taylor et al., 2011). There is a nationally standardized secondary school graduation examination, but grades in this exam only weakly predict real-world outcomes like performance in post-secondary education (Schoer et al., 2010). This means that the existence of this exam does not entirely eliminate the information frictions we study. Schools in these neighborhoods rarely have specialized career counselors and few teachers have received training to offer support in choice of career (Pillay, 2020).

In our data, communication and numeracy quintiles are positively but weakly associated with jobseekers' self-reported grades on their high school leaving exams in respectively English and mathematics (Table A11, columns 1-2). The difference between English and mathematics scores on the exam is also positively associated with jobseekers' comparative advantage in our assessments (column 3).

Even with good information from the schooling system, exam scores should not perfectly predict beliefs: the average jobseeker took the exam multiple years ago, the exam and assessments do not test identical skills, and jobseekers may not perfectly recall their exam scores. Skill beliefs about communication and numeracy are positively correlated with jobseekers' self-reports about their results on the high school graduation exams in respectively English and mathematics (Table A11, columns 4-5). Beliefs about comparative advantage are positively associated with the difference in scores between the two exam subjects (columns 6-7). However, correlations are fairly low, showing that jobseekers' beliefs depend on multiple sources of information.

Second, the employment rate is low, especially for formal jobs. At the time of the tight experiment, unemployment in Johannesburg was 33.7% for the working-age population and 40.5% for ages 15-34 (Statistics South Africa, 2022). In our sample, unemployment is even higher at 67%, most likely because our sample omits people outside the labor force. Jobseekers' limited experience of employment limits their scope to learn about their skills through work experience, the main mechanism in models of worker-side learning in other contexts (Baley et al., 2022; Guvenen et al., 2020). Learning about skills by search is also difficult: only 3% of jobseekers report ever receiving feedback about their skills during an unsuccessful job application.

Third, job search costs are high. One cause is the high cost of transport: Banerjee & Sequeira (2020) find that transport costs make up more than 50% of reported monthly income in a similar sample. Kerr (2017) documents average transport cost of 11% of income among a working population. In our sample, the average person spends half as much on search each week as the average employed person earns, covering transport, mobile phone data, and printing applications. This further limits the scope for jobseekers to learn about their skills through search and raises costs of misdirected job search.

2.7 Beliefs Predict Job Search Targeting by Skill

In the big experiment, we ask candidates what skill is most valuable for the types of jobs they are applying for. In the control group, 47% and 22% answer communication and numeracy respectively and these answers are strongly linked to their beliefs about their comparative advantage (Table A17). Jobseekers are 6-10 percentage points more likely to state that the skill of their comparative advantage was most important for the jobs that they are applying for compared to choosing any of the other skills (17-45% of the mean, $p < 0.01$). The correlations are robust across skill domains and to controlling for measured comparative advantage and demographic controls. We find a similar pattern in the tight experiment using different measures of search direction, which we introduce in Section 3.4.

This link between search direction and comparative advantage beliefs motivates an analytical framework that studies how jobseekers' perceived comparative advantage in skills shapes the direction of their search effort between jobs with different skill requirements.

Jobseekers believe skills are valued: Two patterns in survey data from the big experiment show that jobseekers believe there are positive labor market returns to skills. First, individual jobseekers think their outcomes would improve if their skills were higher. We asked each jobseeker their expected search duration and earnings conditional on finding a job and then asked their expectation for another jobseeker who had better numeracy skills but was otherwise identical to themselves. Control group 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 118.46 USD PPP more (13% of the mean, $p < 0.01$). We did not ask these questions about communication skill due to survey time constraints.

Second, jobseekers with higher skill beliefs expect to earn more and get a job faster. This is a between-jobseeker comparison , so it is vulnerable to omitted variable bias, but the pattern is robust to controlling for actual skill levels, age, gender, education, and work history. A one standard deviation higher average skill belief is associated with 5% higher expected wage ($p < 0.001$) and 3.5% shorter expected search duration ($p < 0.001$) (Table A16).

These two patterns motivate a conceptual framework where information about skills might alter beliefs about the probability of getting a job and wages conditional on getting a job.

3 Tight Experiment: Effects on Beliefs and Behavioral Measures of Directed Search

In this section, we report the results from the tight experiment. This collects rich beliefs data and collects unique measures of skills-directed job search by directly observing jobseekers' choices between jobs with different skills content at the jobseeker$\times$job level. Measuring search direction across jobs with different skill demand is difficult in labor market research that relies on survey data from jobseekers, which typically record aggregated search data such as total applications submitted, not application-level data including skill demand. Job search platform data seldom measure skills, beliefs, or labor market outcomes. Administrative data from unemployment insurance systems do not measure beliefs and seldom measure skills or search.

3.1 Experimental Design

We work with a sample of young jobseekers who have signed up to attend a job search assistance workshop from a social enterprise. We randomize treatment by day of the workshop, so all jobseekers at a workshop receive the same treatment, to reduce risks of spillovers. We assign 17 workshops, each conducted on a separate day, to treatment and 17 to control. Treatment assignments are balanced on baseline covariates, both in the full sample and in the subsample with a clear comparative advantage in skills (Table A19).

The timeline of the day is shown in Figure A3. On the day of the workshop, jobseekers respond to a pre-treatment survey. In this survey 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 of similar jobseekers. They then take assessments of their numeracy, communication, and concept formation skills and complete a brief survey about their perceived performance on the assessments. To cleanly define beliefs about comparative advantage and job choices aligned with comparative advantage in skills 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. Jobseekers who have the same quintile in both skills have no clear comparative advantage, so we cannot sensibly define their job search choices as aligned or misaligned with their comparative advantage.[15]

Treated jobseekers 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 it does not encourage them to apply to any specific types of jobs.

Jobseekers in the control group do not receive a report. They take the assessments and answer skill-focused survey questions, so any priming from these to place more weight on skill levels and skill match during the post-treatment surveys and tasks is constant across treatment groups. They watch a control video, which contains the parts of the treatment video which explains the assessments to hold constant any effect of having better information about communication or numeracy skills. The control video also contains the encouragement to jobseekers to think about what jobs will value their skills to hold constant the general idea of directing search on beliefs or search behavior.

To facilitate comprehension, we piloted reports and videos intensively and gave jobseekers time to ask questions during and after the video. After the video, we ask treated jobseekers to report what the assessment results show about their skill comparative advantage. 95% of treated jobseekers answer this comprehension check correctly.

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 6.3 that information acquisition by firms is unlikely to explain the results of the experiment.

[15] However, including these jobseekers in the sample and classifying all their job search choices as misaligned with comparative advantage produces qualitatively similar treatment effects on skill beliefs (Table A22) and search (Table A23).

Figure 1: Sample Report

REPORT ON CANDIDATE COMPETENCIES
-Personal Copy-

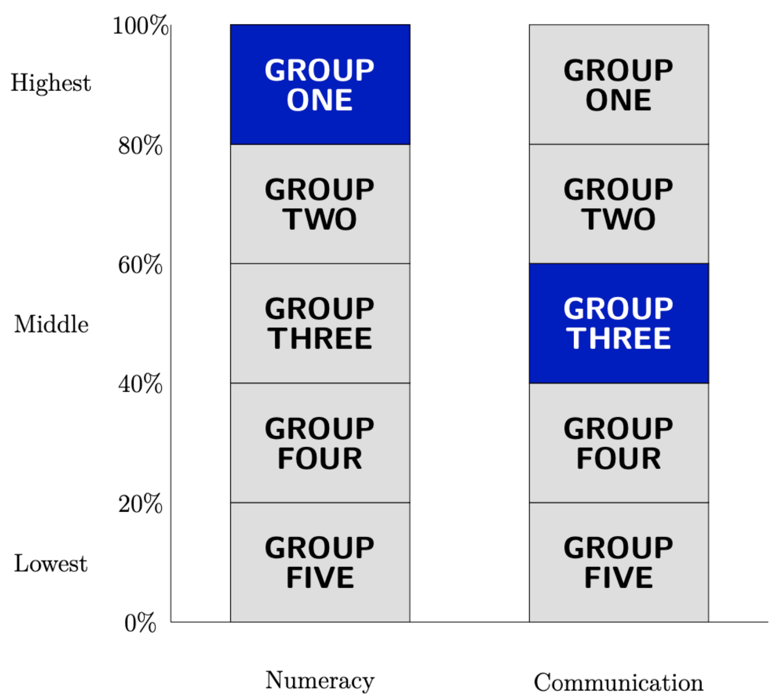
This report contains results from the assessments you took today. These results can help you learn about some of your strengths and weaknesses and inform your job search.

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 and no branding.

3.2 Specification

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

$$Y_{id} = T_d \cdot \beta + \mathbf{X}_{id} \cdot \Gamma + \epsilon_{id}, \tag{4}$$

where 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.[16] We use heteroskedasticity-robust standard errors clustered by assessment date, the unit of treatment assignment.

We also test if increased alignment between perceived and measured skills and of search direction by skill is concentrated among jobseekers whose baseline beliefs are misaligned with their comparative advantage. We estimate models of the form:

$$Y_{id} = T_d \cdot \alpha^{misaligned} + T_d \cdot Aligned_{id} \cdot \alpha^{aligned} + Aligned_{id} \cdot \delta + \mathbf{X}_{id} \cdot \Gamma + \epsilon_{id}, \tag{5}$$

where $Aligned_{id}$ is an indicator for jobseekers whose beliefs about their comparative advantage on the assessments match their measured comparative advantage at baseline (see Appendix C.1 for measurement details). These jobseekers receive less information from treatment. We focus on the average treatment effect for jobseekers with misaligned baseline comparative advantage beliefs, $\alpha^{misaligned}$, and the difference in average treatment effect between jobseekers with misaligned and aligned baseline beliefs, $\alpha^{aligned}$.

As in any heterogeneity analysis, $Aligned_{id}$ may be correlated with other jobseeker-level characteristics, complicating the interpretation of $(\alpha^{misaligned}, \alpha^{aligned})$. This is a relatively minor concern here, because having an aligned comparative advantage belief at baseline is unrelated to gender, age, employment, and work experience. These variables jointly explain only 3% of the variation in aligned comparative advantage beliefs (Table A12). This belief is, unsurprisingly, related to the communication and numeracy assessment scores, but these scores explain only 15% of the belief variation.

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 L.

[16] $\mathbf{X}_{id}$ contains age; a dummy for being female; dummies for having a high school leaving certificate, 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; dummies for having a comparative advantage in either of the two skills; a pre-treatment value of the outcome Y_{id} where available; and block fixed effects, to account for the fact that we randomly assign days to treatment groups within blocks of 4 sequential days.

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

	Aligned comp. adv. 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 comp adv belief (bl)		-0.137		0.080
		(0.082)		(0.050)
Aligned comp adv belief (bl)		0.586***		-0.050
		(0.079)		(0.046)
Treatment effect:		0.072		0.116***
Aligned comp adv 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 the treatment aligns beliefs about skills with their assessed skills 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 average absolute deviation of beliefs about skill quintiles and measured skill quintiles. Columns 2, and 4, show treatment effect heterogeneity by whether jobseekers had aligned comparative advantage beliefs at baseline. Controls include randomization block fixed effects, and prespecified baseline covariates (skill quintile dummies, age, gender, and dummies for having completed a university degree, other post-secondary education, and high school). Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

3.3 Information About Skills Aligns Beliefs with Comparative Advantage

Jobseekers who receive information about their skills are more likely to have beliefs about their skills that are aligned with their assessed 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. We ask about their beliefs about their skills in general, abstracting from our specific assessments. Treated jobseekers are on average 13.5 percentage points more likely to report aligned beliefs, a 28% increase (Table 2, column 1, $p = 0.001$). Section C.1 describes measurement of beliefs.

Most of the treatment effect is driven by jobseekers with misaligned comparative advantage baseline beliefs. Treatment increases the proportion of these jobseekers whose beliefs align with their measured comparative advantage by 21pp (a 44% increase over the control mean, $p < 0.001$). This is the term $\alpha^{misaligned}$ in Equation 5 and the coefficient on the treatment dummy in row 1, column 2, Table 2.

In contrast, treatment has smaller effects on jobseekers with aligned baseline beliefs. The proportion of this group with aligned beliefs increases by an insignificant 7pp (row 4, column 2, Table 2). This is the sum of the treatment effect for misaligned jobseekers $\alpha^{misaligned}$ and the interaction $T_d \cdot Aligned_{id} \cdot \alpha^{aligned}$ in Equation 5. The difference between jobseekers with aligned and misaligned baseline beliefs is large – 14pp, or 30% of the baseline control mean – but marginally insignificant. This is $T_d \cdot Aligned_{id} \cdot \alpha^{aligned}$, shown in row 2, column 2, Table 2.

Treatment shifts the level of skill beliefs as well as beliefs about skill comparative advantage. In the control group, 18% of jobseekers have believed skill quintiles which match their assessed skill quintiles. Treated jobseekers are 7.8pp to have believed skill quintiles which match their assessed skill quintiles, a 43% increase (Table 2, column 5, p=0.005). We focus less on heterogeneity by baseline comparative advantage beliefs for this measure, because aligned comparative advantage beliefs and aligned skill level beliefs differ for many jobseekers so baseline comparative advantage beliefs do not reflect what individuals might learn about their skill levels.

3.4 Job Search with Better Aligned Beliefs about Skills

We now study effects of receiving information on about one's relative ranking on skills on search behavior. We find robust evidence that jobseekers redirect search toward jobs that align with their assessed comparative advantage across four measures of skill-directed search. These differ on several features, such as realism, incentivization, timing, and susceptibility to experimenter demand effects.

Job choice task: We design a novel incentive-compatible job search task in which we asked respondents to make 11 choices between pairs of job advertisements, one of which was coded by recruiters as requiring numeracy skills and one of which was coded as requiring communication skills. Participants were shown each pair of advertisements on a piece of paper, given time to read them, and asked by the enumerator to select which they would like to apply to.

The job advertisements are based on real job advertisements on (SAYouth.mobi). All jobseekers were active platform users so the choice represented a real-life choice they often made during job search. We used only job advertisements for entry-level positions in the Johannesburg area that required no special qualifications or training. Among these, we selected a long-list of 28 vacancy postings which had a clear numeracy or communication skill requirement, recognizing that this is not the case for all jobs to which jobseekers might apply. We asked 13 human resources and recruitment staff with experience hiring for entry-level roles to rate the extent to which vacancies required communication and numeracy. They also predict the expected wage, overall and gender-specific desirability and transparency of skill requirements of each vacancy. Averaging across all pairs, our 13 placement officers scored the job we defined as numeracy-heavy as needing 2.7 standard deviations more numeracy skills and likely to pay only 0.3 standard deviations more. We removed information on the employer name and location to reduce their role in job choices and standardized length and format.

We then created pairs of jobs which had opposite skill requirements but similar expected wages and desirability levels in other respects, based on recruiters' rankings, allowing us to isolate the role of skill-directed job search. The treatment gives information on only communication and numeracy skills. Together, these elements of the design allow simple measurement of how skill-directed job search over jobs requiring two different skills changes in response to information about one's relative ranking on those skills. In contrast, in a field setting, it is possible that jobs with different skill content also differ on unobserved characteristics which drive observed search direction. Similar challenges apply to work studying how job applications on search platforms respond to posted wages (Belot et al., 2022b; He et al., 2021). Figure 2 shows an example pair and Appendix Table A18 shows the full list of job titles. Participants viewed the 11 job advertisement pairs in a randomized order.

We incentivize jobseekers to respond truthfully in the job choice task in two ways. First, one of the 11 pairs of vacancies were live advertisements for jobs with a partner firm. Jobseekers were told when invited to the workshop that they would receive assistance in writing their CV and cover letter for a job. Before the task, we told jobseekers that one of the choices they made would select the job which we would assist them to apply to. They did not know which pair would be selected. Second, before the task, we told jobseekers that after the workshop, they would get a list of entry-level job titles similar to the jobs they preferred in the task to assist them in job search in the future, and gave them the list.

We measure the impact on the share of aligned applications: the share of 11 job choices that were in line with participants' measured comparative advantage. In the control group, 55% of jobseekers make an aligned job choice. Treated jobseekers have, on average, a 4pp higher share of job choices that align with their measured comparative advantage (Table 3, column 3, $p = 0.29$). This effect hides important heterogeneity. Among jobseekers with misaligned baseline beliefs about their comparative advantage, treated

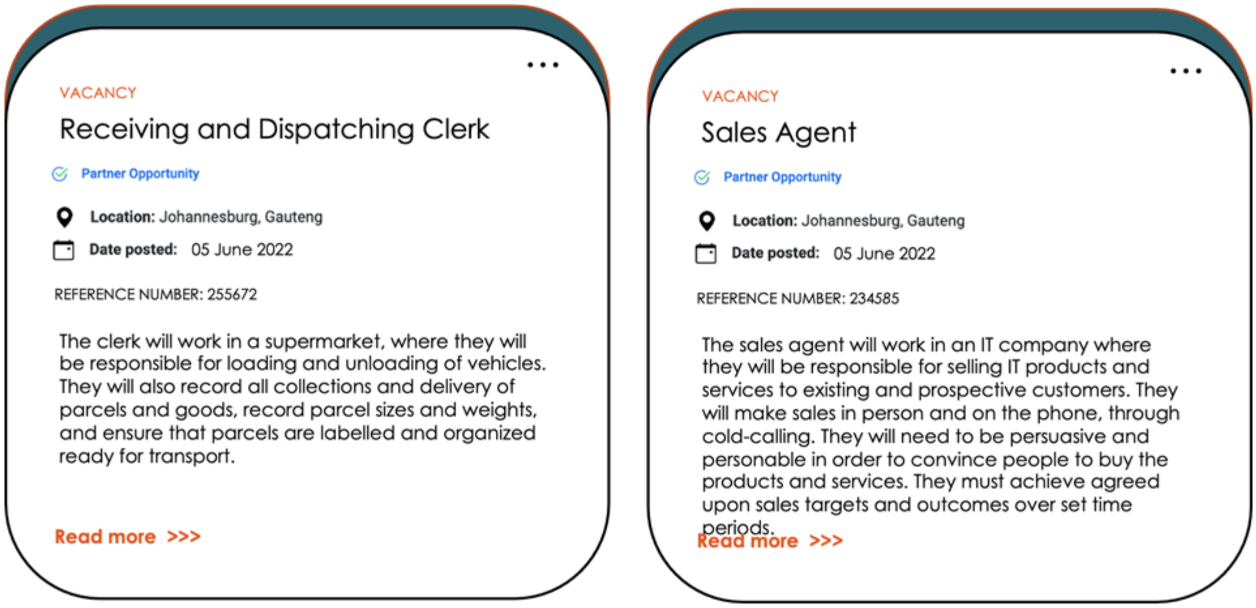


Figure 2: Sample Pair of Jobs from Job Choice Task

jobseekers have an 8.8pp higher share of aligned job choices, a 16% increase relative to the control mean (Table 3, column 4, $p = 0.024$). Jobseekers with initially aligned beliefs make few changes in their search direction in response to treatment, so the difference between jobseekers with aligned and misaligned initial beliefs is large and statistically significant.

This heterogeneity is consistent with heterogeneity in comparative advantage beliefs: jobseekers with initially unaligned beliefs about their comparative advantage in skills were more likely to update their beliefs to align with their comparative advantage than those with initially aligned beliefs. The observed treatment effects are also in line with the model predictions of a shift in search effort towards jobs that align with jobseekers' comparative advantage, particularly among individuals with initially unaligned beliefs.

The search direction results suggest that jobseekers have some information about the relative skill demands of different jobs. However, they might still face information frictions about the jobs' skill requirements. We test this by explicitly revealing relative skill demand for the last two pairs of jobs (recall that order is randomized). Table A32 shows that 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. This suggests that jobseekers have some but not complete information about the skills requirements of some of the jobs.

Application data from online search platform: We recruited the sample from the jobseekers on SAYouth.mobi and can observe their on-platform job search post-treatment.

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)	(0.139)	(0.034)	(0.038)	(0.023)	(0.048)	(0.061)	(0.096)	(1.261)	(1.763)
Treatment × Aligned comp adv belief (bl)		-0.647***		-0.104**		-0.048		-0.163		-6.629**
		(0.205)		(0.039)		(0.080)		(0.128)		(2.651)
Aligned comp adv belief (bl)		0.726***		0.165***		0.013		0.100		8.697***
		(0.153)		(0.035)		(0.049)		(0.102)		(2.412)
Treatment effect:		-0.044		-0.016		0.041		-0.005		-1.883
Aligned comp adv belief (bl)		(0.131)		(0.034)		(0.044)		(0.081)		(1.807)
Control mean	0.000	0.000	0.550	0.550	0.006	0.006	-0.032	-0.032	4.331	4.331
Observations	278	278	278	278	278	278	278	278	278	278

Notes: **Table 3 shows that providing jobseekers with information about their relative comparative advantage in skills aligns their search direction with their assessed comparative advantage in skills in the tight experiment.** Aligned job search is defined as directing search effort toward jobs that mostly require the skill that aligns with job seekers' measured comparative advantage. Columns 1 and 2 show impacts on an index of search direction. The index is constructed as the variance-covariance weighted average of the search alignment measures displayed in columns 3 to 10 (Anderson, 2008). Columns 3 and 4 show the impact on the percentage of 11 incentivized job choices that are aligned with the comparative advantage. Columns 5 and 6 show effects on the difference between the percentage of aligned and non-aligned applications on the online job search platform SAyouth.mobi. Columns 7 and 8 show the effects on the difference in link click rates between aligned and non-aligned jobs sent to job seekers via SMS. All job seekers receive three links with at least one job for each skill. Columns 9 and 10 show effects on the difference between aligned and non-aligned planned applications for the 30 days after the workshop. Even columns show heterogeneity by whether individuals have aligned comparative advantage beliefs at baseline. Controls include randomization block fixed effects, and pre-specified baseline covariates (skill quintile dummies, age, gender, dummies for having completed a university degree, other post-secondary certificate or diploma, and high school, and a dummy indicating a comparative advantage in numeracy). Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Of the 69,000 vacancies available on SAYouth.mobi in Johannesburg during our study period, we classify 14% as communication-heavy jobs and 13% as numeracy-heavy jobs, suggesting demand is roughly equally distributed across skills. [17]

To study search direction, we focus on whether a jobseeker initiated an application to a job, as we do not observe completed applications for all job types. Post-treatment, 23% of applications initiated by jobseekers in our sample are to numeracy or communication jobs. Jobseekers in our sample actively used the platform in the 30 days prior to treatment, visiting it on 3.2 of those days and initiating 6.5 job applications (Table A1). In the 30 days after the workshop, both the treatment and control group become more active, initiating applications to 15 jobs on the platform.

Treated jobseekers are more likely to initiate applications to jobs that match their measured comparative advantage. The outcome is the difference between the fraction of applications to vacancies coded as aligned with jobseekers' comparative advantage in skills and the fraction of applications requiring the opposite skill. Pre-treatment, the difference is close to zero, suggesting jobseekers rarely direct jobs toward vacancies matching their comparative advantage in skills. Treated jobseekers submit 6.3 percentage points more aligned applications than non-aligned applications (Table 3, column 5, $p = 0.01$). Again, this effect is driven by jobseekers with initially misaligned comparative advantage beliefs (8.9pp, $p = 0.072$), compared to jobseekers with initially aligned comparative advantage beliefs (4.1pp, $p > 0.1$), though the difference is not significant. Participants' information was obtained after the experiment from administrative data and is unlikely to be susceptible to experimenter demand effects.

Clicks on links to real jobs: Our third measure captures whether jobseekers can target search according to job skill requirements based on limited information about jobs. 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 that aligned with the skill content of the majority of their choices in the job-choice task. The messages only contain the job title.[18] With SAYouth.mobi, we track whether jobseekers clicked on these links.

[17] We only classify jobs if they clearly mainly require one of the skills. Specifically, we can classify any job titles already coded by the recruitment staff who classify vacancies for the job choice task. We then also classify any jobs where the vacancy specifically mentions at least one of a list of synonyms for communication or numeracy skills and does not mention both skills. The second approach mimics the skill classification surveys used by O*NET, as in other work on skill classification on platforms (Deming & Kahn, 2018; Hershbein & Kahn, 2018).

[18] The message text was: Hi [firstname], we are contacting you in relation to the workshop you participated in at the [venue of workshop]. We found a job opportunity for a [job title] post on the SAyouth.mobi site that you might be interested in. The job is within traveling distance of [area around venue of workshop]. Please follow the link if you would like to apply.

We again find an increase in skills-based search direction. Treatment increases the difference in click rates between jobs which are aligned and misaligned with the jobseekers' comparative advantage by 7 percentage points (Table 3, column 7, $p = 0.25$). The effect is again driven by jobseekers with initially misaligned beliefs, who exhibit a treatment effect of 16pp (column 8, $p = 0.11$), with few effects on those with initially aligned beliefs. However, results are less precisely estimated because our estimation uses one observation per participant, compared to 11 observations per participant in the job choice task.

Planned applications for 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. Treatment increases the difference between the number of planned aligned applications vs misaligned applications by 1.42 applications, a 32 % increase on the control mean (Table 3, column 9, $p = 0.27$). Effects are driven by people with initially misaligned comparative advantage beliefs: the difference between their planned aligned and non-aligned applications increases by 4.8 applications (Table 3, column 10, $p = 0.011$). The observed change in search direction is a result of both a slight increase in planned aligned applications and a reduction in planned non-aligned applications. This measure is perhaps more sensitive to experimenter demand effects than our other measures but findings are consistent across measures.

Search direction index: We combine these measures into an index to avoid multiple hypothesis testing and to increase power (Anderson, 2008). We see a strong, positive, and significant treatment effect on the index of our four measures of directed search of 0.27 standard deviations (Table 3, column 1, $p = 0.013$). This effect is entirely driven by jobseekers with initially misaligned comparative advantage beliefs, for whom we observe a shift in search direction of 0.6 standard deviations (Table 3, column 1, $p < 0.001$). Jobseekers with initially aligned beliefs see few effects. The difference in treatment effects between jobseekers with initially aligned and misaligned beliefs is significant at the one percent level (-0.65 standard deviations, $p = 0.003$).

The size of the search direction effect for jobseekers with initially misaligned beliefs is similar to the correlation of aligned baseline beliefs with search direction in the control group (0.65 standard deviations, $p < 0.001$). The strong correlation in the control group further supports the notion that it is indeed the belief about the comparative advantage that drives measured search direction in this context.

3.5 Robustness and Additional Results

Beliefs About Returns to Directed Search: The model predicts that better information about one's relative rank in skills will also change beliefs about the relative returns to searching for jobs that require different skills. To evaluate this prediction, we measure jobseekers' beliefs about the outcomes from applying to communication- and numeracy-heavy jobs, both in the job choice task during the workshop and in their planned search after the workshop. We measure the believed return to skill-directed search as the difference in expected outcomes between CA-aligned and CA-nonaligned jobs, e.g., for a jobseeker with numeracy CA, the expected wage for numeracy-heavy jobs minus the expected wage for communication-heavy jobs, averaging over all job pairs.

Treatment increases the believed returns to most measures of skill-directed job search. Jobseekers' beliefs about job-specific expected wages also predict which jobs they choose in the job choice task. These results are consistent with the model, although not all are precisely estimated. See Appendix F for details of the measurement and results.

Additional Beliefs Results: For those particularly interested in processes of belief updating, we note some patterns relating to other literature. Treatment reduces the variance of beliefs about skills, constructed by measuring jobseekers' belief distribution over all the quintiles, consistent with treated jobseekers acquiring more precise information (Table A14 and Figure A2). Treatment updates underconfident beliefs more than overconfident beliefs (Table A15), matching asymmetric belief updating patterns in other research (e.g. Zimmermann 2020). Belief updating reflects beliefs about both absolute performance on the assessments and performance relative to the reference group, suggesting that jobseekers learn about both their own skills and their place in the skill distribution (Table A28).

Belief updating does not differ by jobseekers' cognitive skills, proxied by the concept formation score (Table A29), speaking to work on the role of cognitive skills in belief formation (e.g. Kruger & Dunning 1999 and Kahan 2013).

We find no evidence of differential belief updating by gender (Table A50). This, combined with the lack of gender heterogeneity in baseline skill beliefs, means that we pool both genders for the rest of the paper.

Robustness for Beliefs and Search Results: Hypothesis test results are robust to using a wild cluster bootstrap to account for the low number of clusters (34) and multiple hypothesis testing adjustment using sharpened q-values (Benjamini et al. (2006), Table A24). Treatment effect estimates are robust to using more continuous baseline measures of skill beliefs (Table A25) and to controlling for baseline 'confidence' – the difference between belief and measured skill quintiles – and its interaction with treatment (Table A31).

In Section 5, we explore heterogeneous treatment effects by the combination of baseline-aligned comparative advantage belief and over-/underconfidence in skill level, showing that this analysis does not substantially change the findings of the paper.

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

Having demonstrated that information about skills can shift 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 on average.

4.1 Sample

The big experiment took place in the same labor market in 2016/7, five years prior to the tight experiment. Recruitment is from the same population, active jobseekers on the database of our partner Harambee, a social enterprise providing job search assistance, described in Section 2.2.

Table A2 shows both samples have very similar search behavior: 37% (compared to 33% in the tight experiment) had done some work or income-generating activity in the past seven days and 97% (compared to 96%) were actively searching for work. The average jobseeker submitted 9 (compared to 10) job applications and spent 17 (compared to 14) hours searching for work in the last seven days.

There are some differences in demographic characteristics. Broadly, both populations face limited information about their skills in the labor market, with limited post-secondary education and formal work experience. Jobseekers in the big experiment are slightly younger (average 24) than in the tight experiment (26) and slightly more have likely to have completed secondary education relative to having dropped out of secondary education. Only 9% had ever held a permanent or long-term job (lower than the 23% in the tight experiment), perhaps linked to their younger age. Table A3 shows this mainly reflects changes in Johannesburg demographics, with a gradually aging population and declining levels of secondary completion. Table A24 shows the main treatment effect estimates are robust to reweighting the big experiment sample to have the same distribution of baseline age, education and gender as the tight experiment sample.

4.2 Experimental Design

Like the tight experiment, the big experiment also randomly varies whether jobseekers receive a report about their comparative advantage in skills. Treatment is administered at the day level with between 25 and 95 jobseekers taking part on a given day. We assign 2,114 jobseekers assessed over 27 days to the treatment group and 2,274 jobseekers assessed over a different 27 days to the control group. Treatment groups are balanced on covariates both for the full and the sample recontacted for the endline survey (Table A20).

We collect data on beliefs, search, and labor market outcomes at baseline and follow-up. The baseline is a self-administered questionnaire that candidates complete on desktop computers at Harambee under supervision. Endline data is collected through 25-minute phone surveys on average 3.5 months after treatment, reaching 96% of the respondents.[19] The response rate is balanced across treatment groups (Table A21).

The design of the big experiment is very similar to that of the tight experiment, which was designed to mimic it. In both, jobseekers complete a self-administered baseline and do skills assessments. The jobseekers' experience, including the ordering of belief measurement and assessment is shown in Figure A4. In both, treated jobseekers receive an unbranded report with information about their performance on the skills assessments; control jobseekers receive information about what skills are assessed but not their own scores. The control group are not told that jobseekers assessed on other days received assessment results. Therefore, it is unlikely that jobseekers assigned to the control group were discouraged by the lack of assessment results. In both, jobseekers received some job search assistance in a workshop, although its place in the day differs.

The experiments differ in three ways. First, we assess jobseekers on three "soft" skills in addition to communication, numeracy and concept formation: focus, grit, and planning, described in more detail in Appendix B. The reports given to treated jobseekers show results on all six assessments (Figure A5), rather than the two assessments in the tight experiment. This allows jobseekers to direct their search across a broader range of job types, taking advantage of the substantial heterogeneity in firms' skill demands in this context, documented by Carranza et al. (2022). The six assessments horizontally differentiate jobseekers based on their relative skills rather than only ranking or vertically differentiating them in a single dimension of skills, allowing skill-directed search in many different directions. Assessment results are weakly correlated across skills within candidates, with 12 of the 15 pairwise correlations below 0.2 (Table A6). As a result, most jobseekers learn they have substantial variation across skills: 85% have at least one top tercile but only 1.7% have six top terciles, 73% have at least one bottom tercile but only 0.5% have six bottom terciles, and 58% have both top and bottom terciles.

Second, we report assessment results in terciles and not quintiles to allow quick administration with larger groups. The coarseness of terciles relative to quintiles and using

[19] Garlick et al. (2020) provides an experimental validation of phone-based labor market surveys in this setting. Respondents receive mobile phone airtime payments for answering the phone surveys.

six rather than two skills means that only 23% of the jobseekers have a unique comparative advantage (CA). The tight and big experiments thus deliberately capture contrasting approaches: the tight experiment uses fewer skills to allow cleaner definitions of CA and clearer tests for skill-directed search, while the big experiment uses more skills to more closely approximate real world settings where jobseekers could get information about skills, like education or job centres, which tend to assess a larger number of skills.

Third, there are some logistical differences that are unlikely to affect the core mechanism activated by the treatment. In the big experiment, candidates are assessed in bigger groups (80, not 10-15). The briefing after candidates receive the report is delivered in person instead of by video. Jobseekers receive a CV template, interview tips, and job search tips at the beginning of the day in the big experiment and the end of the day in the tight experiment.

4.3 Specification

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

$$Y_{id} = T_d \cdot \beta + \mathbf{X}_{id} \cdot \Gamma + \varepsilon_i \tag{6}$$

where Y_{id} is the outcome of interest for jobseeker i assessed on date d and $\mathbf{X}_{id}$ is a vector of prespecified control variables.[21] The object of interest is the average treatment effect β. We cluster standard errors at the level of treatment (assessment days). As in the tight experiment, we combine families of outcomes in Anderson (2008) indices to limit the number of tested hypotheses. We show in Table A24 that our main hypothesis test results are robust to using a wild cluster bootstrap and to adjusting for multiple hypothesis testing using sharpened q-values, following Benjamini et al. (2006).

We do not emphasize heterogeneous treatment effects by baseline aligned comparative advantage beliefs as much as we did in the tight experiment. In the big experiment, with six skills, there are many more ways for assessed and believed comparative advantage to differ. Furthermore, we only observe beliefs about three of the six skills in the big experiment. Hence, not having baseline aligned CA beliefs is a noisier proxy for having beliefs that can be shifted by treatment in the big than tight experiment.

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

[21] $\mathbf{X}_{id}$ contains measured skills, self-reported skills, education, age, gender, employment, above median discount rate, above-median risk aversion, baseline values for the outcome where available, and fixed effects for the blocks of days within which treatment was randomized.

4.4 Better Information About Skills Aligns Beliefs with Comparative Advantage

Receiving information about skills substantially increases alignment between measured skills and skill beliefs at the time of the endline survey. This is in line with the effects on beliefs observed straight after treatment in the tight experiment.

In the endline survey of the big experiment, we ask candidates in which tercile they believe they ranked for each of the communication, concept formation, and numeracy assessments (see Appendix C.1 for measurement details). We define jobseekers as having aligned CA beliefs if their believed top skill(s) matches their assessed top skill(s) 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 CA beliefs on this measure. Treatment increases this by 13.9pp, a 72% increase (Table 4, column 1, $p < 0.001$).[22] This result shows that receiving information about their skills enables jobseekers to learn about their comparative advantage across, even when substantially more information is provided (on six skills rather than three). Furthermore, their updated beliefs have persisted for 3.5 months after receiving the information. 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 (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 A51).

4.5 Job Search with Better Aligned Beliefs about Skills

We collect a self-reported measure of whether jobseekers are directing their search according to skill CA. We ask jobseekers "Think about the types of jobs that you are applying for. Think about what skills are important for these jobs. Rank the following skills from most important to least important." The skill options are numeracy/maths skills, literacy/communications skills, problem-solving skills, and soft skills. We define aligned search as a dummy indicating that jobseekers' highest ranked skill on this measure (ignoring soft skills) matches the skill in which they have an assessed CA: the most important skill listed for jobs that they are applying for matches the skill in which they have an assessed CA.[23] We construct this measure from separate questions about jobseekers' skill

[22]The treatment effects on aligned CA beliefs are almost identical across the tight and big experiments. However, the control group mean is substantially higher in the tight than big experiment because the big experiment uses three skills, creating more ways for believed and assessed CA to differ. In both experiments, the control group's alignment rate is only slightly higher than random guessing, accounting for the possibility of ties: 19 versus 15% in the big experiment and 48 versus 40% in the tight experiment.

[23]For individuals without a clear CA (i.e., individuals with a tie for highest skill) our main aligned search measure is set to zero. However, all our results are robust to recoding this measure to one when any of the highest skills is listed first.

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

	Beliefs		Search direction
	(1) Aligned comp. adv. belief	(2) Fraction aligned beliefs	(3) Aligned search
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 providing jobseekers with information about their comparative advantage in skills aligns skill beliefs with assessed skills and shifts self-reported search direction in the big experiment.** Column 1 shows the impact on aligned comparative advantage beliefs. Column 2 shows the impact on the fraction of aligned skill tercile beliefs across three domains (numeracy, communication, and concept formation). Column 3 shows the impact on a dummy indicating search direction aligned with comparative advantage in skills (self-reported). Controls include randomization block fixed effects, and prespecified baseline covariates (measured skills, self-reported skills, education, age, gender, employment, discount rate, and risk aversion). Standard errors clustered at the treatment-day level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

beliefs and the skill demand of jobs to which they apply, rather than asking them directly if they are searching for jobs that match their CA to avoid experimenter demand effects.

Treatment increases the share of jobseekers searching mainly for jobs that match the skill in which they score highest from 17% to 22% three months after treatment (Table 4, column 3, $p < 0.001$). This self-reported search direction result is qualitatively consistent with the four measures of directed search in the tight experiment, collected within the first month after treatment.[24]

4.6 Improved Labor Market Outcomes

Taken together, the results on beliefs and search from the two experiments provide strong support for our key argument: information about skills improves the alignment of jobseekers' job search strategy with their measured skill CA. This is in line with the predic-

[24] The effect in the big experiment converts to a 0.14 standard deviation increase in directed search, compared to a 0.27 standard deviation increase in the directed search index in the tight experiment (Table 3, column 1). The latter effect is likely to be larger because the index averages over four measures, producing a smaller standard deviation and hence larger standardized effect size.

tions of the model that a shift in CA beliefs should change the search direction of jobseekers toward jobs that require the skills in which they perform well. The model predicts that the change in search direction should then affect jobseekers' labor market outcomes as they direct their search toward jobs with higher returns to search. Here we show evidence consistent with this prediction and then evaluate possible mechanisms that link skill-directed search to 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 month after the workshop, second month after the workshop, and week before the endline survey are positive but mostly small and mostly not statistically significant (Table 5, panel A, columns 2-5). Our prespecified Anderson (2008) index of work quantity increases by 0.05 standard deviations with $p = 0.13$ (column 1). This result is sensitive to reweighting: the effect is marginally significant when we reweight the big experiment sample to match the demographic composition of the tight experiment (Table A24).

Treatment has substantial effects on employment quality: Treatment increases earnings in the seven days before the endline by 6.52 USD, 26% of the control group mean (Table 5, panel B, column 2, $p = 0.019$).

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. Interpreting treatment effects on such measures can be complex (Mullahy & Norton, 2022). Hence we show that treatment effects on earnings are robust across multiple different approaches. First, effects are weakly positive on the entire earnings distribution (Figure A6). 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 transformations (Table A26, columns 1 and 3-5). Third, restricting the sample to the employed produces only slightly larger treatment effects (Table A26, columns 6-10). This third approach can create a sample selection problem when treatment affects the probability of employment and hence the probability of having non-missing earnings. But, in this case, treatment has little impact on employment, so it is unsurprising that the results are robust to different ways of handling the earnings of the nonemployed. The earnings effects are at the upper end but still in line with effect sizes in related research.[25]

[25] For example, interventions that combine learning about skills with the ability to signal skills increase earnings by 11-34% of the control group means (Abebe et al., 2021a; Bassi & Nansamba, 2022; Carranza et al., 2022). Bandiera et al. (2021) find an 11% change in earnings from a matching intervention that they attribute to changes in 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 wages by 11 percent.

We show in Appendix G that the earnings effect implies that the average treated jobseeker earns enough extra in the time between treatment and endline to cover 1.8 times the average variable cost of the assessment operation.

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, a marker of job formality in this context, by 1.7 percentage points from a 12% mean in the control group ($p = 0.095$, Table 5, panel B, column 4). Our index of work quality increases by 0.085 standard deviations ($p = 0.019$, Table 5, panel B, column 1).

Heterogeneous treatment effects: As our model suggests, these treatment effects are driven by jobseekers who get new information from treatment about the comparative advantage. Table A33 shows that the treatment effect on every labor market outcome is larger for jobseekers with misaligned than aligned CA beliefs at baseline. 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 misaligned and aligned beliefs respectively. Similarly, the treatment effect on skill-directed job search is driven by jobseekers with misaligned baseline CA beliefs, although the pattern of heterogeneity is less clear for updating the CA beliefs themselves (Table A30).

Linking skill-directed search to improved labor market outcomes: What specific economic mechanisms link skill-directed search to 'better' jobs – in terms of earnings, hourly wages, and formality – but not to a higher employment probability? Our data do not allow a conclusive answer to this question. But we provide a framework for thinking about these mechanisms and empirical tests of parts of this framework.

We begin by noting three patterns that motivate our focus on search-based mechanisms. In each case, we estimate treatment effects on interactions between multiple labor market outcomes, e.g., the treatment effects on earnings interacted with indicators for different start dates of current jobs. First, the majority of the positive effects on earnings and written contracts are in jobs with start dates after treatment (Table A27, columns 1-2 and 6-7). This suggests that changes to search behavior are more important than wage bargaining with existing employers. Second, the entire earnings effect comes from wage employment, not self employment (columns 4-5). This suggests that changes to job application behavior and application outcomes are an important mechanism. Third, search and employment are uncorrelated in the control group and treatment has effects of less than 1pp on all four possible search $\times$ work states. This suggests that the distinction between on-the-job and off-the-job search actions is not an important consideration, although the outcomes of the on- and off-the-job search may obviously differ.

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

	(1)	(2)	(3)	(4)	(5)
Panel A: Work quantity					
	Index	Job offers (w)	Worked month 1	Worked month 2	Worked last 7 days
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	
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 work quality but not necessarily work quantity in the big experiment.** Panel A shows impacts on employment quantity. Panel A, column 1 shows impacts on an Anderson (2008) index of the four employment quantity measures. Panel A, column 2 shows the impact on the number of job offers in the last 30 days (winsorized). Panel A, column 3 shows impacts on a dummy indicating any work for pay in month 1 after treatment. Panel A, column 4 shows impacts on a dummy indicating any work for pay in month 2 after treatment. Panel A, column 5 shows the impact on a dummy indicating any work for pay in the last seven days. Panel B show impacts on employment quality. Panel B, column 1 shows impacts on an Anderson (2008) index of the three employment quality measures. Panel B, column 2 shows the impact on earnings in the last seven days (winsorized). Panel B, column 3 shows the impact on hourly wages in the last seven days (winsorized). Panel B, column 4 shows the impact on a dummy indicating a written contract. Controls include randomization block fixed effects, and prespecified baseline covariates (measured skills, self-reported skills, education, age, gender, employment, discount rate, and risk aversion). 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Given these patterns, we evaluate two mechanisms centered on two standard concepts in the search literature: the **wage offer distribution** and **job offer probability**. Returning to our model from Section 2.1, we assume that there are multiple types of jobs 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 allocation of search effort across different types of jobs generates a wage offer distribution $F(w)$ and job offer probability P that is a weighted average across the different job types. As in standard job models, any jobseeker who receives an offer accepts it if the wage is above their reservation wage. Treatment, by shifting search toward job types that use jobseekers' comparative advantage, can shift F to the right and/or raise P.

In the first potential mechanism, skill-directed job search shifts at least part of $F(w)$ to the right, increasing average earnings conditional on employment. If the entire wage offer distribution shifted to the right, then more jobseekers would receive offers above their reservation wage, leading to higher employment. To match our null effect on employment, this mechanism requires an additional feature: either an increase in the reservation wage, so some of the higher wage offers are rejected; or that the shift in F occurs only above reservation wages. We do not observe the terms of job offers, so we cannot directly test that part of the mechanism. But the treatment effect on the CDF of earnings suggests a broad rightward shift in F (Figure A6). Treatment increases reservation wages in both experiments but the effects are under 2% of the control group mean and not statistically significant. We conclude that the data don't offer strong evidence for or against this mechanism.

In the second potential mechanism, skill-directed search raises P, leading to more job offers. To match the positive effect on earnings, some jobseekers must receive multiple offers, so the additional offers allow them to select higher-wage offers and hence earn more. To match the null effect on employment, treatment must also increase reservation wages, so most of the additional offers are rejected. We see some evidence against this explanation: treatment effects on earnings and hourly wages are substantially larger than the near-zero effects on reservation wages, the treatment effect on job offers is close to zero, and only 4% of jobseekers in both the control and treatment groups received more than one job offer in the month before endline. We conclude that the data provide more evidence against than for this mechanism.

If treatment does lead to higher wage offers, as in the first economic mechanism, why might it do so? The model suggests that jobseekers apply to jobs that better match their skills where they can be more productive. We can't directly test this in the big experiment, because we don't observe detailed data on occupation or job tasks. But we do find that treatment increases tenure by 0.44 months, a substantial increase given the control group mean of 1.9 months and the average 3.5 month gap between treatment and endline. This higher tenure provides suggestive but imprecise evidence for higher productivity due to better skill match.

Taken together, the treatment effects across the two experiments align with the model prediction that jobseekers' beliefs about their comparative advantages matter for search direction and, ultimately, labor market outcomes. The tight experiment provides stronger evidence for the link between comparative advantage beliefs and search direction, while the big experiment shows the labor market implications of this process.

5 Search Effort and Beliefs About Skill Levels

In principle, new information about jobseekers' skills might change their search effort as well as search direction, contributing to changes in labor market outcomes. We find very limited treatment effects on search effort, suggesting that search effort does not explain the relationship between treatment and labor market outcomes.[26] We very briefly summarize this evidence here and present the full details in Appendix H, including tables and a generalized conceptual framework allowing endogenous search effort choices.

First, treatment lowers jobseekers' believed skill levels in both experiments (Table A39, 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).

Second, a generalized version of our conceptual framework predicts that this change in beliefs has a theoretically ambiguous effect on search effort. In this framework, jobseekers endogenously chose 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 treatment effect on beliefs: a substitution effect – jobseekers search less because the expected return to each unit of search effort is lower – and 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.

Third, treatment effects on search effort are negligible in both experiments. In the tight experiment, treatment effects are close to zero, not statistically significant, and mostly negative on six different measures of search effort: a survey question on post-workshop planned applications, time spent on a job search task during the workshop, and four measures of job search on the SAYouth.mobi platform after the workshop (Tables A40 and A41). In the big experiment, treatment effects are precisely estimated near-zeros on

[26]This does not, of course, imply that search effort plays no role in determining labor market outcomes in this or other settings. It merely suggests that shifts in search effort are less likely to explain the treatment effects on labor market outcomes than shifts in search direction to align with skill comparative advantage. See the review by Mueller & Spinnewijn (2022) for other evidence on the relationship between labor market outcomes, search effort, and beliefs about labor market prospects, such as job offer probabilities.

three survey measures of search effort: self-reported number of applications and time and money spent on search (Table A42).

6 Additional Mechanisms

In this section, we evaluate four other mechanisms that might account for treatment effects on labor market outcomes. These are not in principle mutually exclusive with the directed search mechanism. We find little evidence for any of these four mechanisms.[27]

6.1 Self-Esteem

If jobseekers' self-esteem – general beliefs about self-worth – respond to new skill information, treatment might affect search behaviour and labor market outcomes. However, treatment has near-zero effects on self-esteem in the big experiment (Table A45). We find this result both in a text message survey 2-3 days after treatment (columns 1-2) and in the endline phone survey roughly 3.5 months after treatment (columns 3-4). The endline survey uses five questions from the Rosenberg (1965) scale and the text message survey uses a single question from this scale, all with five-point Likert response options.

6.2 Human Capital Investment

Treatment might affect labor market outcomes if jobseekers respond to new skill information by investing in additional or different skills. We test this in both experiments. In the big experiment, treatment has near-zero effects on enrollment in both formal and vocational education in the time between treatment and endline (Table A45, columns 5-7). So the treatment effects on earnings in the big experiment are unlikely to be driven by changes in education investment.

owever, some results in the tight experiment suggest that jobseekers might be willing to pay to invest in skills under other conditions. In the tight experiment, we measure jobseekers' willingness-to-pay (WTP) for numeracy and communication training materials (see Appendix I). Treatment reduces 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. But credit constraints or other factors may prevent jobseekers making actual investments in education and training.

[27] We prespecified testing for the "human capital investment" explanation in both experiments and for the "self-esteem" and "information transmission to firms" explanations in the big experiment.

6.3 Skill Information Transmission to Firms

Treatment might affect labor market outcomes if jobseekers share assessment results with firms during job applications, leading to firm-side learning about jobseekers' skills. Some jobseekers in the big experiment do share assessment results with firms: 29% of treated jobseekers self-report that they included a copy of their assessment result with at least one job application, although only 0.8% of treated jobseekers actually included their assessment result in applications to vacancies we created for a companion study.

However, sharing assessment results with firms is unlikely to explain the treatment effects on labor market outcome, for two reasons. First, the reports showing assessment results 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. None of the 15 hiring managers we interviewed during piloting said they would view these reports as credible. Second, the labor market effects are driven by the jobseekers who say they did not use the report in applications (Table A46). This suggests that using reports with applications does not affect labor market outcomes, although we interpret this result cautiously because this analysis conditions on report use, a post-treatment outcome.[28]

6.4 Feedback Loop from Labor Market Experiences to Skill Beliefs

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 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 2.5. 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.

[28] See Carranza et al. (2022) for the labor market effects of providing a different type of skill assessment result that is deliberately designed to be credible to firms.

7 Conclusion

We provide evidence of a relatively understudied type of job search friction: misdirected search due to limited information about the comparative advantage in skills. We show that skill-directed job search can be constrained by jobseekers' limited information about their own comparative advantage in skills. Giving jobseekers additional information about their 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 the 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 market prospects implicitly condition on beliefs about their skills and their search strategies given these skills, so directly capturing these can better understand how beliefs and search behavior co-evolve. 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 types of experimental designs with multiple types of measurement. This allows us to precisely observe shifts in beliefs and search direction in a simpler, two-skill setting with more precise measurement, but also study effects on labour market outcomes in a larger sample with a longer follow-up period.

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? The average treated participant's earnings gain likely exceeds the average variable cost of running the assessment system, suggesting the possibility of profitable market provision. Assessment allows substantial economies of scale, particularly if implemented by online job search and matching platforms, rather than the in-person assessment service we study. 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 in at least the general skills we study in this project, 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, which sometimes include skill assessments, could fill this gap. Researchers have not yet studied the role of this specific component of job search counseling. Future work might examine the economics information provision by both public- and private-sector actors about comparative advantage across different types of skills.

References

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

Abebe, G., A. S. Caria & E. Ortiz-Ospina (2021b) "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 & 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.

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*.

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. & 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., E. Aucejo, A. Maurel & T. Ransom (2016) "College Attrition and the Dynamics of Information Revelation," *NBER Working Paper 22325*.

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

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 (2021) "The Search for Good Jobs: Evidence from a Six-Year Field Experiment in Uganda," *SSRN Working Paper 3910330*.

Banerjee, A. V. & S. Sequeira (2020) "Spatial Mismatches and Imperfect Information in the Job Search," *CEPR Discussion Paper 14414*.

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 International Migration," *NBER Working Paper 29588*.

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.

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.

Chen, J. & J. Roth (2022) "Log-like? Identified ATEs Defined with Zero-Valued Outcomes Are (Arbitrarily) Scale-Dependent," *Working Paper*.

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 (2021) "Gender Differences in Job Search and the Earnings Gap: Evidence from Business Majors," *NBER Working Paper 28820*.

Delavande, A. (2022) "Expectations in Development Economics," in Van der Klaauw, W., G. Topa & R. Bachman eds. *Handbook of Economic Expectations*, 261–291: Elsevier.

Deming, D. & L. Kahn (2018) "Skill Requirements across Firms and Labor Markets: Evidence from Job Postings for Professionals," *Journal of Labor Economics*, 36 (S1), S337–S369.

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

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*.

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*.

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.

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.

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

He, H., D. Neumark & Q. Weng (2021) ""I Still Haven't Found What I'm Looking For": Evidence of Directed Search from a Field Experiment," *NBER Working Paper 28660*.

Heath, R., L. Boudreau & T. McCormick (2022) "Migrants, Experience, and Working Conditions in Bangladeshi Garment Factories," *SSRN Working Paper 4067206*.

Hershbein, B. & L. Kahn (2018) "Do Recessions Accelerate Routine-Biased Technological Change? Evidence from Vacancy Postings," *American Economic Review*, 108 (7), 1737–1772.

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

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

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

Kahan, D. M. (2013) "Ideology, Motivated Reasoning, and Cognitive Reflection," *Judgment and Decision Making*, 8 (4), 407–424.

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.

Kruger, J. & D. Dunning (1999) "Unskilled and Unaware of It: How Difficulties in Recognizing One's own Incompetence Lead to Inflated Self-Assessments.," *Journal of Personality and Social Psychology*, 77 (6), 1121.

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/a5

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.
Mahoney, N. (2022) "Principles for Combining Descriptive and Model-Based Analysis in Applied Microeconomics Research," *Journal of Economic Perspectives*, 36 (3), 211–222.
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.
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*.
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.
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 (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 (2020) "What's the Inside Scoop? Challenges in the Supply and Demand for Information about Job Attributes," *Working Paper*.
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 (2022) *Quarterly Labor Force Survey Quarter 3 2022*, Pretoria: Statis-

tics South Africa.

Stinebrickner, R. & T. Stinebrickner (2014a) "Academic Performance and College Dropout: Using Longitudinal Expectations Data to Estimate a Learning Model," *Journal of Labor Economics*, 32, 601–644.

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

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

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

Taylor, S., S. van der Berg, V. Reddy & D. J. van Rensburg (2011) "How Well Do South African Schools Convert Grade 8 Achievement into Matric Outcomes?" *Stellenbosch University, Department of Economics Working Papers 13/2011*.

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.

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 details and results of the willingness-to-pay elicitation in the tight experiment. Appendix J contains exhibits supporting the additional mechanism analyses from Section 6. Appendix K shows the results related to gender heterogeneity. Finally, Appendix L describes how our analysis relates to the pre-analysis plans.

Contents

A Sample description

This appendix contains two additional exhibits that describe the two samples respectively. Table A1 shows that jobseekers are active on the platform 30 days prior to treatment in the tight experiment. Table A2 confirms that jobseekers in the big experiment are also actively searching and are similar to the jobseekers of the tight experiment shown in Table 1. Based on Table A3 both experimental samples broadly resemble the population of Johannesburg, except that jobseekers in the sample are more likely to be unemployed and searching for jobs.

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

	mean	sd	median	min	max	N
# days active on platform	3.17	4.06	2.00	0.00	25.00	278
# applications clicks (w)	6.51	11.21	2.00	0.00	65.00	278
# applications clicks for numeracy heavy jobs	0.39	1.10	0.00	0.00	11.00	278
# applications clicks for communication heavy jobs	0.82	1.64	0.00	0.00	9.00	278
Fraction of skill coded application clicks	0.12	0.20	0.00	0.00	1.00	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".

Table A2: Summary Statistics - Big Experiment

	(1) Mean	(2) Median	(3) SD	(4) Min	(5) Max	(6) Obs.
Panel A: Demographics						
Black African	0.92	1.00	0.27	0.00	1.00	4389
Male	0.38	0.00	0.48	0.00	1.00	4389
Age	23.67	23.14	3.28	18.04	35.08	4389
Completed secondary education only	0.61	1.00	0.49	0.00	1.00	4389
Any other post-secondary qualification	0.22	0.00	0.41	0.00	1.00	4389
University degree / diploma	0.17	0.00	0.37	0.00	1.00	4389
Panel B: Labor market background						
Worked in past 7 days	0.37	0.00	0.48	0.00	1.00	4389
Earnings in USD (last 7 days, winsorized)	31.26	0.00	75.72	0.00	476.00	4389
Has worked in permanent wage job before	0.09	0.00	0.29	0.00	1.00	4377
Panel C: Search behavior						
Did any search in past 7 days	0.97	1.00	0.17	0.00	1.00	4389
# applications (last 30 days, winsorized)	9.34	5.00	12.85	0.00	90.00	4346
Search expenditure in USD (last 7 days, winsorized)	30.97	20.40	33.05	0.00	204.00	3995
# job offers (last 30 days, winsorized)	0.80	0.00	2.66	0.00	20.00	4335
Panel D: Skills beliefs						
Aligned belief about comparative advantage	0.20	0.00	0.40	0.00	1.00	4312
Fraction of aligned skill beliefs	0.38	0.33	0.31	0.00	1.00	4378
beliefs and assessments						

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

Table A3: Summary Statistics for Experimental and External Comparison Samples - both experiment

	QLFS Johannesburg (2016-17)		Sample (Big experiment)	QLFS Johannesburg (2020)		Sample (Tight experiment)
	Age-restricted	Reweighted		Age-restricted	Reweighted	
Age	26.5	23.7	23.7	26.6	26.6	26.8
	(4.7)	(3.3)	(3.3)	(4.7)	(3.9)	(4.1)
Gender	0.500	0.377	0.378	0.527	0.323	0.285
Black	0.824	0.985	0.985	0.868	0.989	0.992
Highest Education Level						
Less than Secondary	0.388	0.009	0.009	0.376	0.054	0.094
Completed Secondary	0.432	0.606	0.605	0.495	0.645	0.677
More than Secondary	0.163	0.385	0.386	0.116	0.299	0.339
Employed	0.445	0.375	0.374	0.357	0.402	0.333
Searching	0.536	0.535	0.971	0.426	0.491	0.949
Earnings	151	121	33	84	90	53
	(537)	(391)	(91)	(474)	(306)	(146)
N			4389			372

Notes: Table A3 compares the sample of jobseekers in the big experiment (column 3) to several external benchmarks: people in Johannesburg in the eligible age range for the study (column 1), and people in Johannesburg in the eligible age range for the study, reweighted with propensity scores to approximate the experimental sample on age, education, sex, and race (column 2). External benchmarks are calculated from the Quarterly Labour Force Survey (QLFS), averaging over all 2016 and 2017 waves and using post-stratification weights provided by Statistics South Africa. Age is calculated from ID numbers and as of baseline and workshop completion dates. Columns 5-7 repeat the same exercise for the tight experiment and uses the QLFS in 2020. Standard deviations are shown in parentheses for all continuous variables. 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.

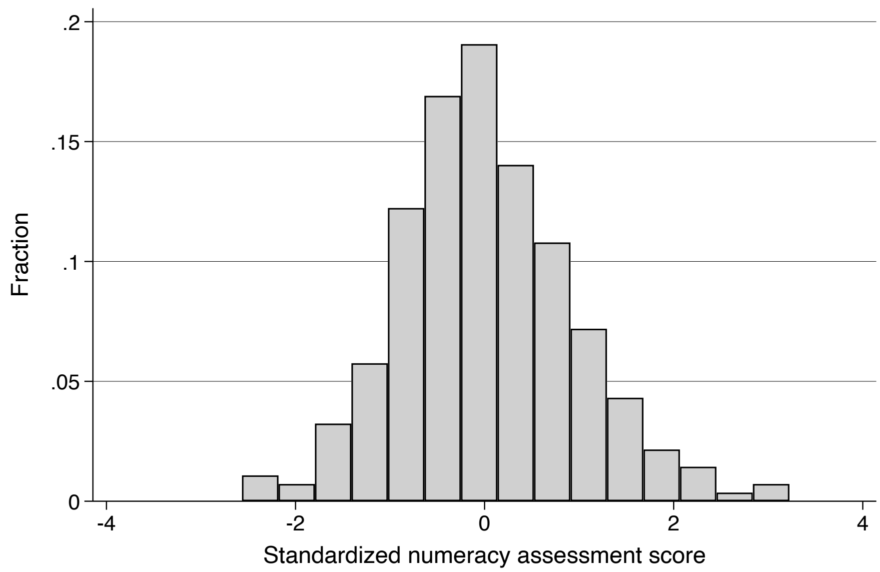
The *communication* assessment quantifies jobseekers' listening and reading comprehension. The assessment focuses on high school English proficiency and was developed by a local adult education provider (www.mediaworks.co.za).

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.

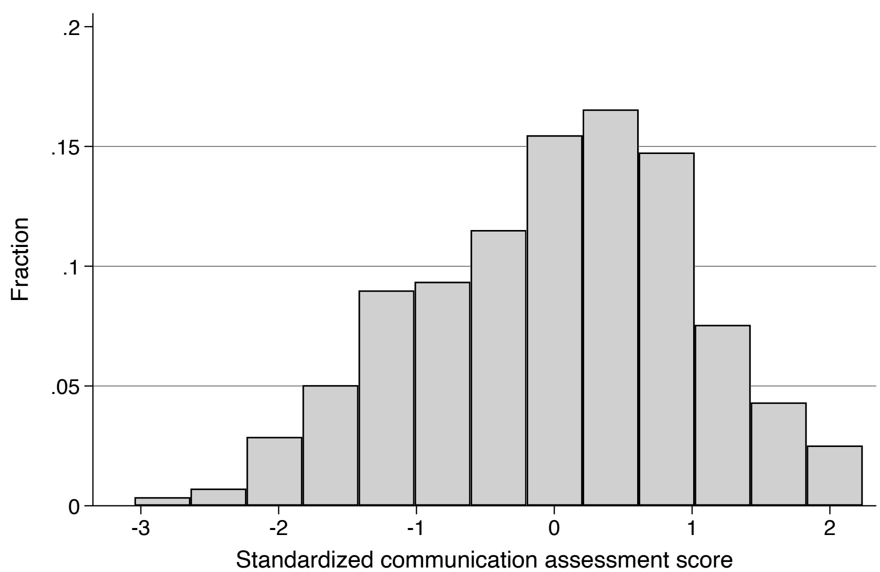
We find ample variation within skills and across skill quintiles. Figure A1 shows the distribution of measured numeracy and communication scores after standardization in the tight experiment. Table A4 show the joint distribution of numeracy skill quintiles and communication skill quintiles. Table A5 shows the correlations between skill measures in the tight experiment. We find moderate positive correlations between all three skill measures.

Figure A1: Distribution of Measured Skills - 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 numeracy scores. Panel B displays communication scores.

Table A4: 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 A4 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.

Table A5: Correlation Between Skill Quintiles - Tight Experiment

	Numeracy	Communication	Concept formation
	(1)	(2)	(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 A5 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. Panel B shows results for the full sample.

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.

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 candidates' 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 A6 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.

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

Table A6: Correlation Between Skill Terciles - Big Experiment

	Concept formation	Communication	Numeracy	Grit	Planning	Focus
	(1)	(2)	(3)	(4)	(5)	(6)
Concept formation	1.000	0.298	0.435	0.156	0.120	0.140
Communication		1.000	0.331	0.111	0.094	0.119
Numeracy			1.000	0.172	0.124	0.116
Grit				1.000	0.157	0.168
Planning					1.000	0.132
Focus						1.000

Notes: **Table A6 shows positive but moderate Pearson correlation coefficients between the skill terciles used in the big experiment.**

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. 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.

To summarize the 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. They 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.

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 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 A7 shows the ranking of numeracy, communication and education, averaging over the 67 firms. There are six different possible rankings of these three elements, each shown in a row. Column 6 shows that 57% of firms prefer a candidate with top-tercile numeracy skills, 34% prefer a candidate with top-tercile communication skills and 9% firms prefer a candidate with a relatively better educational achievement (middle-tercile communication and numeracy skills and with a diploma).

Table A7: 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	Com	Educ	52.24%	Numeracy	56.72%
2	Num	Educ	Com	4.48%		
3	Com	Num	Educ	28.36%	Communication	34.33%
4	Com	Educ	Num	5.97%		
5	Educ	Num	Com	1.49%	Educaction	8.95%
6	Educ	Com	Num	7.46%		
				100.00%		100.00%

Notes: **Table A7 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. Column 5 and 6 does the same collapsing 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 members 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 A8, panel A, column 6 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.

Table A8: Employer Evaluation of Job Applicants Based on Skills

	Mean			Difference			Obs.
	(1) Aligned	(2) Non-aligned	(3) SD	(4) Δ	(5) Δ/SD	(6) p($\Delta = 0$)	(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 A8 shows that an employer can observe jobseekers' skills and evaluates applicants more highly if their assessed comparative advantage in skills matches the job's requirements.** It compares employers' evaluation of jobseekers for two jobs: one that aligns with their assessed comparative advantage in skills and one that does not. As part of the experiment, we ask jobseekers to choose between applying for a real communication or numeracy job at a firm that hires for a range of entry-level roles. They prepare a resume and a cover letter during the workshop, both designed for general use rather than tailored to these specific jobs. Two members of the firm's HR team then evaluate every applicant for both jobs. Evaluators are blind to which applicant applied for which job and are not shown applicants' skill assessment results. We average the scores across both evaluators for each jobseeker. Panel A, column 1 displays the employer's valuation of a jobseeker's skill that is aligned with their assessed comparative advantage (on a scale from 1 to 5, where higher numbers indicate higher skills). Panel A, column 2 displays the employer's evaluation for the misaligned skill on the same scale. Panel B displays the employer's rating of employees for the job whose skill requirement is aligned (column 1) and misaligned (column 2) with the jobseekers' assessed comparative advantage in skills. Evaluators assess applicants' general suitability for the job (on a scale from 1 to 5) and whether they would invite the candidate for an interview. Column 3 shows the pooled standard deviation of the measures in columns 1 and 2. Column 4 shows the difference between columns 1 and 2. Column 5 shows the same difference in terms of standard deviations. Column 6 shows the p-value associated with a test of equality across columns 1 and 2. Column 7 shows the number of observations.

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 A9 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.

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 A3 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 A10, A15, A22, A28, A29 (column 1), A31, A39 (columns 1 to 4), and A50). After treatment, we further measure the distribution of general skill beliefs and use it to show the impact on belief variances in Table A14 and Figure A2.

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 comparative advantage beliefs defined using general skills. We find that treatment increases aligned comparative advantage 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 as part of this study in both experiments. We consider these beliefs a *proxy for the information content of the interventions* for each individual because the treatment provides information about assessment results (following Haaland et al., 2023). Jobseekers who have inaccurate beliefs about their 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 5). 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 measures of comparative 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 A3). We use these 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 20. 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 A4.) 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.

Table A9: 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 (prob. distribution over quintiles)	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. Here are five boxes showing the five groups of people ranked according to their [numeracy/communication] skills. Please place buttons on each of the boxes to show what you think the chances are that you are in each group. For example, if you place 10 buttons in Group 3, you are completely sure you are in this group. But if you are not completely sure, you would place fewer than 10 buttons.
General skills, post-treatment (most likely quintile)	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 A9 displays the exact wording of our skill belief measures.** 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."

C.2 Skill Belief Descriptive Statistics and Treatment Effects

In this section we provide descriptive statistics about jobseekers' skill beliefs. We then show addition treatment effect exhibits and correlational evidence that treatment shifts skill beliefs and skill beliefs correlate with believed retuns to search and search direction.

Table A10 displays correlations between pre-treatment beliefs about skills and assessment results in the tight experiment. We find that numeracy beliefs are strongly correlated with assessment results but communication skills are not significantly correlated with assessment results. Table A11 displays correlations between self-reported grades in English and math in the secondary school leaving exam (matric) and assessment results and beliefs. Columns 1-3 indicate that the secondary school leaving exam results correlate with our assessments and the score differences on those exams positively correlate with the measured comparative advantage in the tight experiment.

We also find that matric results predict jobseekers' beliefs (Table A11, Columns 4-7). However, we do not find that any baseline demographic varaibles, including labor market exposure, meaningfully predict having aligned CA beliefs at baseline in the tight experiment (Table A12). Table A13 shows the development of skill beliefs over time in the control group of the big experiment. Here we observe that beleifs do not get updated over time in the control group. Overall, these results – matric predicts beliefs but not other demographics, and beliefs do not get updated over time in the control group – suggests that jobseekers have limited capacity to learn about their comparative advantage in this labor market.

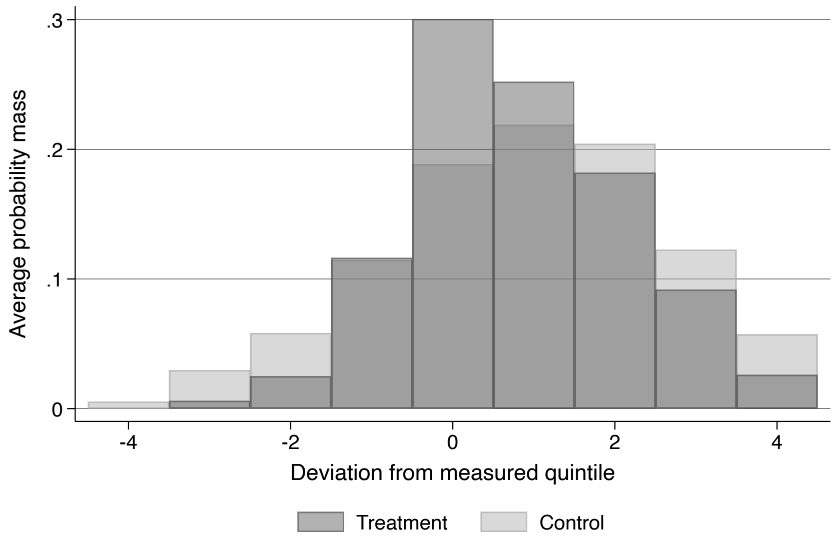
We elicit respondents' full discrete distribution of beliefs about skill quintiles in the tight experiment. Figure A2 shows that the distribution of skill beliefs is closer to assessed skills for the treatment than the control group. Table A14 shows that treated participants also become somewhat more certain in their skill beliefs.[29] Average variance in beliefs goes down by 11%, with similar decreases across numeracy and communication beliefs, although this is noisily estimated.

Table A15 displays treatment effects on dummies indicating the fraction of aligned, under, and overconfident skill beliefs. Table A16 shows the correlation between jobseekers' beliefs about skills and beliefs about the returns to search using the control group in the big experiment. Table A17 shows the correlation between jobseekers' beliefs about comparative advantage and search direction in the control group of the big experiment.

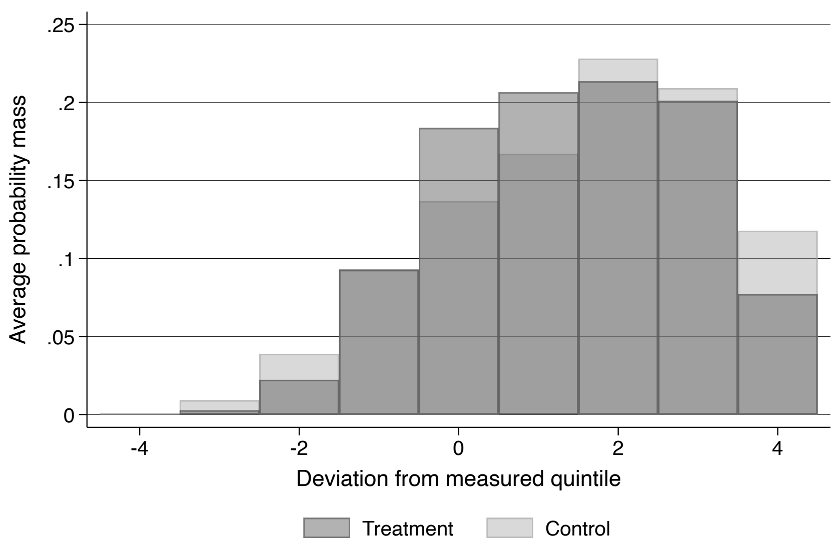
[29] Specifically, we ask respondents to allocate ten tokens across the five quintiles to represent the likelihood of them falling in that group (see specific wording in Table A9). We use this data to calculate the expected skill quintile ($\tilde{s}$) and calculate the variance in beliefs as $var(s_l) = \sum_{i=0}^{4} p_i(i-\tilde{s})^2$. We then average the variance across the two skill domains.

Figure A2: Treatment Effects on Distribution of Beliefs About Skills - Tight Experiment

Panel A: Numeracy beliefs



Panel B: Communication beliefs



Notes: **Figure A2 shows how the treatment shifted beliefs about skills toward assessed skills in the tight experiment.** Beliefs are elicited by asking respondents to place 10 buttons across the five possible quintiles. x-axis measures the deviation from the assessed quintile. y-axis measures the average probability mass. Panel A displays the distribution of beliefs for numeracy beliefs. Panel B displays the distribution of communication beliefs.

Table A10: Association Between Assessed and Believed Skill Quintiles - Tight Experiment

	Skill quintile beliefs	
	(1) Numeracy	(2) Communication
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 A10 analyzes the correlations between assessed skill quintiles and baseline skill quintile beliefs in the tight experiment.** No control variables are included. Robust standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A11: Association of High School Graduation Exam Results with Assessed and Believed Skills - Tight Experiment

	Assessed skills			Beliefs about skills			
	Skill quintile		Comp. adv.	Skill quintile		Comp. adv.	
	(1) Num.	(2) Com.	(3) Num.	(4) Num.	(5) Com.	(6) Num.	(7) Com.
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 A11 shows that self-reported grades in English and math in the secondary school leaving exam (matric) correlate positively with jobseekers' assessed skills and their baseline beliefs about skills.** It displays correlations between comparative advantage assessed based on matric scores and comparative advantage assessed based on assessment results and beliefs in the tight experiment. The sample includes all individuals with a clear comparative advantage who reported matric scores in English and math. Columns 1 to 3 correlate matric grades with assessment results. Columns 4 to 7 correlate matric grades with baseline beliefs about skills. Columns 3, 6, and 7 correlate the difference in math and English matric grades with comparative advantage measures. Columns 1, 2, 4, and 5 correlate math and English matric scores with assessed skill quintiles (columns 1 and 2) and beliefs about skill quintiles (columns 4 and 5). Matric grades are rescaled to range from 0 to 1. No further control variables are included. Robust standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A12: Association Between Baseline Aligned Comparative Advantage Belief and Other Baseline Characteristics - Tight Experiment

	(1) Aligned CA belief	(2) Aligned CA belief
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 A12 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. Heteroskedasticity-robust standard errors shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A13: Development of Skill Beliefs Over Time - Big Experiment Control Group

	Aligned comparative advantage belief			% aligned skill belief		
	(1)	(2)	(3)	(4)	(5)	(6)
Endline	-0.004	-0.001	-0.015	0.009	0.009	0.004
	(0.011)	(0.016)	(0.013)	(0.008)	(0.011)	(0.010)
Endline × Above median search effort		-0.008			0.000	
		(0.025)			(0.014)	
Above median search effort		0.035*			0.003	
		(0.019)			(0.012)	
Endline × Worked last 7 days			0.034			0.015
			(0.022)			(0.014)
Worked last 7 days			-0.005			0.013
			(0.014)			(0.016)
Constant	0.200***	0.183***	0.202***	0.379***	0.377***	0.375***
	(0.009)	(0.014)	(0.010)	(0.007)	(0.007)	(0.009)
Observations	4405	4315	4315	4456	4365	4365

Notes: **Table A13 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 job-seeker level. Columns 1 to 3 show results for aligned beliefs about comparative advantage. Columns 4 to 6 show results for the fraction of aligned skill beliefs. Columns 2 and 5 show heterogeneity by whether individuals exerted above median search effort as measured on the search effort index described in Table A42. Columns 3 and 6 show heterogeneity by whether an individual worked in the last seven days. Heteroskedasticity robust standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A14: Treatment Effects on The Variance of Beliefs About Skills - Tight Experiment

	Average variance in beliefs		Variance numeracy beliefs		Variance communication beliefs	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-0.049	-0.071	-0.029	-0.062	-0.068	-0.080
	(0.054)	(0.077)	(0.058)	(0.087)	(0.055)	(0.076)
Treatment × Aligned comp adv belief (bl)		0.056		0.077		0.036
		(0.129)		(0.147)		(0.129)
Aligned comp adv belief (bl)		-0.148		-0.145		-0.152
		(0.095)		(0.101)		(0.099)
Treatment effect:		-0.015		0.014		-0.045
Aligned comp adv belief (bl)		(0.089)		(0.098)		(0.092)
Control mean	0.644	0.644	0.679	0.679	0.609	0.609
Observations	278	278	278	278	278	278

Notes: **Table A14 shows treatment effects on the variance of general, non-assessment specific, skill beliefs in the tight experiment.** Domain-specific beliefs about skill quintiles were elicited using the histogram method. Column 1 shows effects on the average variance. Column 2 shows effects on the variance of numeracy beliefs. Column 3 shows effects on the variance of communication skills. All specifications include randomization block fixed effects. Controls further include prespecified baseline covariates (skill quintile dummies, age, and dummies for having completed a university degree, other post-secondary education, and high school). Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A15: Treatment Effects on Over- and Underconfident Beliefs - Both Experiments

	Tight experiment (quintiles)		Big experiment (terciles)	
	(1)	(2)	(3)	(4)
	Underconfident	Overconfident	Underconfident	Overconfident
Treatment	-0.057***	-0.022	-0.042***	-0.101***
	(0.016)	(0.018)	(0.005)	(0.007)
Treatment effect/ control mean	-0.306	-0.035	-0.282	-0.219
	(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 A15 shows that underconfident beliefs are more likely to be updated than overconfident beliefs.** It displays treatment effects on dummies indicating the fraction of aligned, under, and overconfident skill beliefs. Columns 1 to 2 show results for the tight experiment. Columns 3 and 4 show results of the big experiment. Columns 1 and 3 show effects on a dummy indicating the fraction of skill domains on which beliefs about respondents' skills are underconfident relative to the assessment results. Columns 2 and 4 show effects on a dummy indicating the fraction of skill domains on which beliefs about respondents' skills are overconfident relative to the assessment results. 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. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A16: 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 A16 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. Heteroskedasticity robust standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A17: Association between Skill Beliefs and Search Direction - Big Experiment Control Group

	Target numeracy jobs		Target communication jobs		Target concept formation jobs	
	(1)	(2)	(3)	(4)	(5)	(6)
Believed numeracy comp. adv.	0.100***	0.090***	-0.094***	-0.080***	-0.003	-0.008
	(0.014)	(0.014)	(0.023)	(0.022)	(0.015)	(0.014)
Believed communication comp. adv.	-0.079***	-0.071***	0.087***	0.080***	-0.000	0.002
	(0.021)	(0.020)	(0.024)	(0.025)	(0.022)	(0.022)
Believed concept formation comp. adv.	-0.013	-0.018	-0.029	-0.024	0.067***	0.064***
	(0.019)	(0.018)	(0.017)	(0.019)	(0.019)	(0.020)
Control mean	0.222	0.222	0.471	0.471	0.220	0.220
Observations	2183	2179	2183	2179	2183	2179
Controls	No	Yes	No	Yes	No	Yes

Notes: **Table A17 shows that jobseekers' beliefs about comparative advantage correlate positively with skill-directed search in the control group of the big experiments.** 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. Independent variables are dummies for beliefs indicating that individuals have a clear comparative advantage in each skill. Even columns control for age, gender, having worked in a wage job, as well as dummies for three education categories, and measured comparative advantage. Heteroskedasticity robust standard errors 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), in Appendix C (skill belief measurement) and in Appendix I (willingness-to-pay 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[30], 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 A3 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[31]. 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

[30]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).

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

questions from the surveyor. On control days, participants still viewed a minimally modified version of the video that omitted the explanation of the results. The scripts and the video was thoroughly piloted to ease participants' understanding. The treatment video is available at https://www.dropbox.com/s/9qug1673wxefdxj/treatment_video_final.mp4?dl=0 and the control video is available at https://www.dropbox.com/s/t4d57xmwopy8r4q/control_video_final.mp4?dl=0. The corresponding scripts are available at https://www.dropbox.com/s/d2p3zh1raflwgsk/video_scripts_control_treatment.pdf?dl=0. Baseline covariates are balanced across treatment arms (Table A19).

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 A18 for the job title pairs). The job descriptions and 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 for three products described in detail in Appendix I. We then 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.

Table A18: 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 A18 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.

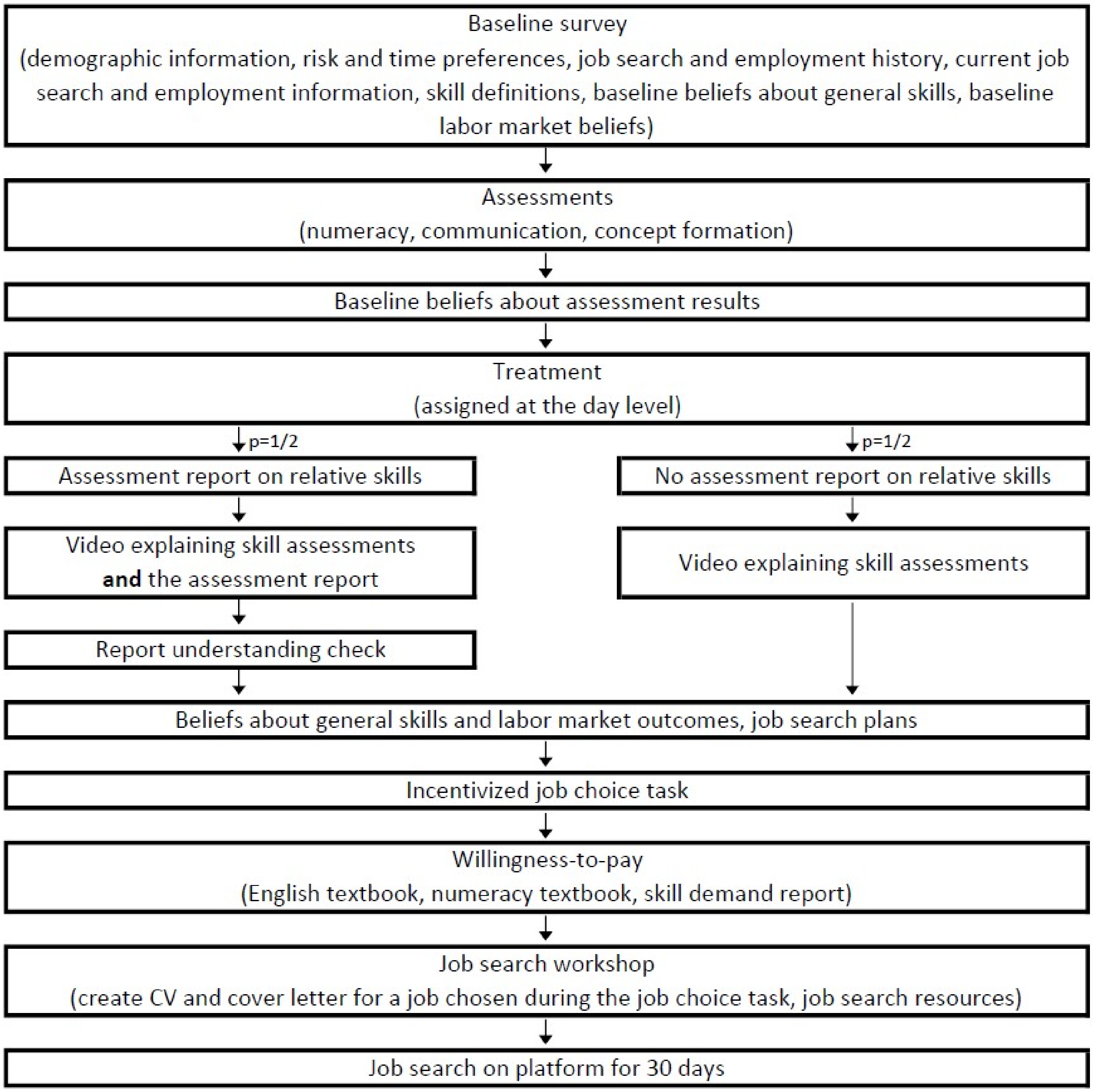
The final session of the workshop measured search effort. Participants had the opportunity to spend up to 15 minutes to write an optional short application email to the employer of the real job in the job choice task that they preferred. The message, along with participants' CV that we created, were 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 A4, 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

Figure A3: Tight Experiment Design



Notes: **Figure A3 shows the experimental design of the tight experiment.**

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 A5) 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

Table A19: Balance Table - Tight Experiment

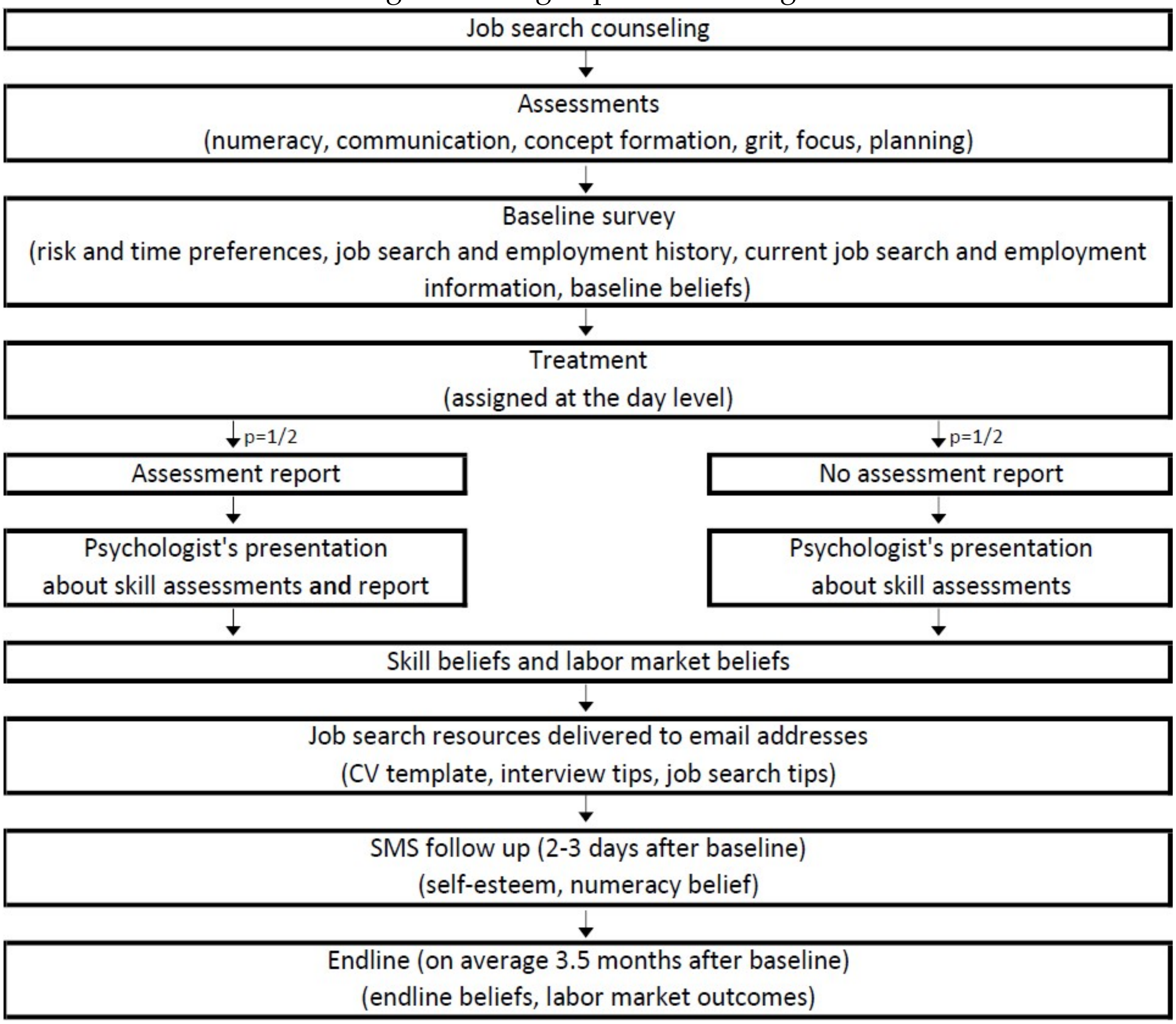
	Restricted sample					Full sample				
	(1) Control	(2) Treatment	(3) Δ	(4) $p(\Delta=0)$	(5) N	(6) Control	(7) Treatment	(8) Δ	(9) $p(\Delta=0)$	(10) N
Panel A: Demographics										
Black African	1.00	1.00	0.00	.	278	0.99	0.99	0.01	0.41	372
Male	0.33	0.32	-0.01	0.74	278	0.28	0.29	0.00	0.92	372
Age	26.42	26.40	-0.02	0.84	278	26.89	26.79	-0.09	0.91	372
Completed secondary education only	0.67	0.71	0.04	0.48	278	0.67	0.68	0.01	0.92	372
University degree / diploma	0.09	0.08	-0.01	0.91	278	0.10	0.06	-0.04	0.14	372
Other post-secondary education	0.14	0.14	-0.01	0.61	278	0.14	0.16	0.02	0.40	372
Panel B: Labor market background										
Any work in the 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 2021 USD (last 7 days, winsorized)	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, winsorized)	11.31	8.69	-2.62	0.10	278	11.46	10.49	-0.97	0.54	372
Search expenditure in 2021 USD (last 7 days, winsorized)	23.98	21.47	-2.51	0.25	278	24.42	22.65	-1.77	0.28	372
Hours spent searching (last 7 days, winsorized)	13.75	13.90	0.15	0.92	278	14.06	14.23	0.17	0.88	371
Job offers (last 30 days, winsorized)	0.14	0.21	0.07	0.11	278	0.20	0.23	0.03	0.51	372
Panel D: Search alignment with comparative advantage										
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	0.76	278	-0.00	-0.00	0.00	0.77	372
Panel E: Aligned comp. adv. belief										
General skill	0.47	0.50	0.04	0.68	278	0.43	0.46	0.03	0.69	369
Assessment result	0.53	0.54	0.01	0.95	278	0.45	0.46	0.02	0.82	371
Panel F: % aligned beliefs										
Assessment result	0.18	0.15	-0.03	0.34	278	0.15	0.14	-0.02	0.55	371
General skill	0.18	0.26	0.08	0.12	278	0.18	0.24	0.06	0.17	369

Notes: **Table A19 shows that covariates are balanced across treatment groups in the tight experiment.** Columns 1 to 5 show results for the sample of individuals with a clear comparative advantage in skills. Columns 6 to 10 show results for the full sample of individuals including those without unique comparative advantage in skills.

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://www.dropbox.com/s/b3lumodiz1gxadq/4c_Presentation_Private_24Oct2016.pptx?dl=0 and the script is available at https://www.dropbox.com/s/4lhxf502sy5kw3r/4d_Private%20Briefing_24Oct2016.docx?dl=0. Baseline covariates are balanced across treatment arms (Table A20).

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 A21).

Figure A4: Big Experiment Design



Notes: **Figure A3 shows the experimental design of the big experiment.**

Figure A5: 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.

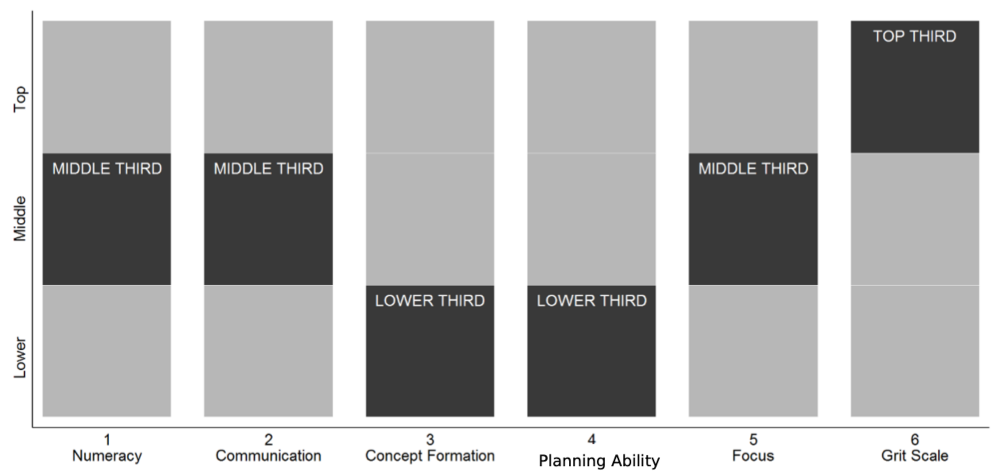
1. The Numeracy tests measure various maths abilities. Your score is the average of the two maths tests you did today at Harambee.
2. The Communication test measures English language ability through listening, reading and comprehension.
3. The Concept Formation test measures the ability to understand 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 A5 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 A20: Balance Table - Big Experiment

	Full sample					Non-attritted sample				
	(1) Control	(2) Treatment	(3) Δ	(4) $p(\Delta = 0)$	(5) N	(6) Control	(7) Treatment	(8) Δ	(9) $p(\Delta = 0)$	(10) N
Panel A: Demographics										
Black African	0.93	0.91	-0.03	0.34	4389	0.93	0.91	-0.03	0.30	4206
Male	0.39	0.36	-0.03	0.04	4389	0.39	0.36	-0.03	0.03	4206
Age	23.55	23.79	0.29	0.07	4389	23.53	23.80	0.31	0.05	4206
Completed secondary education only	0.62	0.59	-0.04	0.20	4389	0.62	0.59	-0.04	0.17	4206
Any other post-secondary qualification	0.21	0.22	0.01	0.61	4389	0.22	0.23	0.01	0.63	4206
University degree / diploma	0.16	0.18	0.03	0.29	4389	0.15	0.18	0.04	0.24	4206
Panel B: Labor market background										
Worked in past 7 days	0.36	0.39	0.02	0.22	4389	0.36	0.39	0.02	0.19	4206
Earnings in USD (last 7 days, winsorized)	30.09	32.52	3.58	0.13	4389	29.84	32.49	3.79	0.11	4206
Has worked in permanent wage job before	0.09	0.09	-0.01	0.61	4377	0.10	0.09	-0.01	0.35	4195
Panel C: Search behavior										
Did any search in past 7 days	0.97	0.97	0.01	0.07	4389	0.97	0.98	0.01	0.02	4206
# applications (last 30 days, winsorized)	9.31	9.37	0.03	0.95	4346	9.13	9.27	0.13	0.79	4165
Search expenditure in USD (last 7 days, winsorized)	32.09	31.14	-1.11	0.34	3912	32.02	31.15	-0.95	0.42	3747
# job offers (last 30 days, winsorized)	0.77	0.84	0.04	0.64	4335	0.78	0.87	0.07	0.44	4152
Panel D: Skills beliefs										
Aligned belief about comparative advantage	0.20	0.21	0.01	0.49	4312	0.20	0.21	0.01	0.67	4132
% aligned beliefs	0.38	0.38	-0.01	0.49	4378	0.38	0.37	-0.01	0.45	4196

Notes: **Table A20 shows that covariates are balanced across treatment arms for the big experiment.** Columns 1 to 5 show statistics for the baseline sample. Columns 6 to 10 show statistics for the sample of individuals reached at endline.

Table A21: 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 A21 shows that attrition is low and balanced across treatment groups in the big experiment.** Dependent variable is a dummy indicating whether an individual was recontacted at for the endline survey. Column 2 includes pre-specified baseline covariates (measured skills, self-reported skills, education, age, gender, employment, discount rate, and risk aversion). Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, ***

E Additional Treatment Effects

E.1 Robustness Checks

In this section we first evaluate robustness for sample definitions. In the tight experiment we restrict the sample to jobseekers who have a clear comparative advantage to allow us to define whether their comparative advantage beliefs are aligned with their assessed CA. In this section we show that our results are robust to this sample restriction. To complete this analysis we define the aligned comparative advantage indicator to be zero for those who have tied skill quintile and reestimate the treatment effects. Table A22 shows the robustness for the skill belief treatment effects and Table A23 shows it for search direction in the tight experiment. In Table A24 we present our main results for the tight and big experiment side-by-side.

Next, we analyze robustness for using a continuous measure for belief alignment in Table A25, after which we explore robustness on alternative ways to define our outcome measures. Table A26 shows that our results are robust to different transformations of earnings. Figure A6 supports this finding showing that the treatment effects on conditional and unconditional earnings are driven by outliers. Table A27 explores treatment effects on work quality outcomes (earnings and written contract) multiplied by potential mediators, such as tenure and start date.

Table A22: Treatment Effects on Skill Beliefs - Tight Experiment Including Jobseekers Without Unique Comparative Advantage

	Aligned comp. adv. 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 comp adv belief (bl)		-0.130		0.057
		(0.089)		(0.041)
Aligned comp adv belief (bl)		0.447***		-0.039
		(0.081)		(0.037)
Treatment effect:		0.050		0.121***
Aligned comp adv belief (bl)		(0.061)		(0.036)
Control mean	0.446	0.446	0.171	0.171
Observations	368	368	368	368

Notes: **Table A22 shows that treatment effects on beliefs about skills are robust to including people without a clear comparative advantage in skills.** 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 quintile beliefs that are aligned with the assessment results. Columns 2 and 4 show treatment effect heterogeneity by whether individuals had aligned comparative advantage beliefs at baseline. Controls include randomization block fixed effects, and prespecified baseline covariates (skill quintile dummies, age, gender, and dummies for having completed a university degree, other post-secondary education, and high school). Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A23: 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***	0.435***	0.032	0.062**	0.052***	0.062**	0.071	0.122*	1.121	3.363***
	(0.078)	(0.094)	(0.025)	(0.026)	(0.017)	(0.030)	(0.046)	(0.063)	(0.878)	(1.140)
Treatment × Aligned comp adv belief (bl)		-0.445***		-0.070**		-0.022		-0.113		-5.190**
		(0.156)		(0.029)		(0.060)		(0.098)		(2.162)
Aligned comp adv belief (bl)		0.504***		0.111***		0.010		0.070		6.113***
		(0.121)		(0.026)		(0.035)		(0.076)		(1.872)
Treatment effect:		-0.010		-0.008		0.040		0.009		-1.827
Aligned comp adv 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 A23 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. Aligned job search is defined as directing search effort toward jobs that mostly require the skill that aligns with job seekers' measured comparative advantage. Columns 1 and 2 show impacts on an index of search direction. The index is constructed as the variance-covariance weighted average of the search alignment measures displayed in columns 3 to 10 (Anderson, 2008). Columns 3 and 4 show the impact on the percentage of 11 incentivized job choices that are aligned with the comparative advantage. Columns 5 and 6 show impact on the difference between the percentage of aligned and non-aligned applications on the online job search platform SAyouth.mobi. Jobs are classified using a list of 20 numeracy-heavy and 20 communication-heavy common job titles classified by experts. Columns 7 and 8 show impact on the difference between aligned and non-aligned applications (winsorized). Columns 9 and 10 show the impact on the difference in link click rates between aligned and non-aligned jobs sent to job seekers via SMS. All job seekers receive three links with at least one job for each skill. Even columns show heterogeneity by whether individuals have aligned comparative advantage beliefs at baseline. Controls include randomization block fixed effects, and prespecified baseline covariates (skill quintile dummies, age, gender, dummies for having completed a university degree, other post-secondary certificate or diploma, and high school, and a dummy indicating a comparative advantage in numeracy). Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A24: Robustness Checks for Main Outcomes - Comparing Tight and Big Experiments

	Tight experiment		Big experiment			
	(1) Aligned comp. adv. belief	(2) Aligned search index	(3) Aligned comp. adv. belief	(4) Aligned search direction	(5) LM quantity index	(6) LM quality index
Panel A: Robustness for inference						
Treatment	0.135	0.241	0.139	0.050	0.045	0.085
unadjusted p-value	(0.001)	(0.023)	(0.000)	(0.000)	(0.165)	(0.017)
wild-bootstrapped p-value	[0.002]	[0.015]	[0.000]	[0.000]	[0.390]	[0.004]
q-value	{0.002}	{0.014}	{0.001}	{0.001}	{0.029}	{0.013}
Panel B: Reweighting to match other sample						
Treatment	0.152***	0.241**	0.142***	0.050***	0.054*	0.089**
	(0.044)	(0.101)	(0.011)	(0.010)	(0.031)	(0.036)
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 A24 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 unadjusted, wild-bootstrapped p-values, as well as sharpened q-values (Benjamini et al., 2006) 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 dummies for being male, just having matric, having a post-secondary degree, or having a post-secondary certificate, having worked in the last seven days, having searched for work in the last seven days, and for having ever worked in a wage job as well as the number of job applications 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. Controls are the same as in the main specifications. 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 A25: Heterogeneous Treatment Effects by Non-binary Baseline Comparative Advantage Beliefs - Tight Experiment

	Main outcomes				Additional belief outcomes			
	Aligned search index		Aligned comp. adv. belief		Degree of CA alignment		Overall alignment of belief levels	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.269**	0.841**	0.135***	0.390***	0.063***	0.324***	0.073***	0.138**
	(0.103)	(0.413)	(0.035)	(0.125)	(0.012)	(0.091)	(0.012)	(0.062)
Treatment × Baseline degree of CA belief alignment		-0.680		-0.302*		-0.315***		-0.078
		(0.503)		(0.155)		(0.106)		(0.072)
Baseline degree of CA belief alignment		1.592***		0.460**		0.828***		0.045
		(0.495)		(0.221)		(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 A25 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.** The non-binary measure of comparative advantage beliefs captures the degree to which comparative advantage beliefs are misaligned and ranges from 0 to 1. If comparative advantage beliefs are aligned it is set to 1. If they are misaligned 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). A value of zero means that jobseekers need to shift their skill beliefs by a total of five quintiles to align their comparative advantage belief with the measured comparative advantage. The mean of this measure is 0.83 and the standard deviation is 0.22. Columns 1 to 4 show robustness of the main result heterogeneity to using a non-binary measure of aligned comparative advantage beliefs. Columns 1 and 2 show effects on the aligned search direction index described in table 3. Columns 3 and 4 show effects on a dummy indicating aligned comparative advantage beliefs. Columns 5 to 8 show treatment effects on additional measures of belief alignment. Columns 5 and 6 show treatment effects on a non-binary measure of aligned comparative advantage beliefs. Columns 7 and 8 show treatment effects on a measure of overall alignment of skill beliefs with measured skills. This 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. Controls include randomization block fixed effects, and prespecified baseline covariates (skill quintile dummies, age, gender, and dummies for having completed a university degree, other post-secondary education, and high school). Standard errors clustered at the treatment-day level in parenthesis. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

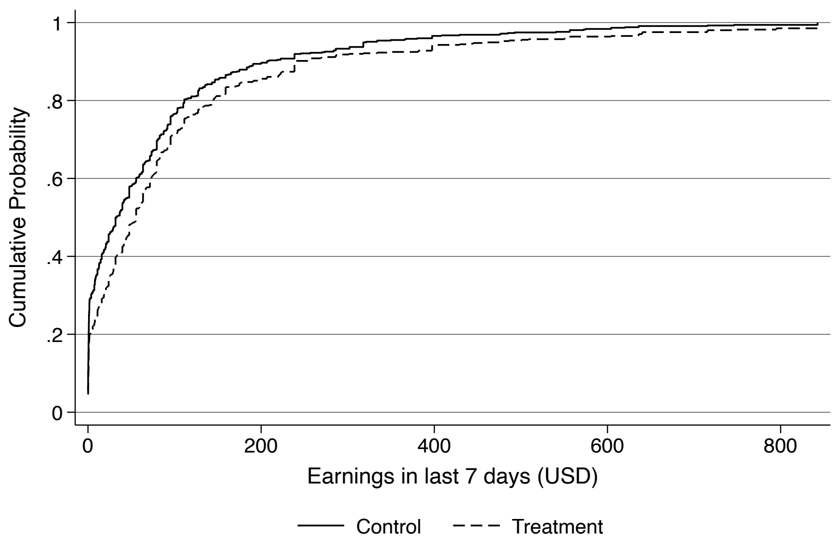
Table A26: Treatment Effects on Different Transformations of Earnings - Big Experiment

	Unconditional earnings					Conditional earnings				
	(1) raw	(2) wins	(3) z	(4) ln(earn+1)	(5) ihs	(6) raw	(7) wins	(8) z	(9) ln(earn+1)	(10) ihs
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

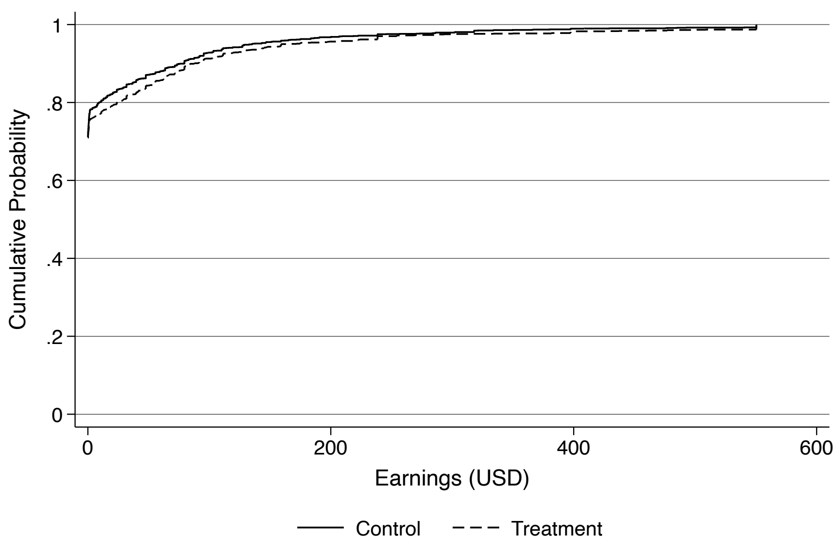
Notes: **Table A26 shows that treatment effects on earnings in the last seven days are robust to a range of functional forms.** Earnings are measured in 2021 USD PPP. Columns 1 to 5 show effects on unconditional earnings. Columns 6 to 10 show effects on earnings conditional on doing any work. Columns 1 and 6 show effects on the raw variable. Columns 2 and 7 show effects on the variable winsorized at the 99^{th} percentile. Columns 3 and 8 show effects on the variable standardized using the control group mean and standard deviation. Columns 4 and 9 show effects on the natural logarithm of the variable plus one. Columns 5 and 10 show effects on the inverse hyperbolic sine of the variable. Controls include randomization block fixed effects, and prespecified baseline covariates (measured skills, self-reported skills, education, age, gender, employment, discount rate, and risk aversion). Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure A6: Cumulative Distributions of Earnings by Treatment Group - Big Experiment

Panel A: Conditional earnings



Panel B: Unconditional earnings



Notes: **Figure A6 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.

Table A27: Treatment Effects on Combinations of Labor Market Outcome Measures - Big Experiment

	Earnings (w)					Written contract	
	(1) Start before treatment	(2) Start after treatment	(3) Tenure	(4) Wage employment	(5) Self employment	(6) Start before treatment	(7) Start after treatment
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 A27 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 multiplied by potential mediators. Columns 1 to 5 show impacts on earnings in the last seven days (winsorized at the 99^{th} percentile). Columns 6 and 7 show impacts on a written contract dummy. Column titles indicate the variable with which the outcome is multiplied. Controls include randomization block fixed effects, and pre-specified baseline covariates (measured skills, self-reported skills, education, age, gender, employment, discount rate, and risk aversion). Standard errors clustered at the treatment-day level in parenthesis. All monetary figures are reported in 2021 USD PPP. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

E.2 Heterogenous Treatment Effects

This section collects various heterogeneity analysis. In Table A28 we assess whether jobseekers learn about themselves or the distribution of skills in the reference population. To do so we ask jobseekers after the assessments but before treatment how many questions they correctly answered in the assessments.Table A29 combines the tight and big experiment to show that there are no heterogeneous treatment effects on belief updating by the level of concept formation skill. Table A30 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 A31 shows treatment effect heterogeneity by baseline comparative advantage beliefs and confidence simultaneously. Table A32 displays treatment effects on search direction in the tight experiment by whether skill requirements were revealed. Table A33 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 A28: Learning About Oneself or The Distribution of Skills? - Tight Experiment

	Aligned comp. adv. belief			Av. abs. deviation from assessments		
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	0.171**	0.208***	0.176***	-0.090***	-0.086***	-0.090***
	(0.065)	(0.050)	(0.060)	(0.014)	(0.014)	(0.015)
Treatment × Aligned abs. comp adv.	-0.037		0.139			
	(0.096)		(0.150)			
Aligned abs. comp adv.	0.473***		0.117			
	(0.118)		(0.138)			
Treatment × Aligned rel. comp adv.		-0.137	-0.240			
		(0.082)	(0.148)			
Aligned rel. comp adv.		0.586***	0.543***			
		(0.079)	(0.112)			
Treatment × Abs. dev. % correct (z)				-0.030**		-0.029**
				(0.013)		(0.013)
Abs. dev. % correct (z)				0.032***		0.025*
				(0.011)		(0.013)
Treatment × Abs. dev. quintile (z)					0.012	0.006
					(0.013)	(0.012)
Abs. dev. quintile (z)					-0.022**	-0.014
					(0.009)	(0.009)
Control mean	0.475	0.475	0.475	0.360	0.360	0.360
Observations	278	278	278	278	278	278

Notes: **Table A28 explores whether belief updating is driven by learning about oneself or about the distribution of skills by others.** Columns 1 to 3 show effects on beliefs about jobseekers' comparative advantage. Columns 4 to 6 show effects on the expected absolute deviation of beliefs about skills (elicited using the histogram method) from measured quintiles. Columns 1 and 4 display heterogeneity by baseline measures defined using beliefs about the absolute performance on the assessments (in percent correctly answered questions). Columns 2 and 5 display heterogeneity by baseline measures defined using the relative performance on assessments (in quintiles). Columns 3 and 6 display heterogeneity by both measures in the same regression. All specifications include randomization block fixed effects. Controls further include prespecified baseline covariates (skill quintile dummies, age, and dummies for having completed a university degree, other post-secondary education, and high school). 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 Belief Updating by Cognitive Skill - Both Experiments

	Fraction aligned	
	(1) Tight experiment	(2) Big experiment
Treatment	0.079	0.164
	(0.026)	(0.011)
Treatment × CFT	0.018	0.005
	(0.020)	(0.008)
CFT	-0.004	0.005
	(0.016)	(0.007)
Control mean	0.183	0.410
Observations	278	4205
Number of clusters	34	54

Notes: **Table A29 shows that treatment effects on belief updating do not vary by a proxy for cognitive skill.** The proxy is the standardized score on the concept formation assessment, which is similar to the Raven's matrices assessment. Column 1 regresses the share of aligned beliefs on treatment, standardized CFT score (Raven's matrices), their interaction and the usual controls in the tight experiment. Column 2 does the same for the big experiment, using the standardized concept formation test score and the respective controls. Standard errors clustered at the treatment-day level.

Table A30: Heterogeneous Treatment Effects on Skill Beliefs and Search Direction by Aligned Baseline Comparative Advantage Belief - Big Experiment

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

Notes: **Table A30 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.** Column 1 shows the impact on aligned comparative advantage beliefs. Column 2 shows the impact on the fraction of aligned belief domains. Column 3 shows the impact on a dummy indicating search direction aligned with comparative advantage in skills (self-reported). Controls include randomization block fixed effects, and prespecified baseline covariates (measured skills, self-reported skills, education, age, gender, employment, discount rate, and risk aversion). Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A31: Heterogeneous Treatment Effects on Main Outcomes by Comparative Advantage Beliefs and Confidence - Tight Experiment

	(1) Aligned comp. adv. belief	(2) Aligned search index	(3) Search effort index
Treatment	0.216***	0.551***	-0.194
	(0.066)	(0.155)	(0.177)
Treatment × Aligned comp adv belief (bl)	-0.161	-0.651**	0.276
	(0.112)	(0.247)	(0.298)
Aligned comp adv belief (bl)	0.737***	0.539**	-0.056
	(0.110)	(0.212)	(0.226)
Treatment × Average confidence (bl)	-0.003	0.084	0.082
	(0.079)	(0.179)	(0.176)
Average confidence (bl)	0.072	0.061	0.200
	(0.058)	(0.142)	(0.143)
Treatment × Aligned comp adv belief (bl) × Average confidence (bl)	0.013	-0.064	-0.170
	(0.111)	(0.208)	(0.230)
Control mean	0.475	0.000	0.000
Observations	278	278	278

Notes: **Table A31 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 between the treatment dummy and baseline comparative advantage does not change much relative to Tables 2 and 3. Moreover, the magnitude of the three-way interaction term is small and insignificant throughout suggesting that the confidence levels do not affect treatment effect heterogeneity by baseline comparative advantage beliefs. 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. Column 1 shows effects on a dummy indicating aligned comparative advantage beliefs. Column 2 shows effects on the aligned search index. Column 3 shows effects on the search effort index. Controls are the same as in the main specifications. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A32: Heterogeneous Treatment Effects on Search Direction for Jobs With and Without Revealed Skill Demands - Tight Experiment

	Choice aligned with comparative advantage	
	(1)	(2)
	Misaligned baseline beliefs	Aligned baseline beliefs
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	0.413	0.413
Observations	1573	1485

Notes: **Table A32 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 table shows the impact on dummy variables indicating job choices aligned with jobseekers' assessed comparative advantage estimated at the job-choice individual level. Column 1 shows the impact for individuals with initially misaligned beliefs. Column 2 shows the impact for individuals with initially aligned beliefs. Controls include randomization block fixed effects, job pair and job pair order fixed effects, and pre-specified baseline covariates (skill quintile dummies, age, gender, dummies for having completed a university degree, other post-secondary certificate or diploma, and high school, and a dummy indicating a comparative advantage in numeracy). Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A33: Heterogeneous Treatment Effects on Labor Market Outcomes by Aligned Baseline Comparative Advantage Beliefs - Big Experiment

	Work quantity					Work quality			
	(1) Index	(2) Job offers (w)	(3) Worked month 1	(4) Worked month 2	(5) Worked last 7 days	(6) Index	(7) Earnings (w)	(8) Hourly wage (w)	(9) Written contract
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 comp adv 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 comp adv 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: Aligned comp adv belief (bl)	-0.015	0.013	-0.019	-0.006	-0.004	0.043	2.770	0.051	0.016
	(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 A33 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 1 to 5 show impacts on employment quantity. Column 1 shows impacts on an Anderson (2008) index of the four employment quantity measures. Column 2 shows the impact on the number of job offers in the last 30 days (winsorized). Column 3 shows the impact on a dummy indicating any work for pay in month 1 after treatment. Column 4 shows the impact on a dummy indicating any work for pay in month 2 after treatment. Column 5 shows the impact on a dummy indicating any work for pay in the last seven days. Columns 6 to 9 show impacts on employment quality. Column 6 shows impacts on an Anderson (2008) index of the three employment quality measures. Column 7 shows the impact on earnings in the last seven days (winsorized). Column 8 shows the impact on hourly wages in the last seven days (winsorized). Column 9 shows the impact on a dummy indicating a written contract. Controls include randomization block fixed effects, and prespecified baseline covariates (measured skills, self-reported skills, education, age, gender, employment, discount rate, and risk aversion). 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 → expected return to skill-directed search → search direction → 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 separately 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 CA. For example, for a jobseeker with numeracy CA, 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 CA beliefs (Table A34, columns 3-4). Treated jobseekers with baseline misaligned CA 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 103.

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 CA beliefs increase their expected returns by 0.26 standard deviations ($p = 0.054$) while jobseekers with initially aligned CA beliefs do not increase their expected returns (column 2).

Reassuringly, we see a 'sensible' relationship between baseline CA beliefs and expected returns to skill-directed job search. Jobseekers whose baseline CA belief matches their assessed CA 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 A35). 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 A36).

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 on post-treatment beliefs. Column 1 of Table A37 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 (5).

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 A36 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).[32] These positive effects on wage beliefs in the tight experiment are consis-

[32]In standard job search models, the reservation wage is both a decision rule and a feature of the wage

tent 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 A38 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.[33] 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.

distribution. This is not inconsistent with our interpretation of reservation wages as another proxy for wage expectations.

[33] 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 A34: 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 comp adv belief (bl)		-0.341*		-0.023		0.114		-45.738***
		(0.180)		(0.034)		(0.223)		(13.960)
Aligned comp adv 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 comp adv 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	278	273	273	272	272	278	278

Notes: **Table A34 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 higher beliefs about the returns to skill-directed search. Columns 1 and 2 show treatment effects on an index of belief about the returns to directed search. Columns 3 and 4 display effects on the difference in expected offers per application for jobs that align with jobseekers' comparative advantage and those that do not align with their comparative advantage. Columns 5 and 6 display effects on the difference in expected search duration for jobs between searching for jobs that *do not* align with jobseekers' assessed comparative advantage in and jobs that align with their assessed comparative advantage (winsorized). Columns 7 and 8 display effects on expected weekly wages for jobs that align with jobseekers' comparative advantage and those that do not align with their comparative advantage measured in 2021 PPP USD (winsorized). Even columns show heterogeneity by whether individuals have aligned comparative advantage beliefs at baseline. Controls include randomization block fixed effects, and prespecified baseline covariates (skill quintile dummies, age, gender, dummies for having completed a university degree, other post-secondary certificate or diploma, and high school, and a dummy indicating a comparative advantage in numeracy). Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A35: 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
	(0.031)	(6.739)	(0.022)
Treatment× Aligned skill req.	0.008	3.873	0.011
	(0.030)	(3.491)	(0.015)
Aligned skill req.	0.030	-3.297	0.017
	(0.023)	(2.461)	(0.010)
Control mean	-0.000	193.588	0.544
Observations	2770	2770	2770

Notes: **Table A35 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. Column 1 shows effects on a standardized measure of desirability of jobs measured on a 0 to 10 Likert-scale. Column 2 shows effects on expected weekly earnings winsorized at the 99^{th} percentile. Column 3 shows effects on the perceived job offer probability. Analysis is at the job-individual level and includes pre-specified control variables (skill quintile dummies, age, gender, and dummies for having completed a university degree, other post-secondary education, and high school) as well as randomization block, job choice order and job fixed effects. Standard errors clustered at the treatment-day level are in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A36: 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 comp adv belief (bl)		-0.461*		-26.041		-21.065		-6.425		-5.366
		(0.258)		(21.034)		(17.531)		(38.832)		(6.578)
Aligned comp adv 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 A36 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. Columns 1 and 2 show effects on an Anderson (2008) wage expectation index. All monetary figures are reported in 2021 USD PPP per week. (w) indicates measures winsorized at the 99^{th} percentile. Columns 3 to 10 show treatment effects on the individual index components. Columns 3 and 4 show impacts on expected wages. Columns 5 and 6 show impacts on the minimum expected wage. Columns 7 and 8 show impacts on the minimum expected wage. Columns 9 and 10 show impacts on the reservation wage. Even columns show heterogeneity by whether jobseekers had aligned comparative advantage beliefs at baseline. Controls include randomization block fixed effects, and prespecified baseline covariates (measured skills, self-reported skills, education, age, gender, employment, discount rate, and risk aversion). Standard errors clustered at the treatment-day level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A37: Association Between Job Choices, Offer Probabilities, and Expected Wages - Tight Experiment

	Marginal effects on choice of numeracy job (logit estimation)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$E[Wage^{num}] - E[Wage^{com}]$ (w)	0.067***	0.075***	0.065**	0.075**	0.067***	0.077***	0.065**	0.073**
	(0.025)	(0.026)	(0.029)	(0.031)	(0.025)	(0.025)	(0.030)	(0.030)
Treatment $\times$ $E[Wage^{num}] - E[Wage^{com}]$ (w)					0.045	0.042	-0.017	-0.021
					(0.038)	(0.038)	(0.045)	(0.044)
# observations	925	925	800	800	1840	1840	1610	1610
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 A37 shows that, in the job choice task in the tight experiment, expected earnings are predictive of job choice and that treatment does not affect this relationship.** It displays the marginal effects of expected returns to applying for a job and job choices estimated using a logit estimator in the tight experiment. The outcome is a dummy indicating choosing the numeracy job in a job pair. The independent variable is the winsorized difference in expected returns to applying to the numeracy job and the communication job. Expected returns are defined as expected weekly wages (in 100s USD PPP in 2021) times the perceived likelihood of receiving a job offer. Columns 1 to 4 are limited to the control group. Columns 5 to 8 also interact the expected returns with a treatment indicator (columns 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A38: 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 comp adv 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 comp adv 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 A38 shows that there are positive but mostly insignificant treatment effects on beliefs about job offers and search duration in the tight experiment.** Columns 1 and 2 show effects on an Anderson (2008) labor market expectation index. Columns 3 to 12 show effects on the individual index components. Columns 3 and 4 show impacts on expected callbacks per application. Columns 5 and 6 show impacts on expected offers per application. Columns 7 and 8 show impacts on the expected time until jobseekers find a full-time job (this enters the index negatively). Columns 9 and 10 show impacts on the perceived probability of being employed one month after baseline. Columns 11 and 12 show impacts on the perceived probability of being employed three months after baseline. (w) indicates measures winsorized at the 99^{th} percentile. Even columns show heterogeneity by whether jobseekers had aligned comparative advantage beliefs at baseline. Controls include randomization block fixed effects, and prespecified baseline covariates (measured skills, self-reported skills, education, age, gender, employment, discount rate, and risk aversion). 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.

Our treatments contain information about jobseekers' skill levels in addition to information about their assessed comparative advantage. Jobseekers in our samples 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 (Table A15). 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 A39, column 3, $p = 0.025$). 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 A39, column 7, $p < 0.001$). The negative effects make sense, because more jobseekers are overconfident than underconfident at baseline.

Treatment effects vary by baseline beliefs about skill levels. To show this, we construct a measure of baseline confidence levels ($confidence_i$): believed skill level minus assessed skill level. This uses quintiles and averages over two skills in the tight experiment and uses terciles and averages over three skills in the big experiment. We then estimate:

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

If $\beta_2 < 0$, this means that that receiving negative information about skill levels lowers Y_{id}.

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 A39, row 2). One standard deviation higher baseline confidence is associated with a 0.26 quintile more negative treatment effect in the tight experiment (column 4, $p < 0.001$) and a 0.125 terciles more negative treatment in the big experiment (column 8, $p < 0.001$).

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, $A\left(E_C, E_N\right)$. 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, or an intertemporal budget constraint. The jobseeker's problem becomes

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

As in Section 2.1, we assume 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.[34]

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.[35] Treatment lowers an index of these search effort measures by 0.08 standard deviations (Table A40, column 1, $p = 0.47$) and a

[34]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.

[35]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.

prespecified index of search effort on the SAYouth.mobi platform by 0.1 SDs (Table A41, 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 A42, 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.

Treatment effects also do not vary substantially by jobseekers' baseline confidence levels in either experiment. We estimate equation (7) and show heterogeneous treatment effects by baseline skill belief confidence in even-numbered columns in Tables A40, A41, and A42. None of the interaction terms are statistically significant, the effects size 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 A40, 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 A31). The platform-based measure has a somewhat larger interaction effect of 0.14 standard deviations but it is still not statistically significant (Table A40, 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 A42, column 2, $p = 0.41$). The interaction effects on the index components are positive but small and not statistically significant.

Finally, we show that treatment effects on labor market outcomes in the big experiment do not vary substantially by baseline confidence about skills (Table A43) 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. (2021), and Banerjee & Sequeira (2020), and Mueller & Spinnewijn (2022) for mixed results about the sign of the relationship between search effort and beliefs about labor market prospects.

Table A39: Heterogeneous Treatment Effects on Skill Beliefs by Baseline Confidence

	Tight experiment				Big experiment			
	Average abs. deviation		Average skill quintile beliefs (0-4)		Average abs. deviation		Average skill tercile beliefs (0-2)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	-0.085***	-0.080***	-0.082**	0.157*	-0.192***	-0.157***	-0.109***	-0.007
	(0.014)	(0.017)	(0.035)	(0.083)	(0.012)	(0.015)	(0.010)	(0.011)
Treatment × Average confidence (bl)		-0.007		-0.263***		-0.043***		-0.125***
		(0.014)		(0.060)		(0.012)		(0.012)
Average confidence (bl)		0.021		0.512***		0.101***		-0.063***
		(0.016)		(0.071)		(0.019)		(0.017)
Control mean	0.360	0.360	2.693	2.693	0.734	0.734	1.446	1.446
Observations	278	278	278	278	4191	4118	4195	4131

Notes: **Table A39 shows that beliefs about relative skill level update downwards and differentially depending on jobseekers' baseline confidence about their relative skills.** Baseline confidence levels are defined as the average baseline deviation of quintile skill beliefs from the assessed quintile across numeracy and communication (tight experiment) or tercile across numeracy, communication, and concept formation (big experiment). Confidence measures are standardized to have a control group standard deviation of one. Positive values indicate overconfidence and negative values indicate underconfidence. Columns 1 to 4 show results for the tight experiment. Columns 5 to 8 show results for the big experiment. Columns 1, 2, 5, and 6 show impacts on the average absolute deviation of beliefs from measured skills. Columns 3, 4, 7, and 8 show impacts on average skill beliefs. Columns 2, 4, 6, and 8 show effects by baseline confidence. Controls include randomization block fixed effects, and prespecified baseline covariates. 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 Search Effort - Tight Experiment

	Index		Planned apps (w)		Time spent drafting (w)		SMS click rate		Observed apps (w)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment	-0.076	-0.101	-3.854	-3.574	-0.530	-0.377	0.003	-0.055	-1.408	-2.043
	(0.103)	(0.145)	(2.555)	(4.124)	(0.591)	(0.925)	(0.032)	(0.053)	(1.551)	(2.379)
Treatment $\times$ Average confidence (bl)		0.019		-0.429		-0.205		0.062		0.617
		(0.111)		(2.936)		(0.771)		(0.045)		(2.068)
Average confidence (bl)		0.196		4.032		1.280		-0.059		1.233
		(0.124)		(2.675)		(0.900)		(0.046)		(1.912)
Control mean	0.000	0.000	37.878	37.878	8.828	8.828	0.635	0.635	15.187	15.187
Observations	278	278	278	278	267	267	278	278	278	278

Notes: **Table A40 shows that treatment effects on search effort in the tight experiment do not vary significantly by baseline confidence levels.** Baseline confidence levels are defined as the average baseline deviation of quintile skill beliefs from measured quintiles (across numeracy and communication) standardized to have control group standard deviation of one. Positive values indicate overconfidence and negative values indicate underconfidence. Columns 1 and 2 show impacts on an Anderson (2008) index of four search effort measures. Columns 3 and 4 show impacts on the number of planned applications in the next 30 days (winsorized). Columns 5 and 6 show impacts on the number of minutes individuals spent drafting a cover letter during the job-search workshop (winsorized). Columns 7 and 8 show impacts on the click rate for three SMS with links to job adverts sent to individuals. Columns 9 and 10 show impacts on the number of applications sent on the job-search platform in the 30 days following the treatment (winsorized). Controls include randomization block fixed effects, and prespecified baseline covariates. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A41: 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	-0.238	-0.723	-1.451**	-0.206	-2.906	-1.463	-2.058
	(0.092)	(0.156)	(0.477)	(0.598)	(1.519)	(3.119)	(1.553)	(2.385)
Treatment × Average confidence (bl)		0.142		0.759		2.848		0.576
		(0.121)		(0.449)		(2.580)		(2.073)
Average confidence (bl)		-0.066		-0.278		-2.303		1.239
		(0.111)		(0.530)		(2.279)		(1.918)
Control mean	-0.000	-0.000	6.014	6.014	9.058	9.058	15.216	15.216
Observations	278	278	278	278	278	278	278	278

Notes: **Table A41 shows that treatment effects on prespecified search effort measures on the job-search platform SAYouth.mobi in the tight experiment do not vary significantly by baseline confidence levels.** 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 1 and 2 show impacts on an Anderson (2008) index of the three search effort measures. Columns 3 and 4 show impacts on the number of days jobseekers were active on the platform in the 30 days following the treatment. Columns 5 and 6 on the number of job adverts jobseekers clicked on on the platform in the 30 days following the treatment (winsorized). Columns 7 and 8 show impacts on the number of applications sent on the job-search platform in the 30 days following the treatment (winsorized). Controls include randomization block fixed effects, and prespecified baseline covariates. Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A42: 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	-0.018	0.056	0.010	-0.321	-0.601	0.243	-0.077
	(0.032)	(0.038)	(0.488)	(0.572)	(0.318)	(0.498)	(0.818)	(0.855)
Treatment × Average confidence (bl)		0.030		0.141		0.355		0.493
		(0.036)		(0.464)		(0.490)		(0.675)
Average confidence (bl)		-0.034		-0.073		-0.604		-0.416
		(0.049)		(0.679)		(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 A42 shows that treatment effects on search effort in the big experiment do not vary significantly by baseline confidence levels.** Baseline confidence levels are defined as the average baseline deviation of tercile skill beliefs from measured terciles (across numeracy, communication, and concept formation) standardized to have control group standard deviation of one. Positive values indicate overconfidence and negative values indicate underconfidence. Columns 1 and 2 show impacts on an Anderson (2008) index of three search effort measures. Columns 3 and 4 show impacts on the number of applications in the last 30 days (winsorized). Columns 5 and 6 show impacts on the number of hours spent searching for jobs in the last 30 days (winsorized). Columns 7 and 8 show impacts on job search expenditure in the last 30 days (winsorized). Controls include randomization block fixed effects, and prespecified baseline covariates (measured skills, self-reported skills, education, age, gender, employment, discount rate, and risk aversion). All monetary figures are reported in 2021 USD PPP. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A43: Heterogeneous Treatment Effects on Labor Market Outcomes by Confidence - Big Experiment

	Work quantity					Work quality			
	(1) Index	(2) Job offers (w)	(3) Worked month 1	(4) Worked month 2	(5) Worked last 7 days	(6) Index	(7) Earnings (w)	(8) Hourly wage (w)	(9) Written contract
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 A43 shows that treatment effects on labor market outcomes in the big experiment do not vary by baseline confidence levels.** Baseline confidence levels are defined as the average baseline deviation of tercile skill beliefs from measured terciles (across numeracy, communication, and concept formation) standardized to have control group standard deviation of one. Positive values indicate overconfidence and negative values indicate underconfidence. Columns 1 to 5 show impacts on employment quantity. Column 1 shows impacts on an Anderson (2008) index of the four employment quantity measures. Column 2 shows the impact on the number of job offers in the last 30 days (winsorized). Column 3 shows the impact of a dummy indicating any work for pay in month 1 after treatment. Column 4 shows the impact of a dummy indicating any work for pay in month 2 after treatment. Column 5 shows the impact on a dummy indicating any work for pay in the last seven days. Columns 6 to 9 show impacts on employment quality. Column 6 shows the impact on an Anderson (2008) index of the three employment quality measures. Column 7 shows the impact on earnings in the last seven days (winsorized). Column 8 shows the impact on hourly wages in the last seven days (winsorized). Column 9 shows the impact on a dummy indicating a written contract. Controls include randomization block fixed effects, and prespecified baseline covariates (measured skills, self-reported skills, education, age, gender, employment, discount rate, and risk aversion). 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 Willingness-to-Pay Appendix

As part of the job search workshop we elicited jobseekers' willingness-to-pay for three products: a document that reveals the expert-assessed skill requirements of jobs included in the job choice task (delivered after 30 days to avoid it influencing their search behavior), printed self-study materials for to improve numeracy skills, and printed self-study materials for to improve written communication skills. The self-study materials were taken from the "Mind the Gap!" series which targets grade 12 students and was developed by the Department of Basic Education in South Africa.

We elicit willingness-to-pay using multiple price lists with six items and announce that one randomly selected choice will be implemented in practice. The choice is always between receiving the product and receiving a changing monetary amount in airtime for their mobile phone (the "price"). For the skill document revealing the skill requirements of jobs the prices are: 0, 25, 50, 100, 150, and 200 Rand. For the self-study materials, we use the following prices: 0, 15, 30, 50, 75, and 100 Rand. We randomize whether jobseekers are asked about the highest or lowest price first. We force a unique switching point by ending the elicitation for an item at the first question they choose the item (for those starting with the highest price) or the monetary amount (for those starting with the lowest price) and assume that they would choose consistently on the omitted questions. We then randomly choose one question among the 18 questions and implement their choice in practice and announce this at the end of the session.

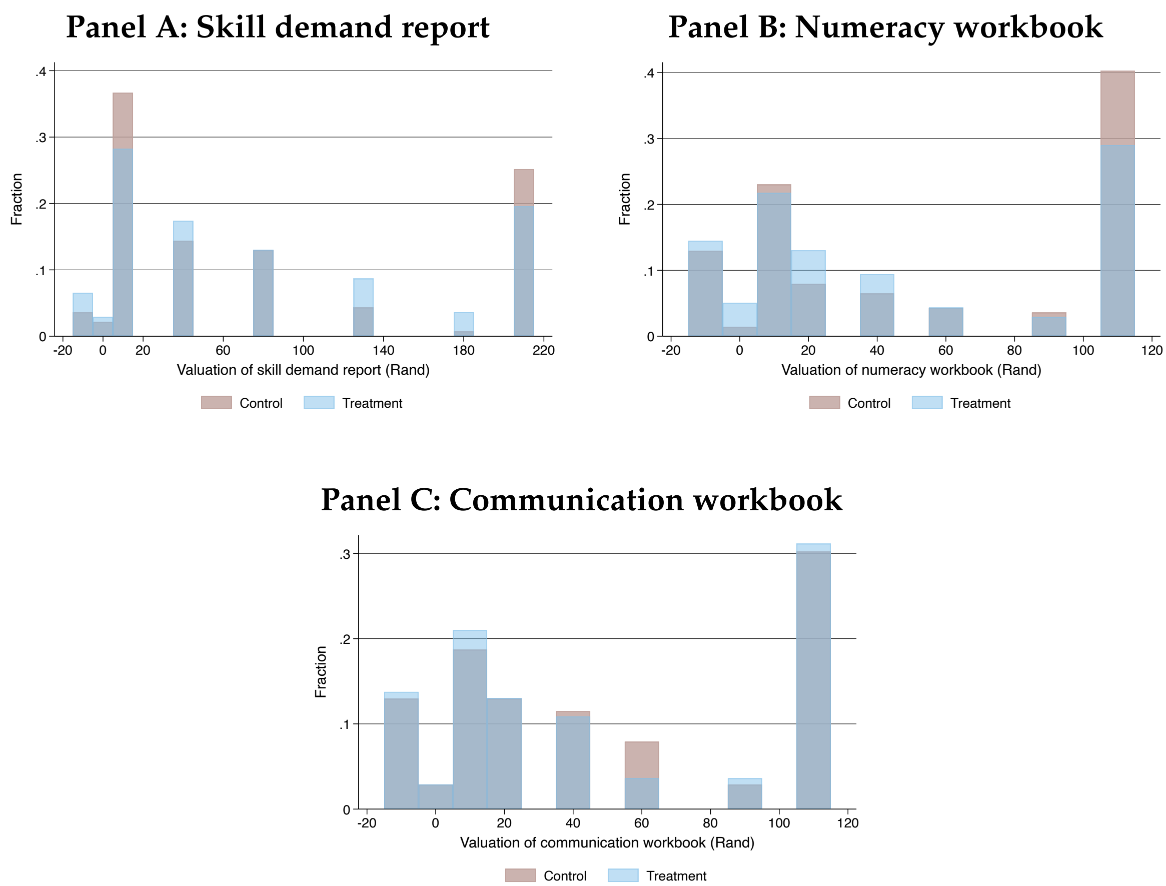
To ensure that jobseekers are familiar with the procedure, they complete a practice round where we elicit the jobseekers' willingness-to-pay for a bar of soap. If their choices are inconsistent, the enumerator explains the elicitation again and asks for the reason for the inconsistency to ensure that they understand the process.

To construct monetary willingness-to-pay measures, we assign jobseekers the average of the price they switched and the previous price. For example, if a jobseeker prefers the item at price zero but the money at price 25 Rand we assign them a willingness-to-pay of 12.5 Rand. For jobseekers who switch immediately, we assign -10 Rand (for those with negative willingness-to-pay) and 110 or 210 Rand (for those whose willingness-to-pay exceeds the maximum price). Given that there is substantially more mass on the upper end of the willingness-to-pay range, this conservative assumption is likely to lead us to underestimate the true willingness to pay. Figure A7 displays the distribution of willingness-to-pay measures by treatment status. Our treatment results are qualitatively robust to using ordered logit estimation rather than assigning monetary values.

Table A44 shows treatment effects on the three different willingness-to-pay measures.

The average willingness-to-pay for the skill report in the control group is 79 Rand (11.06 USD). It is 55 Rand (7.7 USD) for the numeracy materials and 48 Rand (6.72 USD) for the communication materials. We observe no treatment effect on jobseekers' willingness-to-pay for the skill demand report. We find negative treatment effects on jobseekers' willingness-to-pay for the numeracy workbook, driven by individuals with a comparative advantage in numeracy and initially misaligned comparative advantage beliefs. We find negative treatment effects on jobseekers' willingness-to-pay for the numeracy workbook, driven by individuals with a comparative advantage in numeracy and initially misaligned comparative advantage beliefs and jobseekers with a comparative advantage in communication and initially aligned comparative advantage beliefs. We find no average treatment effects on jobseekers' willingness-to-pay for the communication workbook with some indication that jobseekers with a comparative advantage in numeracy and initially aligned comparative advantage beliefs increase their willingness-to-pay. Note that the cell sizes for heterogeneity analysis are relatively small which limits the inference one should draw from these analyses.

Figure A7: Distribution of Willingness-to-pay Measures - Tight Experiment



Notes: **Figure A7 shows the distribution of elicited willingness-to-pay for three goods.** Panel A displays the distribution of willingness-to-pay for a report revealing the skill demand of jobs in the jobs choice task. Panel B displays the distribution of willingness-to-pay for a numeracy workbook. Panel C displays the distribution of willingness-to-pay for a communication workbook. All values are in 2022 South African Rands to illustrate the decisions as made by jobseekers. All other currency values in the paper are reported in 2021 USD purchasing power parity at the exchange rate 1 Rand = 0.140 USD PPP.

Table A44: Treatment Effects on Willingness-to-pay - Tight Experiment

	Info on skill requirements						Numeracy materials						Communication materials					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
	Pooled		Num. CA		Comm. CA		Pooled		Num. CA		Comm. CA		Pooled		Num. CA		Comm. CA	
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 comp adv 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 comp adv 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: Aligned comp adv belief (bl)		5.436		-5.231		5.898		-10.290		16.149		-13.042*		-1.094		43.251**		-8.125
		(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 A44 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. Controls include randomization block fixed effects, and prespecified baseline covariates (skill quintile dummies, age, gender, and dummies for having completed a university degree, other post-secondary education, and high school). Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

J 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 A45 shows evidence for two additional potential mechanisms in the big experiment: self-esteem and education investments. The latter are discussed in Section 6.2 jointly with willingness-to-pay results presented in Appendix Section I. Table A46 shows that average treatment effects on labor market outcomes are driven by jobseekers who attached their skill report with applications. Table A47 shows that the treatment did not induce congestion effects.

Table A45: Treatment Effects on Additional Mechanisms - Big Experiment

	Self-esteem				Education Investment		
	(1) SMS (z)	(2) SMS above med.	(3) Endline (z)	(4) Endline above med.	(5) Any	(6) Apprenticeship	(7) Formal
Treatment	0.003	0.012	0.007	0.012	0.017	0.006	0.014
	(0.009)	(0.015)	(0.022)	(0.015)	(0.012)	(0.005)	(0.011)
Observations	-0.000	0.483	0.000	0.471	0.224	0.036	0.185
Control mean	3334	3334	4206	4206	4205	4205	4205

Notes: **Table A45 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 measures, on a Likert scale, 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. Controls include randomization block fixed effects, and prespecified baseline covariates (measured skills, self-reported skills, education, age, gender, employment, discount rate, and risk aversion). Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A46: Heterogeneous Treatment Effects on Labor Market Outcomes by Skill Report Attachment to Applications - Big Experiment

	Work quantity					Work quality			
	(1) Index	(2) Job offers (w)	(3) Worked month 1	(4) Worked month 2	(5) Worked last 7 days	(6) Index	(7) Earnings (w)	(8) Hourly wage (w)	(9) Written contract
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)
Observations	-0.000	0.195	0.465	0.437	0.309	0.000	25.424	1.267	0.120
Number of clusters	3992	3933	3988	3991	3991	3993	3983	3971	3971

Notes: **Table A46 shows that treatment effects on labor market outcomes in the big experiment are not driven by jobseekers that report attaching their skill reports to at least one application.** *Attached report w. application* is dropped from the regression as no control individual received a report. Columns 1 to 5 show impacts on employment quantity. Column 1 shows impacts on an Anderson (2008) index of the four employment quantity measures. Column 2 shows the impact on the number of job offers in the last 30 days (winsorized). Column 3 shows the impact on a dummy indicating any work for pay in month 1 after treatment. Column 4 shows the impact on a dummy indicating any work for pay in month 2 after treatment. Column 5 shows the impact on a dummy indicating any work for pay in the last seven days. Columns 6 to 9 show impacts on employment quality. Column 6 shows the impact on an Anderson (2008) index of the three employment quality measures. Column 7 shows the impact on earnings in the last seven days (winsorized). Column 8 shows the impact on hourly wages in the last seven days (winsorized). Column 9 shows the impact on a dummy indicating a written contract. Controls include randomization block fixed effects, and prespecified baseline covariates (measured skills, self-reported skills, education, age, gender, employment, discount rate, and risk aversion). 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 A47: Treatment Effects on Concentration of Job Search - Tight Experiment

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

Notes: **Table A47 shows that the treatment (weakly) decreases the concentration of jobseekers' applications in the labor market.** 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. Columns 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 as the outcome of interest is at the labor market level averaged across jobseekers. Controls in columns 1 and 2 include randomization block fixed effects, and prespecified baseline covariates (measured skills, self-reported skills, education, age, gender, employment, discount rate, and risk aversion). Standard errors are clustered at the treatment-day level in columns 1 and 2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A48: Gender Differences in Skill Beliefs - Tight Experiment

	Female	Male	Δ	p($\Delta = 0$)	$\Delta(adjusted)$	p($\Delta(adjusted) = 0$)	N
Aligned comp. adv belief	0.50	0.46	-0.04	0.58	-0.01	0.83	278
Aligned comp. adv belief (assessment result)	0.56	0.48	-0.08	0.22	-0.03	0.46	278
Av. abs(beliefs - assessments) (general skill)	2.36	2.38	0.01	0.88	0.04	0.63	278
Av. abs(beliefs - assessments) (assessment result)	2.12	1.89	-0.23	0.03	-0.16	0.13	278
Fraction overcofident beliefs	0.62	0.59	-0.03	0.52	0.04	0.11	278
Fraction underercofident beliefs	0.15	0.20	0.05	0.15	0.01	0.79	278

Notes: **Table A48 shows that gender differences in beliefs in the tight experiment are small.** Adjusted gender differences control for age, education level, skill quintiles, and randomization block fixed effects.

K Gender

In this appendix we present findings related to gender. Table A48 displays unadjusted and adjusted gender differences in baseline skill beliefs in the tight experiment and Table A49 shows the same for the big experiment. We find that the gender differences in beliefs if anything are small. Table A50 and Table A51 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 A49: Gender Differences in Skill Beliefs - Big Experiment

	Female	Male	Δ	p($\Delta = 0$)	$\Delta(adjusted)$	p($\Delta(adjusted) = 0$)	N
Aligned comp. adv belief	0.19	0.23	0.04	0.01	0.03	0.03	4312
Fraction aligned beliefs	0.35	0.42	0.07	0.00	0.01	0.24	4378
Fraction overconfident beliefs	0.53	0.45	-0.09	0.00	0.00	0.75	4378
Fraction underconfident beliefs	0.11	0.13	0.01	0.06	-0.01	0.04	4378

Notes: **Table A49 shows that gender differences in beliefs in the big experiment are small**. Adjusted differences control for pre-specified covariates (measured skills, self-reported skills, education, age, employment, discount rate, and risk aversion).

Table A50: Heterogeneous Treatment Effects on Skill Beliefs by Gender - Tight Experiment

	Aligned com. adv. 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: Female	0.107	0.116**	0.146***	0.085**
	(0.065)	(0.047)	(0.049)	(0.032)
Control mean	0.475	0.475	0.183	0.183
Observations	278	278	278	278
Controls	No	Yes	No	Yes

Notes: **Table A50 shows that treatment effects on skill beliefs do not differ by gender.** 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. All specifications include randomization block fixed effects. Controls further include prespecified baseline covariates (skill quintile dummies, age, and dummies for having completed a university degree, other post-secondary education, and high school). Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A51: Heterogeneous Treatment Effects on Beliefs about Skills by Gender - Big Experiment

	Aligned comp. adv. 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		X		X

Notes: **Table A51 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. Even columns include prespecified baseline covariates (measured skills, self-reported skills, education, age, gender, employment, discount rate, and risk aversion). Standard errors clustered at the treatment-day level in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

L 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. As the exploratory analysis of the big experiment led us to the research question of the current paper, the corresponding pre-analysis plan did not include the specific research question of the current paper. We deviate from the plan in the following way:

- We restrict the sample to the *private* treatment arm. Participants in this arm received the unbranded report without any identifiers about their relative performance in terciles.
- We study additional outcomes on the topic of comparative advantage beliefs and aligned search. These outcomes were not prespecified.
- The treatment effect analysis was modified to reflect the change in the sample and to study heterogeneous effects by aligned comparative advantage.

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

- We restrict the sample to people who have a clear comparative advantage. In Appendix Tables A22 and A23 we show the main results of the paper for the full sample 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 A41.
- 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.