

Narrow Bracketing in Work Choices*

Francesco Fallucchi

Luxembourg Institute of Socio-Economic Research (LISER)[†]

Marc Kaufmann

Central European University (CEU)[‡]

January 13, 2021

Abstract

Many important economic outcomes result from cumulative effects of smaller choices, so the best outcomes require accounting for other choices at each decision point. We document narrow bracketing — the neglect of such accounting — in work choices in a pre-registered experiment on MTurk: bracketing changes average willingness to work by 13-28%. In our experiment, broad bracketing is so simple to implement that narrow bracketing cannot possibly be due to optimal conservation of cognitive resources, so it must be suboptimal. We jointly estimate disutility of work and bracketing, finding gender differences in convexity of disutility, but not in bracketing.

JEL Classifications: C91, D91, J01

1 Introduction

Many of the most important economic outcomes — such as savings, health, careers and relationships — result from countless interdependent decisions. In the work domain, such cumulative decision-making is bound to become even more important as remote work and online work platforms allowing more flexibility gain in popularity (Farrell et al., 2018, 2019). In order to obtain the best outcomes in these circumstances, a person must, in every single decision, take into account what she did or is going to do in her other decisions. Whenever a person fails to do so, she is engaging in

*We thank Gergő Hajdu, Botond Kőszegi, Tomy Lee, Sergey Lychagin, Ariada Muço, Daniele Nosenzo, Ádám Szeidl, seminar participants at CEU, University of Portsmouth, University of Trento, BEEN 2020 (University of Verona) and ESA 2020 conferences for helpful comments. The study was pre-registered in the AEA RCT Registry with ID 0003412 (see <https://doi.org/10.1257/rct.3412-4.1999999999999999>). The current data and code for the data analysis is currently hosted on GitHub at <https://github.com/MarcKaufmann/narrow-bracketing-in-work-choices>.

[†]francesco.fallucchi@liser.lu

[‡]Corresponding author: kaufmannm@ceu.edu

narrow bracketing (Read et al. (1999)). The literature has not investigated narrow bracketing over such economically essential work choices, and existing research on narrow bracketing — all of which is on monetary outcomes — does not conclusively establish whether behavior is due to a suboptimal neglect of choice interdependence or the optimal conservation of cognitive resources that would be needed to calculate combined outcomes.

In this paper, we document narrow bracketing in work choices in a pre-registered experiment with workers from Amazon Mechanical Turk. Across treatments that fix the set of combined outcomes but vary bracketing by increasing the active choice options and decreasing the passive endowment, we find that participants’ average willingness to work differs by 13-28%. In our experiment, broad bracketing is so simple to implement that narrow bracketing cannot possibly be due to an optimal conservation of cognitive resources, so that it must be suboptimal.

We define narrow and broad bracketing formally in Section 2. A person chooses two bundles X and Y from two arbitrary choice sets. We assume that utility depends on the sum $X + Y$ of the two bundles, which we call the total outcome of the choices. We say that a person brackets broadly if the chosen total outcome is the same whenever the set of possible total outcomes is the same, independent of how the outcome is split across the two choices. We say that a person brackets narrowly if they choose X as if Y was 0 and they choose Y as if X was 0 — they choose as if each choice alone determined total outcomes. We prove that we cannot identify whether a person brackets narrowly or broadly from any choices if and only if the person maximizes (context-independent) *additive* preferences. Additive preferences can be represented via utility functions that satisfy $U(X + Y) = U(X) + U(Y)$ for all bundles X and Y . So under additive preferences, bracketing is unidentified and irrelevant, since it does not affect choices. Depending on the richness of the choice sets we consider, additive preferences together with continuity assumptions imply expected utility, constant-absolute-risk-aversion, or linear preferences, extending and unifying results by Rabin and Weizsäcker (2009) and Ellis and Freeman (2020). We derive general identification strategies to directly test broad and narrow bracketing under non-additive preferences.

In Section 3, we describe the direct tests of both broad and narrow bracketing in work choices where workers on Amazon Mechanical Turk are paid to decode sequences of letters to numbers. We test broad bracketing — whether participants choose based on full outcomes — by comparing two treatments, BROAD and NARROW, that have identical total outcomes. In BROAD, participants choose between different amounts of *total* work and pay, such as between 15 sequences for a total payment of \$6.00 or 30 sequences for a total payment of \$7.00. In NARROW, participants choose between different amounts of *additional* work and pay, such as between 0 additional sequences for an extra \$4.00 or 15 additional sequences for an extra \$5.00, on top of the endowment of 15 sequences and the \$2.00 completion fee. The endowment is prominently displayed at the top of the choice page by noting that required sequences are “in addition to the sequences based on [participants’] choices” and that completing the study leads to “\$2.00 plus a bonus depending on [participants’] choices”. Since the total outcomes are the same, broad bracketing predicts that reservation wages — the lowest extra payments for which participants choose more work — in NARROW and BROAD are the same.

Similarly, we test narrow bracketing — that participants ignore the endowment — by comparing the NARROW to the LOW treatment, which differs from NARROW only by having no required (endowment) sequences. Since the amounts of *additional* work and pay are the same, narrow bracketing predicts that reservation wages for more work in LOW and NARROW treatments are the same.

Our results in Section 4 strongly reject broad bracketing and don't reject narrow bracketing in work choices, as does an estimation that combines the three treatments to directly estimate the *degree* of bracketing. We find that the average reservation wages in NARROW differ from BROAD by 13%-28%. For example, the average reservation wage in NARROW is \$2.07 in one choice scenario, while the corresponding average reservation wage in BROAD is \$2.88. On the other hand, the corresponding average reservation wage in LOW — requiring 15 fewer sequences than NARROW — is \$2.31, which is only significantly different from the NARROW treatment at the 10%-significance level.

In an additional treatment, PARTIAL, we test if there is a difference in bracketing in the money and work dimensions by broadly bracketing the work as in the BROAD, but not the payments. Comparing PARTIAL and BROAD, which differ only in the presentation of money, we reject that participants bracket money broadly; and comparing PARTIAL and NARROW, which differ only in the presentation of work, we reject broad bracketing in all situations in which we find that the identification assumption of non-additive preferences holds. Direct estimation of the degree of narrow bracketing provides similar estimates for bracketing in work alone to bracketing in work and money combined. The estimate for bracketing in work is merely noisier than for work and money combined, but rejects broad bracketing. Unlike direct comparisons of reservation wages, direct estimates of the degree of bracketing distinguish between the impact of bracketing and preferences on reservation wages by jointly estimating them. Joint estimation finds no gender differences in the degree of bracketing, showing that the large gender differences in reservation wages by treatment can be explained by gender differences in work preferences.

In a follow-up experiment aimed at reducing the impact of narrow bracketing, we run two treatments that differ from NARROW by describing additional sequences as being done “before” or “after” the 15 required tasks. Assuming convex disutility of work, we expected participants in the BEFORE treatment to think of the earlier, hence easier, tasks and participants in the AFTER treatment to think of later, hence harder, tasks — although both treatments may lead to broad bracketing by drawing attention to the required tasks. We therefore predicted that reservation wages in BEFORE would be more similar to narrow bracketing, while AFTER would be closer to broad bracketing. The reservation wages are directionally in line with this, but the differences are statistically significant only at the 10% level when they are. For example, in one setting they are \$2.70 in NARROW, \$2.72 in BEFORE, and \$2.96 in AFTER.

In Section 5, we identify narrow bracketing as a suboptimal mistake, ruling out alternative mechanisms. We consider models of optimal allocation of cognitive resources; differences of information; strategic or motivated bracketing; and preferences for bracketing. Since we keep information and the overall outcome sets constant across several

treatments, and since broad bracketing requires merely adding two workload and two payment numbers — noting that $15 + 15 = 30$ and $\$2 + \$4.50 = \$6.50$ — we conclude that narrow bracketing in our experiment is suboptimal. We also discuss how framing never provides an explanation for bracketing, since it can be applied broadly or narrowly.

We conclude with a brief discussion of the challenge of jointly estimating bracketing and preferences outside the lab, and ways to move towards joint estimation in field data. Taking into account narrow bracketing can improve estimates and predictions when policies affect both overall incentives and how people bracket these incentives.

2 Conceptual Framework of Narrow Bracketing

In this section, we formalize narrow and broad bracketing in choices, similar to Read et al. (1999), Rabin and Weizsäcker (2009), and Ellis and Freeman (2020). We show that narrow and broad bracketing are unidentified if a person maximizes context-independent additive preferences over all options, such as expected utility, constant-absolute-risk-aversion, or linear utility preferences depending on how rich the choice set is. Conditional on choices that are not compatible with such context-independent additive preferences, bracketing can be identified, and we describe our strategy for doing so.

Let us describe the setup. A person makes two choices, X from the choice set \mathcal{X} and Y from the choice set \mathcal{Y} , with $\mathcal{X}, \mathcal{Y} \subset \mathbb{X}$ for some choice space \mathbb{X} . We assume that \mathbb{X} is an additive space, so that for any X and Y in \mathbb{X} , we also have $X + Y \in \mathbb{X}$. These two choices together determine the full outcome $O = X + Y$. Then we can write the set of total outcomes from which the person actually chooses as $\mathcal{O} = \mathcal{X} + \mathcal{Y} = \{O : O = X + Y \text{ s.t. } X \in \mathcal{X}, Y \in \mathcal{Y}\}$.

We assume that the utility $U(\cdot)$ depends only on total outcomes – in the experiment will correspond to total effort and total earnings. Thus we can write $U(X, Y) = U(X + Y)$. Hence a broad bracketer combines all decisions and maximizes the following:

$$\max_{X \in \mathcal{X}, Y \in \mathcal{Y}} U(X + Y) = \max_{O \in \mathcal{O}} U(O) \tag{1}$$

A narrow bracketer instead is a person who makes each choice in isolation, by which we mean that they maximize some (possibly different) function $V(\cdot)$ over the *additional* bundles \mathcal{X} , rather than over total outcomes. Formally, they maximize the following:¹

$$\max_{X \in \mathcal{X}} V(X) \tag{2}$$

Based on this, we can now define *narrow* and *broad* bracketing.²

¹Recent work by Vorjohann (2020) suggests a specific form for $V(\cdot)$ in risky choices over multi-dimensional goods, with one reference point per dimension. If the reference points can depend on other choice sets, then this is not covered by our formalization.

²This definition extends naturally to more choices.

Definition 1. Denote by $O(\mathcal{X}, \mathcal{Y})$ the overall choice a person makes who faces two separate choices from choice sets \mathcal{X} and \mathcal{Y} . Let $C_1(\mathcal{X}, \mathcal{Y})$ be the choice made from \mathcal{X} when the other choice set is \mathcal{Y} , and $C_2(\mathcal{X}, \mathcal{Y})$ be the choice made from \mathcal{Y} when the other choice set is \mathcal{X} – so that $C_1(\mathcal{X}, \mathcal{Y}) = C_2(\mathcal{Y}, \mathcal{X})$. Then $O(\mathcal{X}, \mathcal{Y}) = C_1(\mathcal{X}, \mathcal{Y}) + C_2(\mathcal{X}, \mathcal{Y})$.

We say that a person brackets broadly if $O(\mathcal{X}, \mathcal{Y}) = O(\mathcal{X}', \mathcal{Y}')$ for all choice sets $\mathcal{X}, \mathcal{X}', \mathcal{Y}$, and \mathcal{Y} with $\mathcal{X} + \mathcal{Y} = \mathcal{X}' + \mathcal{Y}'$.

We say that a person brackets narrowly if $C_1(\mathcal{X}, \mathcal{Y}) = C_1(\mathcal{X}, \mathcal{Y}')$ for all choices \mathcal{Y} and \mathcal{Y}' .

In our experiment, the second 'choice' set $\mathcal{Y} = \{Y\}$ is a singleton – it is an endowment of required work and money that gets added to their choice $X \in \mathcal{X}$. We describe our specific treatments in Section 3, but conceptually we test for broad and narrow bracketing as follows:³

Hypothesis (Broad Bracketing). Behavior is consistent with broad bracketing if $O(\mathcal{X}, \mathcal{Y}) = O(\mathcal{X}', \mathcal{Y}')$ for all $\mathcal{X}, \mathcal{X}'$, \mathcal{Y} , and \mathcal{Y}' s.t. $\mathcal{X} + \mathcal{Y} = \mathcal{X}' + \mathcal{Y}'$.

Hypothesis (Narrow Bracketing). Behavior is consistent with narrow bracketing if $C_1(\mathcal{X}, \mathcal{Y}) = C_1(\mathcal{X}, \mathcal{Y}')$ for all \mathcal{Y} and \mathcal{Y}' .

We now turn to the question of identifiability, since it is possible for behavior to be consistent with both broad and narrow bracketing. Our proposition 1 characterizes those preferences that are consistent with both broad and narrow bracketing: preferences for which narrow bracketing and broad bracketing *always* lead to identical choices, no matter which choice sets we offer the person.

Proposition 1 (Equivalence of Narrow and Broad Bracketing). Suppose that a person's choices are consistent with both narrow and broad bracketing under all possible choices. Suppose further that there is some numeraire, which we call money. Then (under appropriate continuity assumptions)

1. preferences are context-independent
2. the money-metric $M(\cdot)$ is additive over all options: $M(X + Y) = M(X) + M(Y)$ for all bundles X and Y that we consider. This implies (depending on choice domain):
 - Choices over risk: expected utility
 - Wealth (continuous) choices over risk: constant-absolute-risk-aversion (CARA) preferences
 - Multiple goods (including multiple periods): linear utility

Proof. See Appendix A. □

³These hypotheses naturally extend to more general choice sets \mathcal{Y} .

The proposition highlights the central role played by additivity across preference domains, that $M(X + Y) = M(X) + M(Y)$ for all X, Y , where $M(\cdot)$ is the money metric. Notice that this assumption is usually much stronger than, but implies, additive separability, which requires that additivity holds for bundles across different goods, periods, or dimensions, but not necessarily within each good. In fact, it is well-known that any real-valued continuous function satisfying this additive functional equation (called Cauchy’s functional equation) must be linear, showing how strong this condition is. Two special cases of this result are the result by Rabin and Weizsäcker (2009) who show that CARA-preferences imply no costs to bracketing under expected utility; and the result (contemporaneous with ours) by Ellis and Freeman (2020) that bracketing is unidentified under linear utility over multiple goods.

Our proposition unifies and extends these results as special instances of a general additivity property that holds for arbitrary choice settings (assuming a numeraire). Under different constraints on the choice sets that we observe, additivity of the money metric over these choice sets will lead to different restrictions on the underlying preferences. For example, if the setting is such that people make two choices, one over work and money, and one over food and money, then the first choice determines all the work and the second choice determines all the food. In that case, additivity in such a choice setting implies that the utility can be represented via a utility function that is separable, but not linear, in work and food, and linear in money.

In our experiment with two goods (work and money) our identifying assumption is that either we have context-dependent preferences, or we have non-linear preferences in money and work. To illustrate why linear utilities are a problem, suppose a person is always willing to do exactly one extra task for one extra dollar or more, but never for less. Then their choice from one choice set is not affected by how many tasks they choose in the other choice set, so broad and narrow bracketing lead to identical choices. Both convexity of effort as well as comparison effects (such as Tversky and Simonson (1993); Kőszegi and Szeidl (2012); Bordalo et al. (2013); Bushong et al. (2020)) will ensure that the identification assumption holds. We primarily target increasing tiredness by choosing a reasonably large amounts of difference in work.

3 Experimental Design and Procedures

We conducted two sets of pre-registered experiments online, using the software Lioness (Giamattei et al., 2020). In this section, we first describe our main experiment conducted from December 2019 to January 2020, and our hypotheses for bracketing in real effort choices with detailed instructions in Appendix B. We then highlight some implementation details that were not fixed in or mildly changed from our pre-registration. We finally describe our follow-up experiment conducted in March and August 2020, aimed at reducing the potential impact of narrow bracketing.

3.1 Task and Main Experiment Description

The main experiment consists of four parts.

Part 1: Tutorial Subjects familiarize themselves with the task by completing three tasks correctly which consist of decoding a sequence of twelve letters into numbers, as seen in Figure 1.

Encryption task

Decode sequences into numbers until you got 10 right.

Correctly decoded text sequences: 0

Decoding table

letter	u	t	f	c	g	n	v	q	a	l
corresponds to number	0	1	2	3	4	5	6	7	8	9

Text sequence:

Your answer (enter only numbers):

167757642404

Figure 1: Decoding task

Every sequence is a new table showing a mapping of ten randomly chosen letters to the numbers between 0 and 9. After every attempt we generate a new table, independently of whether the answer was correct or not. This makes it harder for subjects to learn the task than more commonly used encryption or typing tasks (e.g. Erkal et al., 2011; De Quidt et al., 2017; De Quidt, 2018). This should increase the convexity of costs, which is our primary channel for satisfying our identifiability assumptions.

Part 2: Elicit Tediousness We elicit participants' perceived tediousness of the task on a scale from 1 ("not tedious at all") to 10 ("extremely tedious"), providing us a control common to all treatments.

Part 3: Elicit Reservation Wages by Treatment We elicit participants' reservation wage for a high-work option that requires 15 more tasks than the low-work option, depending on treatments. The four treatments (described below) vary participants' endowment of tasks and money and how this endowment is presented. There are two scenarios with Scenario 1 always presented before Scenario 2. We elicit one reservation wage for each scenario through an incentivized price list task, where participants choose to accept or reject the extra work for a list of extra reservation wages between \$0.25 and \$4.00 in \$0.25 increments. The extra wages are in addition to the \$4.00 participants receive for the alternative workload.⁴ We remind subjects that a single binding choice from one of the two scenarios will be chosen at random.

Part 4: Complete Tasks We randomly determine the binding choice and inform subjects about the total payment and number of sequences to decode. Subjects can complete the tasks without time constraints. We then ask a short demographic questionnaire and display a summary of total earnings.

3.2 Main Treatments

Our main experiment consists of four treatments: BROAD, NARROW, LOW, and PARTIAL. Let us illustrate implementation and wording of treatments using Scenario 1. For the three treatments BROAD, NARROW, and PARTIAL, the total outcomes – choice list option plus endowment – are identical; but some of the work and money is shifted from the choice list options to the endowment. Treatment LOW instead has identical choice list wording to NARROW, but no endowment of work. The following list shows how the wording of choice list items changes between treatments for scenario 1. Payments M always start from \$4.25 and go up to \$8, and participants choose between OPTION A and OPTION B:

- BROAD: no endowment, $Y = (0 \text{ sequences}, \$0)$
 - OPTION A “15 sequences for a total HIT payment of \$6”
 - OPTION B “30 sequences for a total HIT payment of $\$M + \2 ”
- NARROW: both money and work endowment, $Y = (15 \text{ sequences}, \$2)$ ⁵
 - OPTION A “0 extra sequences for extra \$4”
 - OPTION B “15 extra sequences for extra $\$M$ ”
- LOW: like NARROW, but without any work endowment, $Y = (0 \text{ sequences}, \$2)$

⁴Details about the individual choices for each treatment can be found in Appendix B.

⁵In earlier sessions of this treatment participants were informed about the endowment of sequences to decode before the choice screen. This information has been dropped in successive sessions. Our main results do not depend on which sample of NARROW (early, late, or pooled) we use.

Table 1: Treatment Summary for both scenarios: the choice set \mathcal{X} gives the number of sequences and amount of money of the options in the price list, the baseline Y gives the (required) baseline number of sequences and money displayed at the top of the choice page, but not integrated into the options in the price list. The value of M goes from \$4.25 to \$8.00 in increments of \$0.25.

TREATMENTS	Choice Set S1 (\mathcal{X}_1)	Endowment S1 (Y_1)	Full Outcome S1
BROAD	(30, $\$(M + 2)$) vs (15, \$6)	(0, 0)	(30, $\$(M + 2)$) vs (15, \$6)
NARROW	(15, $\$M$) vs (0, \$4)	(15, \$2)	(30, $\$(M + 2)$) vs (15, \$6)
LOW	(15, $\$M$) vs (0, \$4)	(0, \$2)	(15, $\$(M + 2)$) vs (0, \$6)
PARTIAL	(30, $\$M$) vs (15, \$4)	(0, \$2)	(30, $\$(M + 2)$) vs (15, \$6)
TREATMENTS	Choice Set S2 (\mathcal{X}_2)	Endowment S2 (Y_2)	Full Outcome S2
BROAD	(45, $\$(M + 2)$) vs (30, \$6)	(0, 0)	(45, $\$(M + 2)$) vs (30, \$6)
NARROW	(30, $\$M$) vs (15, \$4)	(15, \$2)	(45, $\$(M + 2)$) vs (30, \$6)
LOW	(30, $\$M$) vs (15, \$4)	(0, \$2)	(30, $\$(M + 2)$) vs (15, \$6)
PARTIAL	(45, $\$M$) vs (30, \$4)	(0, \$2)	(45, $\$(M + 2)$) vs (30, \$6)

- OPTION A “0 extra sequences for extra \$4”
- OPTION B “15 extra sequences for extra $\$M$ ”
- PARTIAL: no work endowment, $Y = (0 \text{ sequences}, \$2)$
 - OPTION A “15 sequences for extra \$4”
 - OPTION B “30 sequences for extra $\$M$ ”

Table 1 summarises the choice, endowment, and overall outcome for each treatment and each scenario. In the notation of Section 2, we denote by \mathcal{X} the set of options as presented in the isolated choice; by Y the endowment of tasks and money that will be added to the choice made from \mathcal{X} ; and by $\mathcal{O} = \mathcal{X} + \{Y\}$ the actual set of outcomes. All of this information is visible to the participants on their choice page.

In the BROAD treatment, participants are offered the payment $M + 2$, which includes their participation fee. In all other treatments, the payment offered is M and is on top of the \$2.00 participation fee. Across all treatments and scenarios M ranges from \$4.25 to \$8.00 in steps of \$0.25. Participants in the treatments NARROW, PARTIAL, and BROAD treatments must do at least 15 tasks no matter what they choose, which is why we set their lowest additional payment to \$4.00, and why higher work load receives at least \$4.25. Since LOW has identical choice set \mathcal{X} as NARROW, participants in the LOW treatment can receive \$4.00 for no work in Scenario 1.

Table 1 makes clear that BROAD, PARTIAL, and NARROW have identical outcome sets (see Full Outcome column). Broad bracketing therefore predicts identical choices across these treatments. Treatment LOW requires 15 fewer sequences in each scenario, but shares identical choice sets \mathcal{X} with NARROW (see Choice Set column). Narrow bracketing therefore predicts identical choices for NARROW and LOW. Denoting by m_T the average reservation price elicited in treatment T , we get the following hypotheses:

Hypothesis 1 (Broad Bracketing). *Behavior is consistent with broad bracketing if $m_{BROAD} = m_{PARTIAL} = m_{NARROW}$.*

Hypothesis 2 (Narrow Bracketing). *Behavior is consistent with narrow bracketing if $m_{NARROW} = m_{LOW}$.*

As we discussed in Section 2, we need an identification assumption, which in our case implies that $m_{LOW} \neq m_{BROAD}$. If they are equal, we cannot rule out linear (population) preferences: that participants are as willing to do 15 additional sequences on top of 0 sequences, as they are to do them on top of 15 sequences.

3.3 Main Experiment: Summary Statistics

We recruit in total 929 subjects on Amazon Mechanical Turk between the end of December 2019 and the beginning of January 2020. In Table 2 we report a summary of demographics. Of the subjects recruited, 162 did not complete the experiment. While it might itself indicate narrow bracketing, we see no evidence of differential attrition by treatment across different parts of the study (see Table 9 in Appendix C). Across all treatments, between half and two-thirds of attritors drop before completing the practice tasks, the rest essentially drop out after finding out how many tasks they have to do in total. Treatments are similar in terms of gender composition (χ^2 p-value: 0.91), while they are slightly older in the BROAD treatment compared to other treatments (37.8 years vs 35.0-36.0, χ^2 p-value: 0.01). Finally, individuals rate the task on average as 7.33-7.54 out of 10 in tediousness, which does not significantly vary across treatments (χ^2 p-value: 0.86). Roughly 25% of the choices made in the choice-list are inconsistent: in few cases subjects make only one inconsistent choice, while in other cases choices are inconsistent throughout the list of wages offered. In our main analysis we drop a scenario if individuals make one or more inconsistent choices in it. To detect an effect size of 0.40, that we consider economically meaningful, at a 5% level of significance with 90% power, we would need 174 observations in NARROW and 116 in PARTIAL and BROAD treatments. The number of observations collected with consistent choices are above these thresholds and therefore considered sufficient for our treatments comparisons, although as the discussion on identification makes clear, the effect size decreases as preferences become more linear.⁶ Participants earned \$7.30 on average, for an average working time of 35 minutes.

3.4 Main Experiment: Implementation Details

The careful reader will have noticed that we have more participants in the NARROW treatment than in other treatments. This is because in our initial version of the NARROW treatment, we mistakenly informed participants of their endowment on the page *before* the first choice, which we fixed in later version by informing participants of their endowment on the first choice page only. Early information provision might lead to some (fast) reference effects not

⁶This highlights the benefits of running a pilot estimating the effect size by running a small sample of LOW against BROAD – various budgetary constraints kept us from doing so.

Table 2: Summary statistics for main treatments.

	BROAD	NARROW	LOW	PARTIAL	p-value
Participants	200	320	196	213	
Attrition	18%	20.3%	13.3%	16.4%	0.22
Final Participants	164	255	170	178	
Share Female	0.4	0.38	0.4	0.37	0.91
Age	37.8	35	35.1	36	0.01
Tediousness	7.54	7.45	7.33	7.54	0.86
Inconsistent Choices					
Scenario 1	34	60	40	47	0.59
Scenario 2	30	59	37	47	0.33

present in other treatments. In order to have enough observations of each subtreatment, we collected more data. As we show in Appendix C, our results are robust to choosing the early, late, or pooled NARROW sample.

Moreover, in our pre-registration we had planned on using three treatments with equal total outcomes, one with no baseline tasks, one with 8 baseline tasks, and one with 16. We realized however that the middle treatment of 8 baseline tasks would add little value and cost us more, so we dropped it and went with 15 instead of 16 tasks.⁷ When fixing the NARROW treatment, we realized however that instead of having an intermediate number of tasks, we could broadly frame the task but not the money dimension, which led to PARTIAL as a replacement for the intermediate treatment. This PARTIAL treatment allows us to test directly for narrow bracketing in the money or work dimension separately which seems a useful addition. While in line with our pre-registration, it is worth pointing out that this treatment was not pre-registered exactly as run.

Finally, because we added PARTIAL and changed NARROW later, these observations are not balanced against observations in LOW, BROAD, and early NARROW treatments: most were collected in sessions focusing only on these treatments. We don't consider this a major issue, since restricting our analysis to the early balanced data does not meaningfully affect our results.

3.5 Follow-Up Experiment Description

In our follow-up study we explored if we can reduce narrow bracketing by making the increasing costs of the additional tasks more salient through a different presentation of choices. We design two new treatments, BEFORE and AFTER, that are identical to the NARROW treatment, except that we describe the effort choice to participants as extra sequences to decode before or after the mandatory sequences. We reasoned that thinking about doing additional tasks "after" the

⁷For the same reason we decided to restrict ourselves to 2 scenarios rather than 5 scenarios as suggested in the pre-registration. This is enough to see if participants have convex disutility, and can span the same range of workloads, while avoiding intermediate effort levels which would lower effect size due to smaller changes in reservation wages.

required ones would lead participants to think about higher marginal disutility.⁸

Concretely, we presented choices as follows: OPTION A "0 additional sequences before (after) the 15 required for an extra \$4" versus OPTION B "15 additional sequences before (after) the 15 required for an extra \$X" with X starting from \$4.25 and up to \$8. In total, 302 participants were recruited and started the HIT.⁹ We report in Table 10 in Appendix C the summary statistics, compared to the earlier NARROW sessions. We observe a similar attrition rate to the other treatments. Overall, the composition of the sample in terms of demographics and perceived tediousness of the task is similar to the main treatments. However, we find more inconsistent choices.¹⁰

This leads to the following natural hypotheses:

Hypothesis 3 (Debiasing through highlighting required tasks alone). *Drawing attention to extra tasks reduces narrow bracketing if $m_{BEFORE} \neq m_{NARROW}$.*

Intuitively, the BEFORE phrasing might make participants realize the relevance of the required tasks. This may change their behavior, if they understand that they should add the endowment to their choice and do so.

Hypothesis 4 (Debiasing through highlighting convex cost of additional tasks). *Drawing attention to the (assumed) convexity of costs reduces narrow bracketing if $m_{AFTER} \neq m_{NARROW}$.*

Intuitively, the AFTER phrasing might make people realize that they have to do tasks 16 through 30, rather than 1 through 15, even if they do not realize that they should broadly bracket. So, if it draws attention in this way, then it should lead to a larger change away from NARROW. In both cases, full debiasing requires that the reservation wage equals BROAD — or PARTIAL if it only debiases in the work dimension.

These hypotheses are however not as tightly linked to the theory as our main hypothesis, as we need to make additional assumptions about the shape of the preferences, as well as about our treatments drawing sufficient attention to these features of the preferences.

4 Results

In the analysis below we analyse the reservation wages across treatments: the smallest extra wage for which subjects prefer OPTION B over OPTION A, where OPTION B always requires decoding correctly 15 extra sequences compared to OPTION A. The extra payments start at \$0.25, if a subject always accepts the extra work, and increases in \$0.25 increments to \$4.00. If a subject never accepts the extra work, we code the reservation wage as \$4.25.

⁸Note however, that these follow-up treatments were collected primarily after COVID-19 induced lockdowns had been put in place.

⁹The new study was conducted in two sessions in March and August 2020.

¹⁰These features may be explained by recent evidence about the effects of the COVID-19 pandemics on the composition of the pool of Mturkers. Moss et al. (2020) reports that the demographic composition of MTurkers did not change with the pandemics. However, Arechar and Rand (2020) find that on average Mturkers became less attentive. Our findings are in line with both studies, making us cautious about the comparability of treatments. For this reason we focus exclusively on individuals whose choices are consistent.

First, we report our main results which test and estimate broad and narrow bracketing directly. We then report a non-preregistered heterogeneity analysis, as well our pre-registered follow-up study collected after COVID-19 induced lockdowns, which might affect comparisons with the main treatments.

4.1 Main Results

Table 3: Means of main treatments by scenario

Treatment	Res. Wage	Std Dev	% upper bound	N
Scenario 1				
BROAD	2.88	1.24	24%	137
NARROW	2.07	1.29	12%	207
LOW	2.31	1.33	18%	131
PARTIAL	2.50	1.28	15%	137
Scenario 2				
BROAD	2.99	1.29	33%	141
NARROW	2.70	1.29	20%	205
LOW	2.74	1.32	25%	135
PARTIAL	2.43	1.32	18%	138

Result 1. *We reject Hypothesis 1 that individuals bracket work decisions broadly.*

Both Tables 3 and 4 resoundingly reject broad bracketing as per Hypothesis 1 since reservation wages are different for NARROW, BROAD, and PARTIAL, despite equal outcome sets.¹¹ For example, the average extra reservation wage of the BROAD treatment in Scenario 1 is \$2.88 compared to \$2.07 in NARROW and \$2.50 in PARTIAL. Based on Wilcoxon rank-sum tests reported in Table 4, we reject that BROAD and NARROW are equal with a p-value of less than 0.001 and that BROAD and PARTIAL are equal with a p-value of 0.012 in Scenario 1.¹² Similar results hold for Scenario 2. Notice that the reservation wage of Scenario 2 slightly decreases (statistically remains constant) in PARTIAL, which we attribute to either genuine non-convex disutility or, more likely, to some narrowly bracketed framing effect possibly due to the order of choices.

Result 2. *We fail to reject Hypothesis 2 that individuals bracket work decisions narrowly.*

The same comparisons of mean reservation wages via Wilcoxon rank-sum tests fail to reject narrow bracketing, where we compare means in LOW and NARROW, which have identical choice sets, but different endowments. In scenario 1, the Wilcoxon rank-sum test fails to reject that the mean of \$2.30 in LOW is different from the mean of

¹¹See C.4 in Appendix C for bar plots and kernel density plots of the raw reservation wage data by treatment and scenario.

¹²Since there are ties in our data – there are multiple people with the same reservation wage – the test we compute uses a normal approximation to the test statistic (which does not assume that the data is normal, but that the test statistic itself is normal).

Table 4: Between-treatment p-values for main treatments based on two-sided Wilcoxon rank-sum tests, treating each individual in each scenario as a single independent observation.

Treatments	BROAD	NARROW	LOW
Scenario 1			
NARROW	< 0.001		
LOW	< 0.001	0.117	
PARTIAL	0.012	0.002	0.209
Scenario 2			
NARROW	0.023		
LOW	0.108	0.704	
PARTIAL	< 0.001	0.064	0.054

Table 5: Estimates of κ when the broader option is BROAD or PARTIAL respectively. Narrow option is always LOW, κ estimates the convex combination between the broader and the narrow option, with $\kappa = 0$ indicating equality with the broader option.

Terms	BROAD All	BROAD Male	BROAD Female	PARTIAL All	PARTIAL Male	PARTIAL Female
S1 broader (B_1)	2.89 (0.11)	2.85 (0.14)	2.94 (0.17)	2.52 (0.11)	2.54 (0.14)	2.46 (0.18)
S2 broader (B_2)	2.98 (0.11)	3.09 (0.14)	2.82 (0.17)	2.46 (0.11)	2.53 (0.14)	2.33 (0.18)
S1 narrow (N_1)	2.3 (0.11)	2.43 (0.13)	2.03 (0.18)	2.24 (0.1)	2.36 (0.11)	2 (0.18)
S2 narrow (N_2)	2.77 (0.07)	2.86 (0.09)	2.61 (0.11)	2.64 (0.08)	2.73 (0.12)	2.52 (0.11)
κ	1.38 (0.29)	1.42 (0.5)	1.25 (0.3)	1.53 (0.61)	1.45 (0.9)	1.4 (0.61)
Observations	956	558	392	953	572	371

\$2.07 in NARROW (p-value of 0.117), and in scenario 2 fails to reject that \$2.74 in LOW is different from \$2.70 in NARROW (p-value of 0.704) – see Table 3 for the means and Table 4 for the test p-values.¹³

One problem with comparing means by scenario is that this allows for multiple tests. While we only have two scenarios that both lead to identical results for BROAD compared to the other treatments, we would like a more generally valid test. In our pre-registration, we specified running a linear regression with a single treatment effect averaged across scenarios. This essentially averages the reservation wages across scenarios by treatment and compares them. While this test rejects that BROAD is equal to NARROW or PARTIAL — see Table 16 in Appendix C — it is conceptually the wrong test, as is illustrated by its failure to reject that NARROW and PARTIAL are equal. The reservation wage in PARTIAL is once above and once below the reservation wage in NARROW, which leads to similar

¹³See Appendix C.2 and C.3 for the results when restricting the NARROW treatment to the early or late observations only — those providing information on endowment on or before the choice page.

average reservation wages when averaged over both scenarios. The reason this is conceptually the wrong test is made clear by Hypotheses 2 and 1, which apply at the level of full choices, which means at the level of scenarios. At the scenario-level, PARTIAL and NARROW indeed are statistically significantly different in Scenario 1, but not in Scenario 2.

The general test we propose is to estimate the average *degree* of narrow bracketing across scenarios. Intuitively speaking, instead of averaging the reservation wages first and then estimating narrow bracketing, we would like to estimate narrow bracketing per scenario and then average this. One way to do this is to realize that we can write the reservation wage in NARROW as a convex combination as follows:

$$m_{NARROW} = (1 - \kappa) \cdot m_{BROAD} + \kappa \cdot m_{NARROW}$$

where κ is the degree of narrow bracketing. If $\kappa = 0$ we have broad bracketing, if $\kappa = 1$ we have narrow bracketing. Based on this we run the following non-linear regression:

$$Y_{i,s,t} = B_s \mathbb{1}(t = BROAD) + N_s \mathbb{1}(t = LOW) + ((1 - \kappa_B) \cdot B_s + \kappa_B \cdot N_s) \mathbb{1}(t = NARROW) + \varepsilon_{i,s,t}$$

$$Y_{i,s,t} = B_s \mathbb{1}(t = PARTIAL) + N_s \mathbb{1}(t = LOW) + ((1 - \kappa_P) \cdot B_s + \kappa_P \cdot N_s) \mathbb{1}(t = NARROW) + \varepsilon_{i,s,t}$$

where i is the individual, s the scenario, and t the treatment. If we thought that there is a fraction κ_B of the population that brackets exactly narrowly, and the rest brackets exactly broadly, then the above would identify this κ_B . Similarly, κ_P captures the fraction of people who bracket work narrowly, and we believe everyone brackets money narrowly. This estimation allows us to combine all scenarios, and provides the average degree of narrow bracketing. It also allows κ to fall outside the interval $[0, 1]$, which can happen under non-convex disutility or because people engage in other types of bracketing than full broad or narrow bracketing.¹⁴

We estimate the above regression via non-linear least squares (NLS, Gallant (1975)) and report the results in Table 5, where we use both the BROAD and the PARTIAL as the broader treatment: thus we estimate κ for bracketing over money and work (BROAD) as well as just work (PARTIAL). The main take-away is that we reject equality of PARTIAL and NARROW (column 4, $\kappa = 1.53$, standard error of 0.61), because it takes the reversed ranking of reservation wages into account. In order for NLS to provide credible standard errors however, we need the means between the broader treatment and LOW to be sufficiently different across all scenarios jointly — which is our identification assumption again! Our identification assumption holds for BROAD compared to LOW (since it holds in Scenario 1), but is unlikely to hold for PARTIAL compared to LOW, given how similar they are. This is reflected in the noisy (and un-

¹⁴As an example of non-convex disutility, suppose that the tasks 1 to 10 are unpleasant, that tasks 11 to 20 are easy due to warm-up, while tasks 21 to 30 are really unpleasant. Then a person who has to do 15 tasks but is asked like in NARROW treatment to do 15 vs 0 additional tasks may well think that these 15 additional tasks are *easier* than the first 15 tasks if they take their baseline into account only partially. Thus they would be more willing than participants in LOW who in turn are more willing to work than participants in BROAD, hence κ would fall outside the $[0, 1]$ interval.

trustworthy) standard errors for the PARTIAL treatment in column 4. Despite the standard errors of the NLS-estimate probably being untrustworthy, the estimate of κ strongly favors narrow over broad bracketing.¹⁵

This approach seems the most conceptually sound to us — although there are probably better econometric estimation strategies than NLS. In addition to avoiding multiple tests, it allows for joint identification of bracketing and preferences, the benefit of which we highlight in the next part.

Table 6: Determinants of bracketing across scenarios via Tobit Regressions (right censored at \$4.25).

	<i>Dependent variable:</i>			
	BROAD	NARROW	LOW	PARTIAL
	(1)	(2)	(3)	(4)
Scenario 2	0.20 (0.20)	0.72 (0.15)	0.52 (0.20)	-0.04 (0.19)
Female	0.34 (0.21)	0.43 (0.15)	0.33 (0.20)	0.07 (0.19)
Age	0.02 (0.01)	0.01 (0.01)	-0.004 (0.01)	-0.01 (0.01)
Tediousness	0.12 (0.04)	0.05 (0.03)	0.12 (0.04)	0.08 (0.04)
Constant	1.31 (0.52)	1.04 (0.39)	1.53 (0.48)	2.38 (0.46)
Observations	278	408	264	271

Note: Standard errors in parentheses.

4.2 Heterogeneity, Joint Identification, and Gender Differences

We further explore the determinants of work decisions across treatments with a set of (non pre-registered) estimates of κ by gender, as well as Tobit regressions. The gender-specific estimates of κ for BROAD versus LOW show that there is no major difference in the degree of narrow bracketing as measured by κ . This holds despite the fact that the difference in reservation wage for women in scenario 1 is \$0.91 in Scenario 1 and \$0.21 in Scenario 2 (the difference between S1 broader and S1 narrow in Table 5), while these differences are only \$0.42 and \$0.23 for men. This highlights our next result:

Result 3. *We see no difference in bracketing by gender, but narrow bracketing is costlier for women than men due to*

¹⁵This remains true when we restrict observations of NARROW to late observations where participants find out their baseline on their first choice page only, but the standard errors are even larger, both due to smaller sample size and due to the identification assumption failing badly for males in this subsample.

higher convexity of disutility.

This highlights the need for direct estimation of the degree of narrow bracketing via κ , which jointly estimates the preferences – here the change in reservation wage between scenarios – and bracketing. The reason is that the more linear the preferences are, the more similar behavior under narrow and broad bracketing is, which means that the reservation wage changes less even if the person brackets narrowly. Since in our dataset women have a larger change in their reservation wage as measured by the difference between LOW and BROAD, they pay a higher cost from narrow bracketing. If instead we focused on whether the difference in reservation wage was significant and large as a proxy for narrow bracketing, we would wrongly conclude that women bracket more narrowly than men, another illustration why focusing on reservation wages rather than κ is conceptually unhelpful.

In the Tobit regressions, we regress, separately for each treatment, the reservation wages choices on age, gender and the self-report of perceived tediousness, controlling for the Scenarios. In particular we are interested to check if there are within treatment gender differences, as already found in different settings (e.g. Koch and Nafziger, 2019). Since we have a substantial fraction of choices by individuals that are never willing to decode extra sequences, our data are right-censored at \$4.25. We report the estimation results in Table 6.

The Tobit regressions in Table 6 tell a similar story. Age has no clear effect on the results, while we observe a positive correlation between tediousness and the reservation wage in all treatments, although this is not always significant.

Let us conclude this part by highlighting a field setting where such joint identification is almost possible and is likely to matter. In a nice study of monopsony power, Dube et al. (2020) measure labor supply elasticities on MTurk. In their estimation of employers’ market power, they assume myopic workers who essentially narrowly bracket. While this does not affect the reduced-form measure of labor supply elasticity, if we treat this estimate as if it was all driven by preferences rather than bracketing, this may provide the wrong predictions to responses in labor market policies and designs. Their data does not allow for our simple strategy of joint identification, since they only have a single observation per worker. But in future data collections, it should be straightforward to collect multiple observations per worker or to make and test additional assumptions about population-level bracketing and preferences to allow joint estimation.

4.3 Study 2: an Attempt to Debias

The treatments BEFORE and AFTER are identical to the NARROW treatment, except for describing additional sequences as “additional sequences before” or “additional sequences after” the 15 required tasks. We report the means by Scenario and by treatment, and the Wilcoxon tests in Tables 7 and 8 similarly to our main treatments. We compare BEFORE and AFTER with the NARROW treatment with which they are identical *except* for the additional highlight-

ing of the tasks as “before” or “after” the baseline tasks. In both BEFORE and AFTER the extra reservation wage is higher than in NARROW, but in both cases this difference is not statistically significant ($p - values > 0.091$). In Appendix C, we however show that the AFTER treatment is statistically significantly different from NARROW when we limit ourselves to those observations in NARROW that received their information about baseline on the first choice page only, which indicates a partial success of debiasing.

Result 4. *Reminding participants of additional tasks alone has no discernible affect, while emphasizing the additional costs due to convexity (suggestively) reduces the impact of narrow bracketing.*

Table 7: Means of treatments with identical presentation and outcomes but without framing (NARROW), or framing additional tasks as before (BEFORE) or after (AFTER) the baseline tasks

Treatment	Res. Wage	Std Dev	% upper bound	N
Scenario 1				
NARROW	2.07	1.29	12%	207
BEFORE	2.20	1.38	13%	75
AFTER	2.27	1.38	14%	85
Scenario 2				
NARROW	2.70	1.29	20%	205
BEFORE	2.72	1.40	28%	78
AFTER	2.96	1.23	30%	84

5 Discussion of Bracketing

In this section, we consider alternative mechanisms that can lead to bracketing and explain why they do not explain our results. This provides some of the most conclusive evidence that narrow bracketing *can* be a suboptimal mistake. At the end of the section, we discuss why context and comparison effects *cannot* explain bracketing.

Table 8: Between-treatment p-values for NARROW, BEFORE, and AFTER treatments based on two-sided Wilcoxon rank-sum tests, treating each individual in each scenario as a single independent observation.

Treatments	NARROW	BEFORE
Scenario 1		
BEFORE	0.585	
AFTER	0.278	0.636
Scenario 2		
BEFORE	0.748	
AFTER	0.091	0.299

5.1 Causes of Bracketing

The literature has highlighted four causes for narrow bracketing: due to cognitive costs of broad bracketing; due to strategic concerns; due to preferences that depend on or correlate with choice brackets; or due to unintentional mistakes.

By cognitive cost we have in mind what Handel and Schwartzstein (2018) call frictions: people realize that it would be optimal to combine the choices, yet they decide that it isn't worth doing so given the cognitive costs they expect this would entail. This includes models of rational (in-)attention such as Lian (2018) and Kőszegi and Matějka (2020). Our results are inconsistent with these models, which rely on lack of information or cognitive costs. Lack of information is ruled out by having all choice-relevant information on the choice page, while cognitive costs would have to be implausibly large for MTurkers to 'decide' that computing $15 + 15$ is too costly. Note that we keep the information identical across all treatments (with the exception of early observations in our NARROW treatment – see the discussion in Section 3.4 for details), which rules out any broadly bracketed type of reference effects, since all the information is presented in one go on the first choice page.

Strategic concerns for bracketing have been studied primarily as a means for self-control (Koch and Nafziger (2011, 2016)) including in choices over effort (Koch and Nafziger (2020)). In such models, a person sets narrow goals and bears a cost from missing these goals, which can help them overcome self-control problems. While narrow goal-bracketing can lead people to respond within a given bracket, it cannot lead to different goal-bracketing in our experiment: all the possible outcomes are identical, hence then possibilities for self-control and goal-setting are also identical.

This brings us to preferences as the source of bracketing, such as the model of news utility (Kőszegi and Rabin (2009)) where people get reference-dependent utility from news about investments or gambles. If the news about a choice in one bracket is resolved separately from news in other brackets, people cannot avoid feeling the resulting news utility separately, which leads to narrow bracketing. News utility cannot explain our results, since there is a single piece of news in all treatments.

A more likely candidate for preference-based bracketing is social preferences where brackets serve as a signal of social norms or sanctions. For example, a person who is asked to split \$10 between two people may split this amount equally, even if they know that one person received more money that day than the other. They may not consider it their responsibility, their duty, or their right to affect the pre-existing income difference – and this might depend on the social context in which they are asked to make the choice. In our experiment, this would require people expecting to be treated fairly in each choice, rather than by overall outcome or by employer, which seems unlikely given that people bear the full consequences. Second and more importantly, there is almost no response to the baseline workload, which would surely affect how fair participants perceive their workload to be.

Finally, by mistakes we mean anything from people not understanding that they should combine outcomes, to 'forgetting' or not realizing that they should do so in a situation – what Handel and Schwartzstein (2018) call *mental gaps*. Our study provides some of the most conclusive evidence of narrow bracketing as a suboptimal mistake. While we interpret Tversky and Kahneman (1981) and related narrow bracketing over gambles as mistakes, the choices involved are more complicated to combine (Rabin and Weizsäcker (2009); Ellis and Freeman (2020)) or they don't rule out preference-based explanations as readily (such as Redelmeier and Tversky (1992), if participants expect to find out separate choices separately, or Exley and Kessler (2018), which involves social preferences and thus leaves room for norms).

5.2 Comparison Effects Cannot Explain Bracketing

Context effects such as focusing and range effects, like convex costs, can be applied narrowly or broadly and thus they cannot explain why a person brackets them narrowly or broadly. Since our experiment holds full outcomes constant across most treatments, context effects should be applied equally *if* they are broadly bracketed, and thus not lead to different choices. Context effects and bracketing are thus complementary dimensions of how choice presentations affect decisions. Barberis et al. (2006) explore this point in detail with respect to how it is the interaction of non-expected utility and bracketing that can make sense of observed behavior over moderate-sized gambles, not non-expected utility alone.¹⁶

6 Conclusion

In this paper, we test for and estimate broad and narrow bracketing in work choices. We characterize the identification assumption needed and reject that workers on Amazon Mechanical Turk bracket broadly, since they seem to ignore required baseline tasks. Our experiment rules out optimal conservation of cognitive resources, information, and preferences as alternative mechanisms for our results, which leaves suboptimal mistakes as the candidate explanation.

In order for narrow bracketing to be usefully applied outside the lab, we need to move from identifying narrow bracketing to estimating its impact and identify situations where it matters. One major challenge will be to jointly identify bracketing and preferences in field data. There are at least three ways to improve identification in the field. First, we can combine observational data with targeted field experiments that directly estimate preferences or bracketing (or both) to estimate relevant parameters. Second, we can model or include direct information on how the final outcomes are effectively chosen. For example, our data may contain monthly expenses on retailers even though we know that people buy multiple times per months. Finally, we may relax additivity assumptions which are central to narrow bracketing: choices made by maximizing additive preferences are not affected by bracketing, as our discussion

¹⁶For related points in choices over social preferences, see also Read et al. (1999)'s discussion of Rawlsian preferences, as well as Sobel (2005)

in section 2 makes clear. In choices over money, this issue is less severe, because money is fungible: short of large income shocks or strong liquidity constraints, a dollar is a dollar is a dollar. Thus instances of non-fungibility of money (Heath and Soll (1996); Hastings and Shapiro (2013); Abeler and Marklein (2017)) provide evidence of mental accounting, a form of narrow bracketing. Non-fungibility of work on the other hand does not immediately imply narrow bracketing as a mistake, because an hour of work is not identical to an hour of work on another day, at another hour, or on another task. This requires allowing for non-additive preferences such as habit formation or adaptation, strongly convex costs, complementarities, or non-expected utility preferences. The additional difficulty compared to monetary outcomes shows why direct evidence for narrow bracketing in work choices is important.

References

- Johannes Abeler and Felix Marklein. Fungibility, labels, and consumption. *Journal of the European Economic Association*, 15(1):99–127, 2017.
- Antonio Alonso Arechar and David Rand. Turking in the time of covid. 2020.
- Nicholas Barberis, Ming Huang, and Richard H Thaler. Individual preferences, monetary gambles, and stock market participation: A case for narrow framing. *American economic review*, 96(4):1069–1090, 2006.
- Pedro Bordalo, Nicola Gennaioli, and Andrei Shleifer. Saliency and consumer choice. *Journal of Political Economy*, 121(5):803–843, 2013.
- Benjamin Bushong, Matthew Rabin, and Joshua Schwartzstein. A Model of Relative Thinking. *The Review of Economic Studies*, 10 2020.
- Jonathan De Quidt. Your loss is my gain: a recruitment experiment with framed incentives. *Journal of the European Economic Association*, 16(2):522–559, 2018.
- Jonathan De Quidt, Francesco Fallucchi, Felix Kölle, Daniele Nosenzo, and Simone Quercia. Bonus versus penalty: How robust are the effects of contract framing? *Journal of the Economic Science Association*, 3(2):174–182, 2017.
- Arindrajit Dube, Jeff Jacobs, Suresh Naidu, and Siddharth Suri. Monopsony in online labor markets. *American Economic Review: Insights*, 2(1):33–46, March 2020.
- Andrew Ellis and David J Freeman. Revealing choice bracketing. *arXiv preprint arXiv:2006.14869*, 2020.
- Nisvan Erkal, Lata Gangadharan, and Nikos Nikiforakis. Relative earnings and giving in a real-effort experiment. *American Economic Review*, 101(7):3330–48, 2011.
- Christine L Exley and Judd B Kessler. Equity concerns are narrowly framed. Technical report, National Bureau of Economic Research, 2018.
- Diana Farrell, Fiona Greig, and Amar Hamoudi. The online platform economy in 2018: Drivers, workers, sellers, and lessors. *JPMorgan Chase Institute*, 2018.
- Diana Farrell, Fiona Greig, and Amar Hamoudi. The evolution of the online platform economy: Evidence from five years of banking data. In *AEA Papers and Proceedings*, volume 109, pages 362–66, 2019.
- Gallant. Nonlinear regression. *The American Statistician*, 29(2):73–81, 1975.

- Marcus Giamattei, Kyanoush Seyed Yahosseini, Simon Gächter, and Lucas Molleman. Lioness lab: a free web-based platform for conducting interactive experiments online. *Journal of the Economic Science Association*, pages 1–17, 2020.
- Benjamin Handel and Joshua Schwartzstein. Frictions or mental gaps: what’s behind the information we (don’t) use and when do we care? *Journal of Economic Perspectives*, 32(1):155–78, 2018.
- Justine S Hastings and Jesse M Shapiro. Fungibility and consumer choice: Evidence from commodity price shocks. *The quarterly journal of economics*, 128(4):1449–1498, 2013.
- Chip Heath and Jack B Soll. Mental budgeting and consumer decisions. *Journal of consumer research*, 23(1):40–52, 1996.
- Botond Köszegi and Matthew Rabin. Reference-Dependent Consumption Plans. *American Economic Review*, 99(3): 909–36, 2009.
- Alexander K Koch and Julia Nafziger. Self-regulation through goal setting. *Scandinavian Journal of Economics*, 113 (1):212–227, 2011.
- Alexander K Koch and Julia Nafziger. Goals and bracketing under mental accounting. *Journal of Economic Theory*, 162:305–351, 2016.
- Alexander K Koch and Julia Nafziger. Correlates of narrow bracketing. *The Scandinavian Journal of Economics*, 121 (4):1441–1472, 2019.
- Alexander K. Koch and Julia Nafziger. Motivational goal bracketing: An experiment. *Journal of Economic Theory*, 185:104949, 2020. ISSN 0022-0531. doi: <https://doi.org/10.1016/j.jet.2019.104949>. URL <http://www.sciencedirect.com/science/article/pii/S0022053119301000>.
- Botond Köszegi and Filip Matějka. Choice simplification: A theory of mental budgeting and naive diversification. *The Quarterly Journal of Economics*, 135(2):1153–1207, 2020.
- Botond Köszegi and Adam Szeidl. A model of focusing in economic choice. *The Quarterly journal of economics*, 128(1):53–104, 2012.
- Chen Lian. A theory of narrow thinking. Technical report, Working paper, Massachusetts Institute of Technology, 2018.
- Aaron J Moss, Cheskie Rosenzweig, Jonathan Robinson, and Leib Litman. Demographic stability on mechanical turk despite covid-19. *Trends in Cognitive Sciences*, 2020.

- Matthew Rabin and Georg Weizsäcker. Narrow bracketing and dominated choices. *American Economic Review*, 99(4):1508–43, 2009.
- Daniel Read, George Loewenstein, Matthew Rabin, Gideon Keren, and David Laibson. Choice bracketing. In *Elicitation of preferences*, pages 171–202. Springer, 1999.
- Donald A Redelmeier and Amos Tversky. On the framing of multiple prospects. *Psychological Science*, 3(3):191–193, 1992.
- Joel Sobel. Interdependent preferences and reciprocity. *Journal of economic literature*, 43(2):392–436, 2005.
- Amos Tversky and Daniel Kahneman. The framing of decisions and the psychology of choice. *science*, 211(4481):453–458, 1981.
- Amos Tversky and Itamar Simonson. Context-dependent preferences. *Management science*, 39(10):1179–1189, 1993.
- Pauline Vorjohann. Reference-dependent choice bracketing. 2020.

A Appendix: Proofs

Let us prove proposition 1, which we break up into multiple subresults.

Let us set up more general notation than the simplified notation used in Section 2. The full set of possible outcomes is \mathbb{X} , such as \mathbb{R} or \mathbb{R}_0^2 or the space of lotteries. We denote our two choice sets by $\mathcal{X}, \mathcal{Y} \subset \mathbb{X}$. The option chosen from \mathcal{X} is denoted by X , and the option chosen from \mathcal{Y} is denoted by Y . We call X the first choice and Y the second choice if we need to order them, although this does not represent temporal order, simply an index for referring to choices made. For example, in choices over work and money, we have $\mathbb{X} = \mathbb{R}_0^2$, and if the first choice is between doing 4 tasks for \$2 and doing nothing, and the second choice is between doing 6 tasks for \$3, then $\mathcal{X} = \{(0, 0), (4, 2)\}$ and $\mathcal{Y} = \{(0, 0), (6, 3)\}$.

There are two ways that choices can be presented: *separately* or *aggregated* (i.e. summed together). When presented separately, the person is asked to make two choices X and Y , with $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$. When presented in aggregate, the person is asked to make a single choice O , with $O \in \mathcal{O} := \mathcal{X} + \mathcal{Y}$.

We denote the choice from an aggregate decision as $O_A(\mathcal{X}, \mathcal{Y})$ (O for 'overall outcome'), and we denote by $F_A(\mathcal{X}, \mathcal{Y})$ the option that was chosen from the first choice set, here \mathcal{X} ; and by $S_A(\mathcal{X}, \mathcal{Y})$ the option that was chosen in the second choice set, here \mathcal{Y} . For instance if $\mathcal{X} = \{1, 2\}$ and $\mathcal{Y} = \{3, 5\}$, then $\mathcal{O} = \{4, 5, 6, 7\}$. If $O_A(\mathcal{X}, \mathcal{Y}) = 6$, we know that the choice from \mathcal{X} was 1 and the choice from $\mathcal{Y} = 5$. In general this can be multi-valued, a technical issue that we will ignore.¹⁷ The total outcome is just the sum of the two choices, hence $O_A = F_A + S_A$, where we drop the arguments when they are clear implicitly. Similarly denote by $O_S, F_S,$ and S_S the overall choice, as well as the first and second choices made from \mathcal{X} and \mathcal{Y} when choices are presented separately.

From our definition, a person *brackets broadly* if $O_S(\mathcal{X}, \mathcal{Y}) = O_A(\mathcal{X}, \mathcal{Y})$ for all \mathcal{X} and \mathcal{Y} . A person *brackets narrowly* if $F_S(\mathcal{X}, \mathcal{Y}) = F_S(\mathcal{X}, \mathcal{Y}')$ for all \mathcal{X}, \mathcal{Y} and \mathcal{Y}' (by symmetry this also directly holds for S_S , since we can exchange \mathcal{X} and \mathcal{Y}).

Thus narrow and broad bracketing are unidentifiable if both conditions hold at the same time. Being unidentified means that even if a person brackets narrowly, their behavior will still look consistent with broad bracketing; and similarly if a person brackets broadly, their behavior will look consistent with narrow bracketing. Suppose that choices are unidentifiable, so that both conditions hold for all choice sets. Then we have:

¹⁷If we assume that all choice sets are strictly concave, this will hold; if we assume that choice sets are finite, then by continuity we can slightly disturb choices so as to make them unique, which shows that multi-valuedness is non-generic in such cases. Since in most situations, finite choice sets together with continuity fully can be extended to preferences over all choice sets, this covers most interesting cases.)

$$\begin{aligned}
F_A(\mathcal{X}, \mathcal{Y}) + S_A(\mathcal{X}, \mathcal{Y}) &= O_A(\mathcal{X}, \mathcal{Y}), \text{ by definition of } O_A \\
&= O_S(\mathcal{X}, \mathcal{Y}), \text{ by broad bracketing} \\
&= F_S(\mathcal{X}, \mathcal{Y}) + S_S(\mathcal{X}, \mathcal{Y}), \text{ by definition of } O_S \\
&= F_S(\mathcal{X}, \mathcal{Y}') + S_S(\mathcal{X}, \mathcal{Y}'), \text{ by narrow bracketing} \\
&= F_A(\mathcal{X}, \mathcal{Y}') + S_A(\mathcal{X}, \mathcal{Y}'), \text{ by redoing all the steps in reverse with } \mathcal{Y}'
\end{aligned}$$

This implies that

$$F_A(\mathcal{X}, \mathcal{Y}') = F_A(\mathcal{X}, \mathcal{Y}) = F_S(\mathcal{X}, \mathcal{Y}') = F_S(\mathcal{X}, \mathcal{Y})$$

so that we can simply refer to the first choice by $C(\mathcal{X})$: it depends only on what the choice set is, not on the other choice set, nor on whether it is presented separately or in aggregate or whether it is the first or the second. The same holds for the second choice, which we can write as $C(\mathcal{Y})$.¹⁸ And therefore, the same applies to overall outcomes, which can be written as $C(\mathcal{O})$ for $\mathcal{O} = \mathcal{X} + \mathcal{Y}$.

When neither choice depends on presentation nor on the other choice, it is clear that bracketing is unidentified, since both narrow and broad bracketing hold. This proves the following proposition:

Proposition 2. *Bracketing is unidentified if and only if the following hold for all $\mathcal{X}, \mathcal{Y}, \mathcal{X}', \mathcal{Y}'$ and $\mathcal{O} = \mathcal{X} + \mathcal{Y}$:*

$$F_A(\mathcal{X}, \mathcal{Y}') = F_A(\mathcal{X}, \mathcal{Y}) = F_S(\mathcal{X}, \mathcal{Y}') = F_S(\mathcal{X}, \mathcal{Y}) = C(\mathcal{X})$$

$$S_A(\mathcal{X}, \mathcal{Y}) = S_A(\mathcal{X}', \mathcal{Y}) = S_S(\mathcal{X}', \mathcal{Y}) = S_S(\mathcal{X}, \mathcal{Y}) = C(\mathcal{Y})$$

$$O_S(\mathcal{X}, \mathcal{Y}) = O_A(\mathcal{X}, \mathcal{Y}) = C(\mathcal{O})$$

Let us define the aggregate money metric $M(\cdot)$. First we have to assume that we have a numeraire good, which we'll call money, which exists under sufficiently strong continuity conditions. This may be money itself, some other good or composition of goods, or a continuous measure for the probability over some lotteries. Then for every outcome X , let us define $M(X)$ such that the person is indifferent between paying $M(X)$ and receiving X or receiving 0 and no money *in choices presented in aggregated form*. (The bundle X may itself contain any quantity of the numeraire.) Thus $M(X)$ is defined via the indifference relation $(X, M(X)) \sim (0, 0)$ over aggregate choices, which is independent of whether the person brackets narrowly or broadly. Thus (X, P) can be read as “paying P to receive X ”, where

¹⁸Here we rely on the fact that $F_i(\cdot)$ and $S_i(\cdot)$ are single-valued and hence functions, so that there is only one $X \in \mathcal{X}$ and one $Y \in \mathcal{Y}$ that add up to the given overall outcome.

payments are done in the numeraire.¹⁹ We are not assuming that the numeraire is separable from other goods, only that it spans the whole range of utility.

Definition 2. A function $f(\cdot) : X \rightarrow \mathbb{R}$ is additive if $f(x + y) = f(x) + f(y)$ for all x, y .

With this, we can state the following proposition:

Proposition 3. Suppose we have a money metric $M(\cdot) : \mathbb{X} \rightarrow \mathbb{R}$ for choices presented in aggregate, and suppose that preferences satisfy appropriate continuity assumptions. \mathbb{X} is closed under addition: if $X, Y \in \mathbb{X}$, then $X + Y \in \mathbb{X}$. Then bracketing is unidentifiable if and only if the money metric over aggregately presented choices $M(\cdot)$ is additive.

Proof. It is easy to see that if $M(X + Y) = M(X) + M(Y)$ for all X and Y , then the choice $X \in \mathcal{X}$ does not depend on how the choices are presented, nor on what the second choice set is, hence bracketing is unidentifiable by proposition 2.

For the reverse, simply note that $M(X)$ is a utility representation for the preferences over aggregate presented choices. If preferences are unidentifiable, then the choices over choices that are separately presented are identical to those if they are presented jointly by 2, hence $M(X)$ is also a utility representation for choices in separately presented choices, i.e. people choose the item that maximizes $M(\cdot)$ in each choice set, no matter how it is presented.

Suppose by contradiction that $M(A + B) > M(A) + M(B)$ for some A, B . Denote by (X, M) the bundle that provides the bundle X for the price M , i.e. the person loses M and gets X . Let $\mathcal{X} = \{(0, 0), (A, M(A) + \varepsilon)\}$ and $\mathcal{Y} = \{(0, 0), (B, M(B) + \varepsilon)\}$, where $0 < \varepsilon < M(A + B) - M(A) - M(B)$.²⁰ Then by unidentifiability, $F_S(\mathcal{X}, \mathcal{Y}) = F(\mathcal{X}) = (0, 0)$ since the person is willing to pay at most $M(A)$ for A . Similarly, $S_S(\mathcal{X}, \mathcal{Y}) = (0, 0)$ since the person is willing to pay at most $M(B)$ for B . Yet, $O_A(\mathcal{X}, \mathcal{Y})$ is $A + B$, since the payment required for this is $M(A) + M(B) + 2 \cdot \varepsilon < M(A + B)$, but by unidentifiability, we have that $O_A(\mathcal{X}, \mathcal{Y}) = O_S(\mathcal{X}, \mathcal{Y}) = F(\mathcal{X}) + S(\mathcal{Y}) = (0, 0)$, which contradicts unidentifiability.

The same argument holds if $M(A + B) < M(A) + M(B)$, proving the proposition. □

We will now use proposition 3 to prove the special subcases mentioned in proposition 1.

Proposition 4. Suppose $\mathbb{X} = \mathbb{R}^N$, so that the choice is between two goods, whether different goods, or goods on different dates, or in different states. If M is additive and continuous, then it is linear: $M(X) = \sum_{i=1}^N \lambda_i X_i$ for some $\lambda \in \mathbb{R}^N$.

Proof. Additivity plus continuity for one-dimensional functions in \mathbb{R} implies linearity.²¹ Letting e_i be the unit vector in dimension i , i.e. it is the bundle that provides one unit of good i and nothing else, then for $x \in \mathbb{R}$, we have that

¹⁹We assume that negative amounts of the numeraire are allowed, so that such an $M(X)$ exists, i.e. we allow people to be paid for taking X .

²⁰Here we used continuity of the money metric, i.e. of the numeraire.

²¹Additivity defining a function is also called Cauchy's functional equation and the stated result dates back to Cauchy.

$f_i(x) := M(x \cdot e_i)$ is a one-dimensional function that is additive and continuous, hence $M(x \cdot e_i) = \lambda_i x$ for some λ_i . By additivity we have that $M(X) = M(\sum_i e_i X_i) = \sum_i M(X_i \cdot e_i) = \sum_i \lambda_i X_i$. \square

Proposition 5. *Suppose \mathbb{X} is a set of risky choices, possibly with a finite number of outcomes. If M is additive and continuous in probability, then M satisfies the independence axiom and has an expected utility representation.*

Proof. Denote by X , Y , and Z three random variables, possibly multi-dimensional. Note that if A is an event independent of the three random variables that happens with probability p , and A^C denotes its complement, then $\tilde{X} = \mathbb{1}(A) \cdot X + \mathbb{1}(A^C) \cdot Z$ and $\tilde{Y} = \mathbb{1}(A) \cdot Y + \mathbb{1}(A^C)Z$ are the random variables yielding the value of X respectively Y with probability p and the value of Z with probability $1 - p$.

Let A be distributed uniformly on $[0, 1]$, independently of X , Y , Z . Let $A(p, q)$ be the event that $A \in (p, q)$. Hence

$$\begin{aligned} M(\mathbb{1}(A(0, p+q))X) &= M(\mathbb{1}(A(0, p))X + \mathbb{1}(A(p, p+q))X) \\ &= M(\mathbb{1}(A(0, p))X) + M(\mathbb{1}(A(p, p+q))X), \text{ by additivity} \\ &= M(\mathbb{1}(A(0, p))X) + M(\mathbb{1}(A(0, q))X), \text{ by equal distributions} \end{aligned}$$

where we used the fact that $\mathbb{1}(A(0, q))X$ and $\mathbb{1}(A(p, p+q))X$ have the same distribution, hence also the same utility. Writing $f_X(p) = M(\mathbb{1}(A(0, p))X)$, we have that f_X is additive, i.e. it satisfies $f_X(p+q) = f_X(p) + f_X(q)$. We assumed that it is continuous in p , hence we know that f_X is linear, i.e. $f_X(p) = \lambda p$ for some λ . Since $f_X(1) = M(X)$, we have $\lambda = M(X)$, which shows that $M(\mathbb{1}(A(0, p))X) = pM(X)$.

Let $\tilde{X} = \mathbb{1}(A(0, p))X + \mathbb{1}(A(p, 1))Z$ and similarly for $\tilde{Y} = \mathbb{1}(A(0, p))Y + \mathbb{1}(A(p, 1))Z$. Suppose that $M(X) > M(Y)$. Then we have $M(\tilde{X}) = M(\mathbb{1}(A(0, p))X) + M(\mathbb{1}(A(p, 1))Z) = pM(X) + (1-p)M(Z) > pM(Y) + (1-p)M(Z) = M(\mathbb{1}(A(0, p))Y) + M(\mathbb{1}(A(p, 1))Z) = M(\tilde{Y})$. The same argument applies to any event B with probability p (independent of the lotteries) since the distributions and hence utilities are identical, i.e. $M(\mathbb{1}(B)X) = M(\mathbb{1}(A(0, p))X)$. Thus the independence axiom holds, which with continuity implies expected utility. \square

The next result is a rewording of part of proposition 2 of Rabin and Weizsacker (2009).

Proposition 6. *Suppose \mathbb{X} is a set of risky choices over a single continuous dimension, such as wealth. If M is additive and an expected utility representation, then M has constant absolute risk aversion.*

Proof. Let X be any lottery and w and w' any two deterministic wealth levels. Then $M(X + w) = M(X) + M(w) \implies M(X + w) - M(w) = M(X) = M(X + w') - M(w')$. This is the certainty equivalent, which has to be constant for all wage levels. For expected utility preferences this implies that we have constant absolute risk aversion. \square

From this, we get the following corollaries.

Corollary 1. *Suppose M is additive and defined over risky wealth outcomes. Then M is an expected utility with CARA.*

Proof. Proposition 5 shows that M satisfies the independence axiom, hence by continuity that it has expected utility form, hence we can apply proposition 6. \square

And finally

Corollary 2. *Suppose M is additive and over at least two goods with risk and time. Then M is linear in all goods and with risk-neutral expected utility. If moreover M is stationarity, then we get a single linear factor for each good and the others are determined by exponential discounting.*

Proof. This follows from applying all prior propositions in and across all dimensions. \square

This concludes all results relating to additivity under context-independence from proposition 1. We now prove that if bracketing is unidentified, then WARP holds, hence we cannot have context-dependent preferences, since we have the independence of irrelevant alternatives (IIA) assumption. This rules out context-dependent preferences such as Tversky and Simonson (1993), Kőszegi and Szeidl (2012), Bordalo et al. (2013), or Bushong et al. (2020).

Proposition 7. *Suppose that we have preferences for which bracketing is not identifiable. Then under sufficient continuity assumptions, and assuming that adding positive amounts of numeraire to a single option make it more desirable, WARP holds.*

Proof. Suppose that we have unidentifiability. Suppose by contradiction that WARP fails for choices in overall aggregate choices, so that there are $X, Y \in \mathcal{X} \cap \mathcal{Y}$ s.t. $X \in C(\mathcal{X})$ and $Y \in C(\mathcal{Y})$, yet $X \notin C(\mathcal{Y})$. Then assuming that each $C(\cdot)$ is a singleton set and that we have continuity, this implies that for $\mathcal{Y}' = \mathcal{Y} \setminus \{X\} \cup \{X + \varepsilon\}$, we also have $C(\mathcal{Y}') = C(\mathcal{Y}) = Y$ for sufficiently small ε of the numeraire. (Here we rely strongly on continuity, either by considering only finite choice sets for which singleton choice sets and simple continuity will give these results straightforwardly, or by relying on stronger forms of continuity that rule out overly strong context-dependent effects, such as making one option better makes other options seem better by even more etc.)

Let us now offer two separate choices \mathcal{X} and \mathcal{Y}' . By unidentifiability, we know that $C(\mathcal{X}) = X$ since the choice is identical to the choice made when it is the only choice. Similarly, we have $C(\mathcal{Y}') = Y$. Hence the person chooses $X + Y$ from $\mathcal{O} = \mathcal{X} + \mathcal{Y}'$. But $X + Y + \varepsilon \in \mathcal{O}$, which implies that the person could have gotten something strictly better (again, ruling out context-dependent preferences that violate various forms of dominance). This is a contradiction, hence WARP holds. \square

B Appendix: Instructions

B.1 Welcome Screen

Welcome

Thank you for accepting our HIT.

During the HIT, please do not close this window or leave the HIT's web pages in any other way.

If you do close your browser or leave the HIT, you will not be able to re-enter and we will not be able to pay you!

You will receive a baseline payment of \$2.00 once you complete the HIT. Additionally, you can earn an extra bonus that will depend on your choices.

You will receive a code to enter into MTurk to collect your payment once you have finished.

Please read all instructions carefully.

B.2 Instructions

Thank you for accepting to participate in this HIT. On top of the guaranteed payment of \$2.00 you will have the chance to earn an extra bonus, as explained later.

The task In this HIT you will decode several sequences of random letters into numbers with the given decoding table. For each letters sequence, the decoding table changes. The main part of the HIT will require you to decode several of these tasks.

To gain familiarity with the task you will now have to correctly decode 3 sequences. Note that each letter must be decoded correctly. After entering the decoded sequence, hit the submit button. Subsequently, irrespective of whether the text sequence was decoded correctly or not, a new sequence and decoding table will appear. Once you decode 3 sequences correctly, we describe the main part of the HIT.

Encryption task

Decode sequences into numbers until you got 10 right.

Correctly decoded text sequences: 0

Decoding table

letter	u	t	f	c	g	n	v	q	a	l
corresponds to number	0	1	2	3	4	5	6	7	8	9

Text sequence:

Your answer (enter only numbers):

167757642404

Figure 2: SCREENSHOT

In the example you see the text sequence *tvqqnqvfgug*. The decoding table tells you that u=0, t=1,... This means that you have to decode *tvqqnqvfgug* into *167757642404* and enter this numeric value into the answer field.

B.3 Task

B.4 Main Task instructions

THE TASK

[NARROW] By completing the HIT you will receive \$2.00. To do so you are required to decode some sequences correctly for a bonus.

We will give you two pages of choices, with 16 choices on each page. Each choice is between a low number and a high number of additional sequences to decode before the required sequences for different bonuses.

Example Choice (DOES NOT COUNT):

- 10 additional sequences for an extra \$4.00
- 20 additional sequences for an extra \$5.00

[LOW] By completing the HIT you will receive \$2.00. To do so you are required to decode some sequences correctly for a bonus.

We will give you two pages of choices, with 16 choices on each page. Each choice is between a low number and a high number of sequences to decode for different bonuses.

Example Choice (DOES NOT COUNT):

- 10 sequences for an extra \$4.00
- 20 sequences for an extra \$5.00

[BROAD] To complete the HIT you will be asked to decode a certain number of sequences correctly. The number of sequences you will be required to decode will depend on your choices.

We will give you two pages of choices, with 16 choices on each page. Each choice is between a low number and a high number of sequences to decode for different amounts (which includes the \$2.00 completion fee of the HIT).

- 10 sequences for a total payment of \$6.00
- 20 sequences for a total payment of \$7.00

[PARTIAL] By completing the HIT you will receive \$2.00. To do so you are required to decode some sequences correctly for a bonus.

We will give you two pages of choices, with 16 choices on each page. Each choice is between a low number and a high number of sequences to decode for different bonuses.

Example Choice (DOES NOT COUNT):

- 10 sequences for an extra \$4.00
- 20 sequences for an extra \$5.00

After you made your choice the computer will select randomly one of the 16 choices from one of the 2 pages. That option will be implemented. Thus you should select your preferred option for each choice.

B.5 Main Task - Scenario 1

[BROAD]

Choices to make now: for each choice in this Scenario, choose the preferred option.

By completing the HIT you will receive a total payment (which includes the \$2.00 completion fee) depending on your choices.

		OPTION A	OPTION B	
1)	15 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	30 sequences for a total payment of \$6.25
2)	15 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	30 sequences for a total payment of \$6.50
3)	15 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	30 sequences for a total payment of \$6.75
4)	15 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	30 sequences for a total payment of \$7.00
5)	15 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	30 sequences for a total payment of \$7.25
6)	15 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	30 sequences for a total payment of \$7.50
7)	15 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	30 sequences for a total payment of \$7.75
8)	15 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	30 sequences for a total payment of \$8.00
9)	15 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	30 sequences for a total payment of \$8.25
10)	15 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	30 sequences for a total payment of \$8.50
11)	15 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	30 sequences for a total payment of \$8.75
12)	15 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	30 sequences for a total payment of \$9.00
13)	15 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	30 sequences for a total payment of \$9.25
14)	15 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	30 sequences for a total payment of \$9.50
15)	15 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	30 sequences for a total payment of \$9.75
16)	15 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	30 sequences for a total payment of \$10.00

[PARTIAL]

Choices to make now: for each choice in this Scenario, choose the preferred option.

By completing the HIT you will receive \$2.00 plus a bonus depending on your choices.

		OPTION A	OPTION B	
1)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$4.25
2)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$4.50
3)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$4.75
4)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$5.00
5)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$5.25
6)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$5.50
7)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$5.75
8)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$6.00
9)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$6.25
10)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$6.50
11)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$6.75
12)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$7.00
13)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$7.25
14)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$7.50
15)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$7.75
16)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$8.00

[NARROW]

Note: you are required to decode 15 sequences correctly, in addition to the sequences based on your choices.

Choices to make now: for each choice in this Scenario, choose the preferred option.

By completing the HIT you will receive \$2.00 plus a bonus depending on your choices.

		OPTION A	OPTION B	
1)	0 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	15 additional sequences for an extra \$4.25
2)	0 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	15 additional sequences for an extra \$4.50
3)	0 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	15 additional sequences for an extra \$4.75
4)	0 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	15 additional sequences for an extra \$5.00
5)	0 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	15 additional sequences for an extra \$5.25
6)	0 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	15 additional sequences for an extra \$5.50
7)	0 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	15 additional sequences for an extra \$5.75
8)	0 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	15 additional sequences for an extra \$6.00
9)	0 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	15 additional sequences for an extra \$6.25
10)	0 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	15 additional sequences for an extra \$6.50
11)	0 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	15 additional sequences for an extra \$6.75
12)	0 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	15 additional sequences for an extra \$7.00
13)	0 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	15 additional sequences for an extra \$7.25
14)	0 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	15 additional sequences for an extra \$7.50
15)	0 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	15 additional sequences for an extra \$7.75
16)	0 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	15 additional sequences for an extra \$8.00

[LOW]

Choices to make now: for each choice in this Scenario, choose the preferred option.

By completing the HIT you will receive \$2.00 plus a bonus depending on your choices.

		OPTION A	OPTION B	
1)	0 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	15 sequences for an extra bonus of \$4.25
2)	0 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	15 sequences for an extra bonus of \$4.50
3)	0 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	15 sequences for an extra bonus of \$4.75
4)	0 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	15 sequences for an extra bonus of \$5.00
5)	0 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	15 sequences for an extra bonus of \$5.25
6)	0 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	15 sequences for an extra bonus of \$5.50
7)	0 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	15 sequences for an extra bonus of \$5.75
8)	0 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	15 sequences for an extra bonus of \$6.00
9)	0 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	15 sequences for an extra bonus of \$6.25
10)	0 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	15 sequences for an extra bonus of \$6.50
11)	0 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	15 sequences for an extra bonus of \$6.75
12)	0 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	15 sequences for an extra bonus of \$7.00
13)	0 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	15 sequences for an extra bonus of \$7.25
14)	0 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	15 sequences for an extra bonus of \$7.50
0)	0 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	15 sequences for an extra bonus of \$7.75
16)	0 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	15 sequences for an extra bonus of \$8.00

B.6 Main Task - Scenario 2

[BROAD]

Choices to make now: for each choice in this Scenario, choose the preferred option.

By completing the HIT you will receive a total payment (which includes the \$2.00 completion fee) depending on your choices.

		OPTION A	OPTION B	
1)	30 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	45 sequences for a total payment of \$6.25
2)	30 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	45 sequences for a total payment of \$6.50
3)	30 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	45 sequences for a total payment of \$6.75
4)	30 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	45 sequences for a total payment of \$7.00
5)	30 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	45 sequences for a total payment of \$7.25
6)	30 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	45 sequences for a total payment of \$7.50
7)	30 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	45 sequences for a total payment of \$7.75
8)	30 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	45 sequences for a total payment of \$8.00
9)	30 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	45 sequences for a total payment of \$8.25
10)	30 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	45 sequences for a total payment of \$8.50
11)	30 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	45 sequences for a total payment of \$8.75
12)	30 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	45 sequences for a total payment of \$9.00
13)	30 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	45 sequences for a total payment of \$9.25
14)	30 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	45 sequences for a total payment of \$9.50
15)	30 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	45 sequences for a total payment of \$9.75
16)	30 sequences for a total payment of \$6.00	<input type="radio"/>	<input type="radio"/>	45 sequences for a total payment of \$10.00

[PARTIAL]

Choices to make now: for each choice in this Scenario, choose the preferred option.

By completing the HIT you will receive \$2.00 plus a bonus depending on your choices.

		OPTION A	OPTION B	
1)	30 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	45 sequences for an extra bonus of \$4.25
2)	30 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	45 sequences for an extra bonus of \$4.50
3)	30 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	45 sequences for an extra bonus of \$4.75
4)	30 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	45 sequences for an extra bonus of \$5.00
5)	30 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	45 sequences for an extra bonus of \$5.25
6)	30 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	45 sequences for an extra bonus of \$5.50
7)	30 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	45 sequences for an extra bonus of \$5.75
8)	30 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	45 sequences for an extra bonus of \$6.00
9)	30 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	45 sequences for an extra bonus of \$6.25
10)	30 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	45 sequences for an extra bonus of \$6.50
11)	30 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	45 sequences for an extra bonus of \$6.75
12)	30 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	45 sequences for an extra bonus of \$7.00
13)	30 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	45 sequences for an extra bonus of \$7.25
14)	30 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	45 sequences for an extra bonus of \$7.50
15)	30 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	45 sequences for an extra bonus of \$7.75
16)	30 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	45 sequences for an extra bonus of \$8.00

[NARROW]

Note: you are required to decode 15 sequences correctly, in addition to the sequences based on your choices.

Choices to make now: for each choice in this Scenario, choose the preferred option.

By completing the HIT you will receive \$2.00 plus a bonus depending on your choices.

		OPTION A	OPTION B	
1)	15 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	30 additional sequences for an extra \$4.25
2)	15 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	30 additional sequences for an extra \$4.50
3)	15 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	30 additional sequences for an extra \$4.75
4)	15 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	30 additional sequences for an extra \$5.00
5)	15 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	30 additional sequences for an extra \$5.25
6)	15 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	30 additional sequences for an extra \$5.50
7)	15 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	30 additional sequences for an extra \$5.75
8)	15 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	30 additional sequences for an extra \$6.00
9)	15 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	30 additional sequences for an extra \$6.25
10)	15 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	30 additional sequences for an extra \$6.50
11)	15 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	30 additional sequences for an extra \$6.75
12)	15 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	30 additional sequences for an extra \$7.00
13)	15 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	30 additional sequences for an extra \$7.25
14)	15 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	30 additional sequences for an extra \$7.50
15)	15 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	30 additional sequences for an extra \$7.75
16)	15 additional sequences for an extra \$4.00	<input type="radio"/>	<input type="radio"/>	30 additional sequences for an extra \$8.00

[LOW]

Choices to make now: for each choice in this Scenario, choose the preferred option.

By completing the HIT you will receive \$2.1515 plus a bonus depending on your choices.

		OPTION A	OPTION B	
1)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$4.25
2)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$4.50
3)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$4.75
4)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$5.00
5)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$5.25
6)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$5.50
7)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$5.75
8)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$6.00
9)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$6.25
115)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$6.50
11)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$6.75
12)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$7.00
13)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$7.25
14)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$7.50
15)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$7.75
16)	15 sequences for an extra bonus of \$4.00	<input type="radio"/>	<input type="radio"/>	30 sequences for an extra bonus of \$8.00

B.7 Results

[NARROW & PARTIAL] SUMMARY OF THE TASK

The computer randomly selected the Choice # from Scenario #.

For this option you selected that you are (are not) willing to decode # additional sequences for \$X.XX.

In total you will decode # sequences to receive the HIT payment and the bonus.

[BROAD & LOW] SUMMARY OF THE TASK

The computer randomly selected the Choice # from Scenario #.

For this option you selected that you are (are not) willing to decode # sequences in total for \$X.XX.

In total you will decode # sequences to receive the HIT payment and the bonus.

Table 9: Attrition in % by a given stage

Treatments	Practice	Choice 1	Answer 1	Learn Tasks	End
Main					
BROAD	11%	11%	11%	12%	18%
NARROW	11%	11%	11%	12%	20%
LOW	9%	10%	10%	11%	13%
PARTIAL	8%	8%	8%	8%	16%
Follow Up					
BEFORE	7%	7%	7%	7%	18%
AFTER	7%	7%	7%	7%	14%

B.8 MAIN TASK

B.9 PAYMENT PAGE

Your earnings

In today's HIT you have earned a bonus of \$.

Your guaranteed participation fee is: \$2.00.

So, in total, you have earned \$.

To receive your earnings, please enter this code into MTurk

After you have done that, you can close this window. We thank you for participating in our study.

C Appendix: Additional Results

C.1 BEFORE and AFTER summary statistics

C.2 Baseline Tasks revealed on the choice page only

Here we report the results when we restrict the data from the NARROW treatment to data to those sessions where the baseline endowment is only revealed on the first choice page, rather than on the page right before, as was inadvertently the case for early sessions.

Broad bracketing is rejected as before, narrow bracketing is still not rejected, but when estimating κ with PARTIAL as the broader treatment, the value of κ jumps all over the place and is incredibly noisy (although as mentioned in the main text, PARTIAL and LOW are so close that one shouldn't take the standard errors seriously).

Moreover, NARROW and AFTER are statistically significantly different in this case, as indicated by Table 15. However, the issues around the different sample population for BEFORE/AFTER remain, given that we collected

Table 10: Summary statistics for follow-up treatments

	NARROW	BEFORE	AFTER	p-value
Participants	320	150	152	
Attrition	20.3%	18%	13.8%	0.23
Final Participants	255	123	131	
Share Female	0.38	0.36	0.39	0.68
Age	35	35.1	35.9	0.47
Tediousness	7.45	7.48	7.56	0.93
Inconsistent Choices				
Scenario 1	60	56	51	0
Scenario 2	59	53	53	0

Table 11: Means of main treatments by scenario, restricting NARROW to sessions where baseline is revealed on first choice page.

Treatment	Res. Wage	Std Dev	% upper bound	N
Scenario 1				
BROAD	2.88	1.24	24%	137
NARROW	2.01	1.36	12%	90
LOW	2.31	1.33	18%	131
PARTIAL	2.50	1.28	15%	137
Scenario 2				
BROAD	2.99	1.29	33%	141
NARROW	2.49	1.33	15%	88
LOW	2.74	1.32	25%	135
PARTIAL	2.43	1.32	18%	138

most of the data after COVID-19 induced lockdowns.

C.3 Baseline Tasks revealed right before the choice page only

Next we report results from the initial NARROW treatments where the information was displayed on the page right before. The results are essentially the same, although there is no longer a statistically significant difference in scenario 2, since NARROW lies between LOW and BROAD and is not significantly different from either, reflecting the lower power due to closer to 'linear' preferences in Scenario 2 (the difference between LOW and BROAD is lower).

We compare the means of NARROW treatments with message displayed before the first choice page and on the first choice page by scenario directly in Table 19. This shows that for scenario 2, these two versions are significantly (and sizeably) different, reflecting also that in one case this leads to rejection of broad bracketing in scenario 2 and once it doesn't. No matter which is the accurate treatment, both reject broad bracketing, and neither rejects narrow

Table 12: Between-treatment p-values for main treatments based on two-sided Wilcoxon rank-sum tests, treating each individual in each scenario as a single independent observation. Restricted to those sessions of NARROW where baseline is revealed only on first choice page.

Treatments	BROAD	NARROW	LOW
Scenario 1			
NARROW	< 0.001		
LOW	< 0.001	0.093	
PARTIAL	0.012	0.005	0.209
Scenario 2			
NARROW	0.003		
LOW	0.108	0.156	
PARTIAL	< 0.001	0.719	0.054

Table 13: Estimates of κ when the broader option is BROAD or PARTIAL respectively. Narrow option is always LOW, κ estimates the convex combination between the broader and the narrow option, with $\kappa = 0$ indicating equality with the broader option.

Terms	BROAD All	BROAD Male	BROAD Female	PARTIAL All	PARTIAL Male	PARTIAL Female
S1 broader (B_1)	2.88 (0.11)	2.83 (0.14)	2.94 (0.17)	2.52 (0.11)	2.41 (0.12)	2.46 (0.19)
S2 broader (B_2)	3 (0.11)	3.12 (0.13)	2.83 (0.17)	2.51 (0.11)	2.63 (0.11)	2.34 (0.18)
S1 narrow (N_1)	2.33 (0.11)	2.47 (0.12)	2.04 (0.19)	2.26 (0.11)	2.53 (0.14)	2 (0.18)
S2 narrow (N_2)	2.7 (0.08)	2.8 (0.11)	2.57 (0.14)	2.57 (0.07)	2.72 (0.13)	2.5 (0.13)
κ	1.62 (0.38)	2.01 (0.74)	1.34 (0.38)	1.88 (1.05)	-2.32 (4.1)	1.46 (0.77)
Observations	722	433	285	719	447	264

Table 14: Restricting NARROW to sessions done with presentation on choice page only. Means of treatments with identical presentation and outcomes but without framing (NARROW), or framing additional tasks as before (BEFORE) or after (AFTER) the baseline tasks

Treatment	Res. Wage	Std Dev	% upper bound	N
Scenario 1				
NARROW	2.01	1.36	12%	90
BEFORE	2.20	1.38	13%	75
AFTER	2.27	1.38	14%	85
Scenario 2				
NARROW	2.49	1.33	15%	88
BEFORE	2.72	1.40	28%	78
AFTER	2.96	1.23	30%	84

Table 15: Between-treatment p-values for NARROW, BEFORE, and AFTER treatments based on two-sided Wilcoxon rank-sum tests, treating each individual in each scenario as a single independent observation. Restricted to choices of NARROW when baseline is mentioned on choice page first.

Treatments	NARROW	BEFORE
Scenario 1		
BEFORE	0.411	
AFTER	0.208	0.636
Scenario 2		
BEFORE	0.225	
AFTER	0.013	0.299

bracketing.

There are two possible reasons for the difference: either it is due to the display of information, in which case the later data with information display on the page is the appropriate test, rejecting broad in both scenarios. In this case, the treatments NARROW and BROAD are not balanced within sessions, since we had completed collection of data on BROAD (mostly at least, we have a small overlap between the treatments). Or it is due to changes in the population due to sampling at different times. In this case the earlier data is the appropriate test, and balances observations against the BROAD treatment – i.e. the rejection of broad bracketing cannot be due (or more correctly, is statistically unlikely to be due) to different preferences.

C.4 Plots of reservation wages by treatment

Table 16: Linear Regressions of reservation wages by treatment and scenario, with and without interaction effects

	<i>Dependent variable:</i>	
	Reservation wage	
	Without Interactions	With Interactions
	(1)	(2)
scenarioScenario2	0.31 (0.07)	0.10 (0.16)
treatmentNARROW	-0.55 (0.10)	-0.82 (0.14)
treatmentLOW	-0.41 (0.11)	-0.58 (0.16)
treatmentPARTIAL	-0.47 (0.11)	-0.38 (0.16)
scenarioScenario2:treatmentNARROW		0.53 (0.20)
scenarioScenario2:treatmentLOW		0.33 (0.22)
scenarioScenario2:treatmentPARTIAL		-0.17 (0.22)
Constant	2.78 (0.09)	2.88 (0.11)
Observations	1,231	1,231
R ²	0.04	0.05
F Statistic	12.62 (df = 4; 1226)	9.36 (df = 7; 1223)

Note:

Standard errors in parentheses.

Figure 3: A bar plot of the raw reservation wages by treatment and scenario

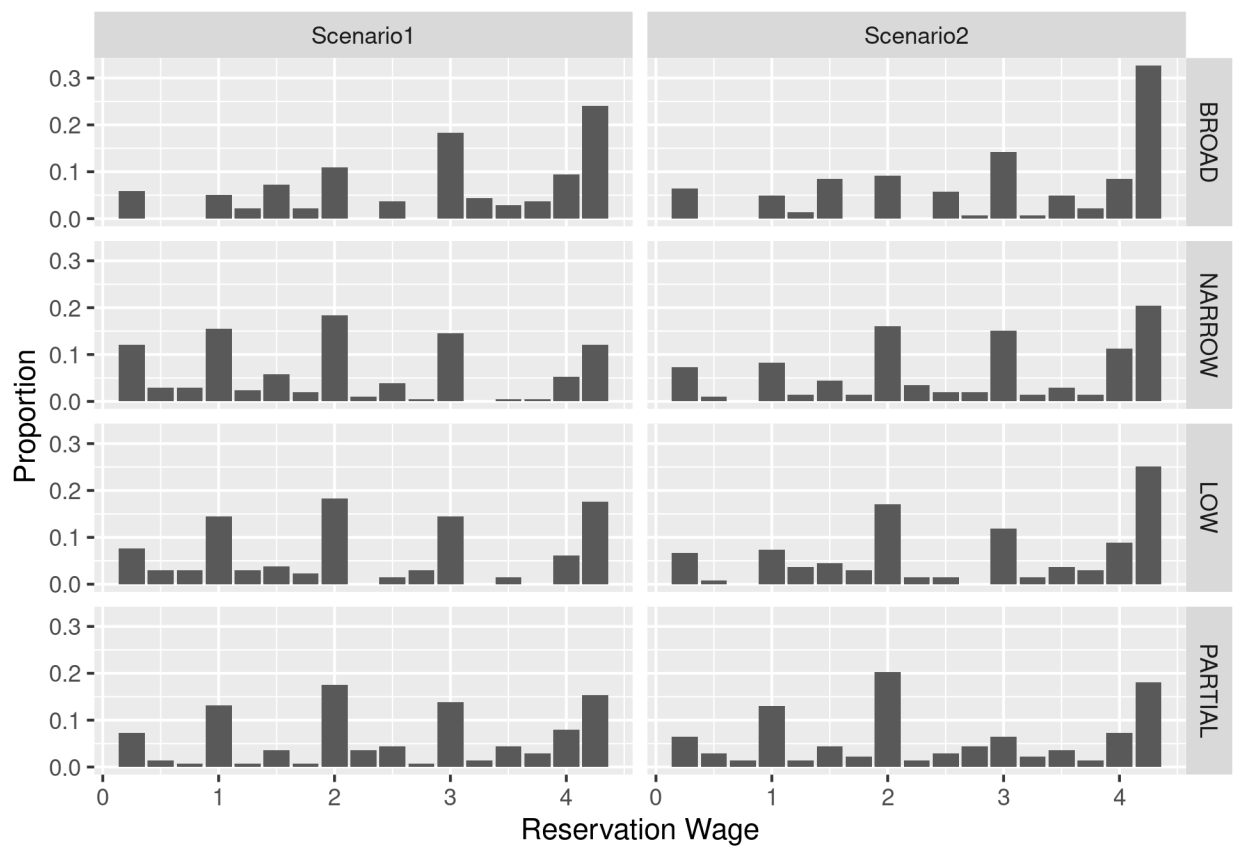


Figure 4: A kernel density plot of the raw reservation wages by treatment and scenario

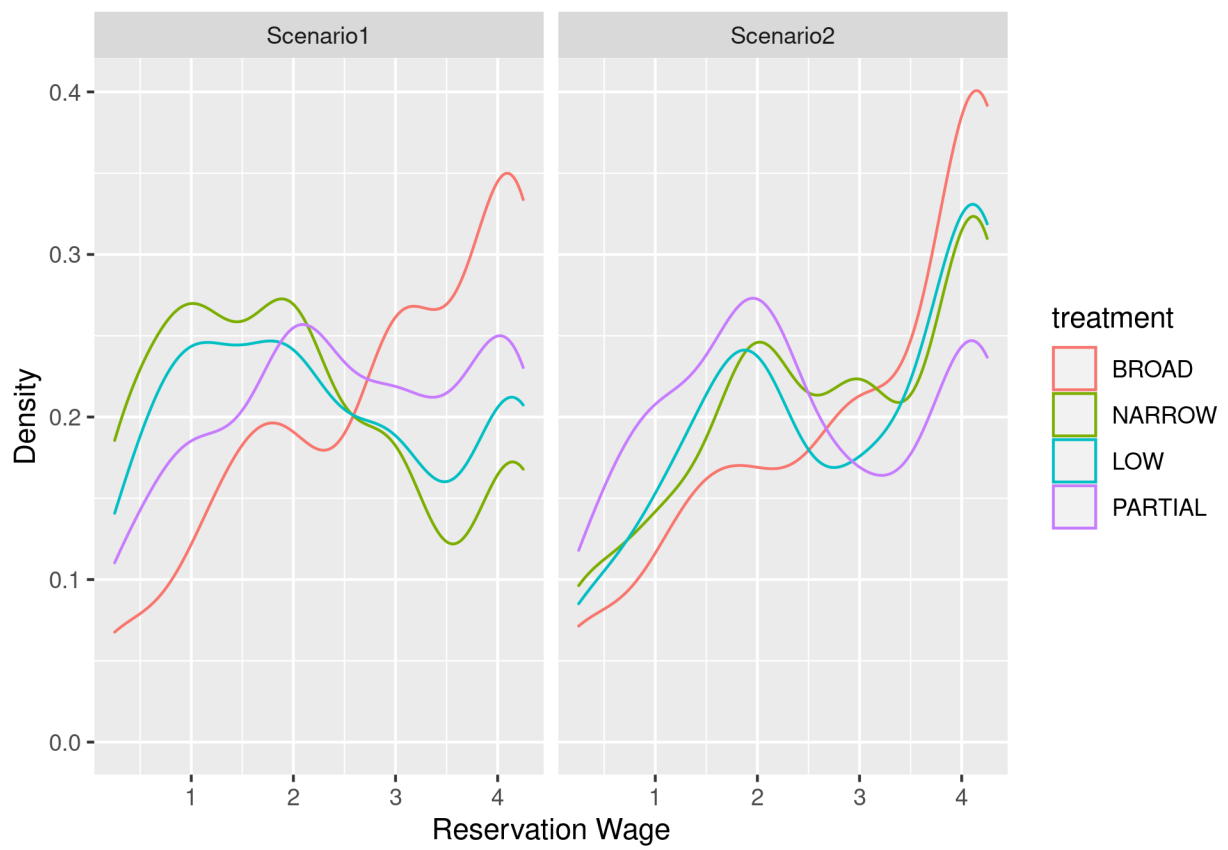


Table 17: Means of main treatments by scenario, restricting NARROW to sessions where baseline is revealed right before the first choice page.

Treatment	Res. Wage	Std Dev	% upper bound	N
Scenario 1				
BROAD	2.88	1.24	24%	137
NARROW	2.11	1.25	12%	117
LOW	2.31	1.33	18%	131
PARTIAL	2.50	1.28	15%	137
Scenario 2				
BROAD	2.99	1.29	33%	141
NARROW	2.85	1.24	25%	117
LOW	2.74	1.32	25%	135
PARTIAL	2.43	1.32	18%	138

Table 18: Between-treatment p-values for main treatments based on two-sided Wilcoxon rank-sum tests, treating each individual in each scenario as a single independent observation. Restricted to those sessions of NARROW where baseline is revealed right before the first choice page.

Treatments	BROAD	NARROW	LOW
Scenario 1			
NARROW	< 0.001		
LOW	< 0.001	0.298	
PARTIAL	0.012	0.015	0.209
Scenario 2			
NARROW	0.291		
LOW	0.108	0.566	
PARTIAL	< 0.001	0.011	0.054

Table 19: Between-treatment p-values for NARROW when information on baseline is presented for the first time right before the first choice or exactly on the first choice page. Based on two-sided Wilcoxon rank-sum tests, treating each individual in each scenario as a single independent observation.

Scenarios	Wilcoxon Test
Scenario 1	0.618
Scenario 2	0.002