Typecast: A Routine Mental Shortcut Causes Party Stereotyping

Supporting Information*

Douglas J. Ahler†
Gaurav Sood‡

January 23, 2022

*We thank the American Politics & Public Policy workshop at Yale University and the Political Behavior & Identities workshop at Duke University for participants' helpful questions, thoughts, and suggestions. We also appreciate our fellow panelists at MPSA 2017 for their feedback. Stephen Goggin, Marcel Garz, Brad Gomez, Bob Jackson, Kabir Khanna, Matt Levendusky, Matt Pietryka, Carrie Roush, and Julie Wronski also offered helpful comments. We are grateful to them.

†Doug is Assistant Professor of Political Science at the Florida State University, dahler@fsu.edu.

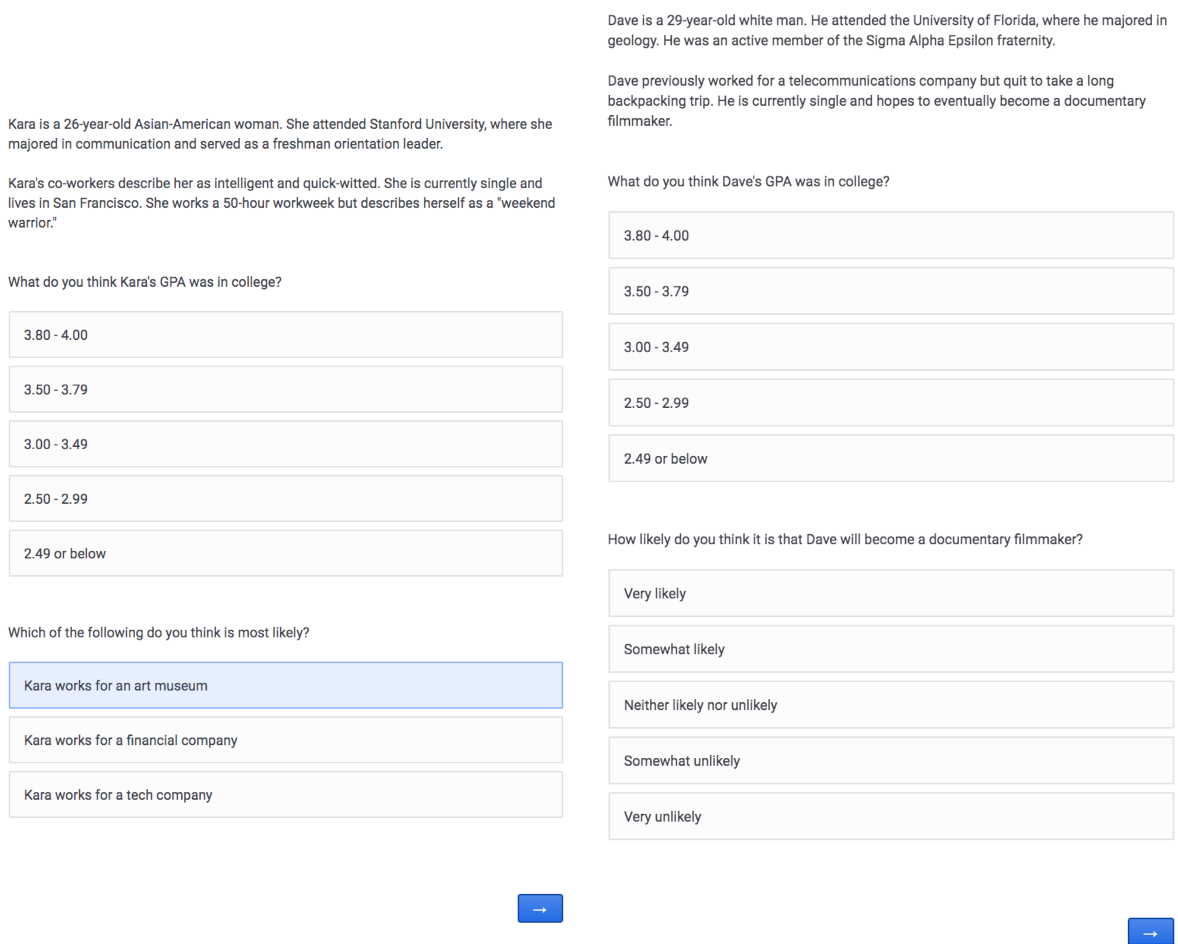
‡Gaurav can be reached at gsood07@gmail.com.

SI 1 Item Text and Survey Design

SI 1.1 Linda and James

SI 1.1.1 All vignettes from Experiment 1

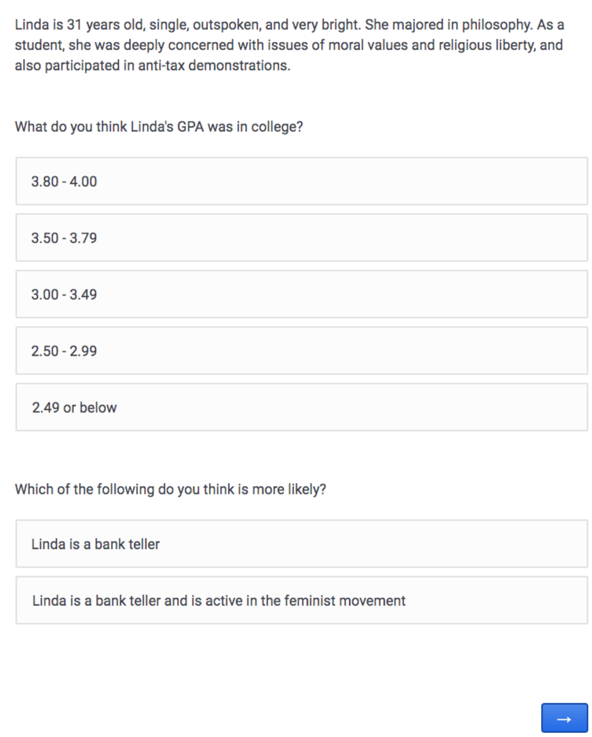
Figure SI 1.1: To preclude suspicion, James and Linda were couched as part of a study ostensibly on people's perceptions of the returns to higher education. These are the "Kara" and "Dave" vignettes that respondents saw first.



(a) (b)

Figure SI 1.2: Vignettes for our modified Linda Problem and the maximal-contrast James experiment

(a)

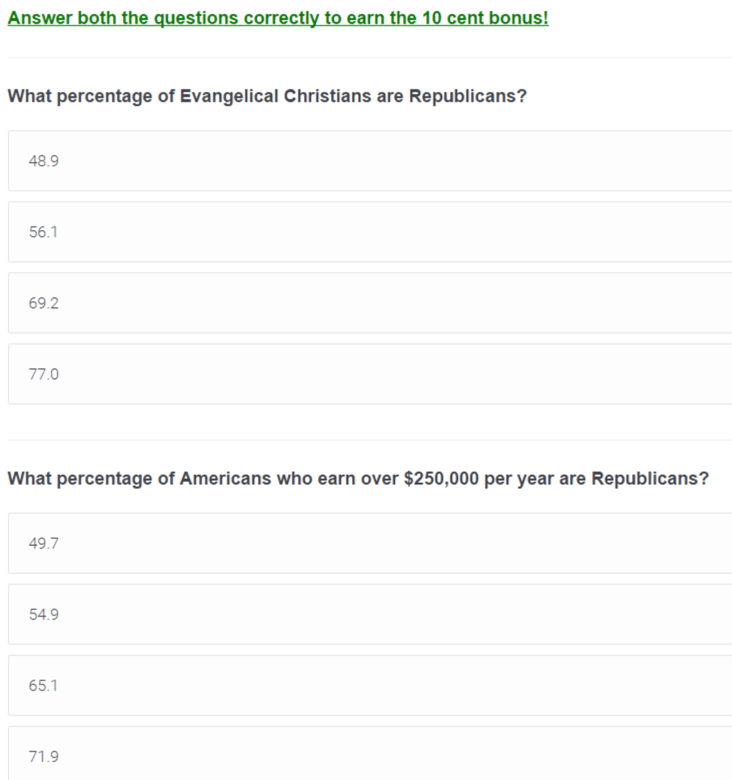


(b)

SI 1.2 Bayesian Perceptions

SI 1.2.1 Recall battery for treatment impact

Figure SI 1.3: Recall Treatment


Answer both the questions correctly to earn the 10 cent bonus!

What percentage of Evangelical Christians are Republicans?

48.9

56.1

69.2

77.0

What percentage of Americans who earn over $250,000 per year are Republicans?

49.7

54.9

65.1

71.9

SI 1.2.2 Numeracy Battery

The four items used for a numeracy battery are drawn from Weller et al. (2013). They are:

1. "Someone rolls a fair, six-sided die 1,000 times. On average, how many times would the die come up as an even number?" (Open-ended text-entry response)

2. "There is a 1% chance of winning a $10 prize in the Megabucks Lottery. On average, how many people would win the $10 prize if 1,000 people each bought a single ticket?" (Open-ended text-entry response)

3. "Which of the following numbers represents a bigger risk of getting a disease?"

    - 1 in 12
    - 1 in 37

4. "In the PCH Sweepstakes, the chances of winning a car are 1 in 1,000. What percent of PCH Sweepstakes tickets win a car?" (Open-ended text-entry response)

SI 2 Filtering protocol for the fully-factorial "James" experiment

At the time we fielded the second "James" study, a panic had broken out among experimental social scientists regarding data quality on Mechanical Turk (e.g., Ahler, Roush and Sood 2019; Bai 2018; Ryan 2018). In particular, concerns about bot respondents, foreign respondents, and "account farms" (locations with multiple individuals taking the same survey, or one individual with multiple accounts) masquerading as genuine survey-takers. We implemented a filtering protocol consistent with (Ahler, Roush and Sood 2019) to preclude these concerns. We recruited 1,991 respondents for our original sample. Using Laohaprapanon and Sood (2018), we discovered 359 IP addresses flagged for being outside the United States or on a known Blacklist, and an additional 106 duplicated addresses and 9 missing addresses. We classified these as suspicious responses. 87 respondents were also flagged for responding to multiple low-incidence screener questions used to identify potential survey "trolls" (Lopez and Hillygus 2018). In all—there was some overlap between the IP flags and the low-incidence screener flags—we found 484 respondents (24% of the sample) to be potentially low-quality, leaving us with our final sample of $n = 1{,}507$.

Although the suspicious respondents add noise to the data, they do not change our findings, either in terms of statistical or substantive significance. All of the treatments significantly affect perceptions of James—and the likelihood of committing the conjunction fallacy—as they do in the analysis limited to non-suspicious respondents. Furthermore, the cues about James's race and sexuality continue to be the strongest treatments.

SI 3 Model results: Fully-factorial "James" experiment

As described in the paper, for the fully-factorial "James" experiment our model takes the following form, with i indexing respondents and j indexing possible values of the dependent variable:

$$p_{ij} = p(y_i = j) = \begin{cases} p(y_i = -1) = p(y_i^* \leq \alpha_{-1}) \\ p(y_i = 0) = p(\alpha_{-1} < y_i^* \leq \alpha_0) \\ p(y_i = 1) = p(\alpha_0 < y_i^*) \end{cases} \tag{1}$$

where y_i^* is the respondent's latent outcome and α_{-1} and α_0 are the model's cutpoints. We model these probabilities as follows:

$$p(y_i = j) \sim \text{logit}^{-1}(\beta_k X_{ik} + \varepsilon_i) \tag{2}$$

where X_k denotes our treatment vector—James's race, sexual orientation, religion, and policy views. Importantly, because these these treatments were assigned randomly and independently of each other, β_k captures the unique effect of attribute k, on average and independent of all other treatments in K.

The results of the model are presented in Table SI 3.1. The coefficients, converted to changes in predicted probability of making the Democratic and Republican conjunction fallacies (depicted in Figure 3, are subsequently presented in Table SI 3.2.

Table SI 3.1: Full model results for the fully-factorial “James” experiment

	DV: Democratic (+1) or Republican (-1) conjunction fallacy
Black (vs. white)	-0.61^{***}
	(0.10)
Gay (vs. straight)	-0.82^{***}
	(0.10)
Evangelical (vs. nothing)	0.26^{**}
	(0.12)
Secular (vs. nothing)	-0.29^{**}
	(0.13)
Liberal (vs. nothing)	-0.41^{***}
	(0.13)
Conservative (vs. nothing)	0.36^{***}
	(0.12)
N	1507

$^{***}p < .01$; $^{**}p < .05$; $^{*}p < .1$

Table SI 3.2: Model coefficients converted to changes in predicted probabilities in the fully-factorial "James" experiment

When James is described as	Effect	Lower CI	Upper CI
Black (vs. white) Dem-CF	15.00	10.10	19.90
Black (vs. white) Rep-CF	-9.70	-12.84	-6.56
Conservative (vs. no cue) Dem-CF	-9.00	-15.08	-2.92
Conservative (vs. no cue) Rep-CF	6.00	1.88	10.12
Evangelical (vs. no cue) Dem-CF	-6.40	-12.28	-0.52
Evangelical (vs. no cue) Rep-CF	4.20	0.28	8.12
Gay (vs. straight) Dem-CF	20.10	15.20	25.00
Gay (vs. straight) Rep-CF	-13.20	-16.53	-9.87
Liberal (vs. no cue) Dem-CF	10.10	4.22	15.98
Liberal (vs. no cue) Rep-CF	-6.40	-10.12	-2.68
Secular (vs. no cue) Dem-CF	7.10	1.02	13.18
Secular (vs. no cue) Rep-CF	-4.50	-8.22	-0.78

SI 3.1 Marginals for the James experiment: How often do respondents make the conjunction fallacy?

An alternative way to present these results is simply to show how often respondents commit the conjunction fallacies when presented with party-representative traits of James. This is shown in the table below, which also implies why changes in predicted probabilities are a better way to assess the effects of this information. For reasons not entirely clear, respondents were far more likely to commit the Democratic conjunction fallacy—that is, to say that James is X and a Democrat—than to commit the equivalent Republican conjunction fallacy, sometimes even for counter-representative groups. (For example, when James is given a conservative cue, respondents are still more likely to commit the Democratic conjunction fallacy. However, the rate at which people commit the Republican conjunction fallacy goes up when James is described as conservative, and is in between these two quantities when James lacks an ideological cue.) We suspect that the Democratic conjunction fallacy is more popular for several reasons but, first and foremost, that James is described as college educated at a time when a serious degree divide has opened up between the two parties. Additionally, if anything, respondents appear ever so slightly more likely to attribute their own partisanship to James, all else equal.

One interesting thing that stands out is Democrats' and Republicans' tendencies to commit the conjunction fallacy at higher rates than independents, even while the rates they commit these errors are quite similar to each other—even when assessing the other party. This is relatively consistent with the results in Ahler & Sood (2018): everyone stereotypes, but partisans are more prone to do so than independents.

Table SI 3.3: Marginal rates at which conjunction fallacies (CFs) are committed, given particular traits of "James"

Group-CF dyad	Full sample	Democrats	Independents	Republicans
Black-Dem. CF	0.60	0.66	0.49	0.54
Black-Rep. CF	0.15	0.12	0.10	0.21
White-Dem. CF	0.49	0.50	0.38	0.49
White-Rep. CF	0.28	0.28	0.20	0.29
Gay-Dem. CF	0.63	0.68	0.48	0.60
Gay-Rep. CF	0.14	0.12	0.08	0.19
Straight-Dem. CF	0.46	0.49	0.38	0.44
Straight-Rep. CF	0.29	0.29	0.25	0.31
Evangelical-Dem. CF	0.49	0.52	0.38	0.48
Evangelical-Rep. CF	0.27	0.25	0.25	0.29
Secular-Dem. CF	0.61	0.67	0.48	0.55
Secular-Rep. CF	0.16	0.16	0.12	0.18
No relig. cue-Dem. CF	0.54	0.56	0.45	0.52
No relig. cue-Rep. CF	0.21	0.20	0.08	0.27
Liberal-Dem. CF	0.63	0.69	0.43	0.60
Liberal-Rep. CF	0.15	0.13	0.15	0.17
Conserv.-Dem. CF	0.47	0.49	0.35	0.46
Conserv.-Rep. CF	0.29	0.29	0.15	0.33
No ideo cue-Dem. CF	0.53	0.55	0.54	0.49
No ideo cue-Rep. CF	0.21	0.19	0.16	0.26

SI 4 The Ubiquity of Party Stereotypes?

But even when people are encouraged to process slowly and deliberately, their party stereotypes still yield cognitive bias. In August 2018, we conducted an experiment showing just this. We presented 138 research participants from Amazon's Mechanical Turk with the image in Figure SI 4.1a, requiring them to stay on that screen for 15 seconds before they could advance the survey. Figure SI 4.1a presents two groups of 25 avatars, one side in blue shirts and the other in red. The avatars differ in race, gender, and general appearance but, importantly, both "parties" contain the exact same 25 images. (We randomized both which side was the "Democratic" side and the order of avatars in the figure.) After advancing the screen, participants indicated the number of "people" in each party they saw who were: men or women, and Black, white, or another race/ethnicity, with a counter tool provided to help with summation to 25.

Figure SI 4.1: People are given identical images but report systematically different beliefs about what they saw

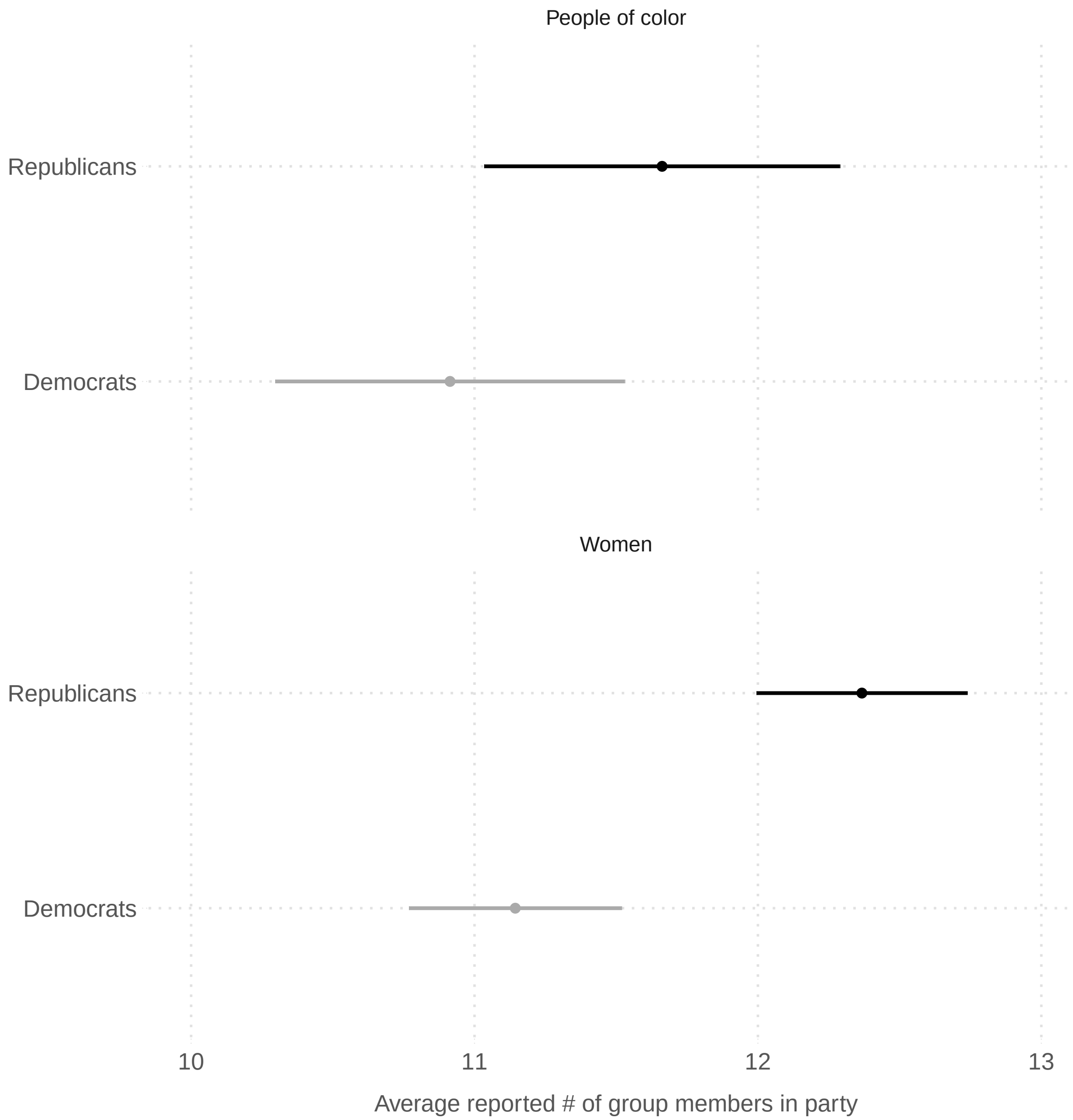
(a) What respondents saw

Based on a recent survey, this image represents the demographics of **Democratic Party supporters (in blue)** and **Republican Party supporters (in red)**.

In just a moment, you'll be able to advance the screen and answer a question about this image.

Please look carefully but **DO NOT take notes, take a screenshot, etc.**

(b) What respondents reported



If people's party stereotypes are not central to how they process political information, then we would expect participants to report equal numbers of women and non-white avatars on each side of the image in Figure SI 4.1a. Despite our alerting participants to having to answer questions about the images and requiring them to spend time on that screen, only 26.4% correctly identified the same number of women across images and just 25.8% correctly reported equal numbers of non-white avatars on both sides. On average, as Figure SI 4.1b shows, participants identified 1.2 more women and 0.8 more people of color wearing blue shirts than red shirts. Although we provided identical images and encouraged more thoughtful processing, respondents still erred systematically, and consistent with party stereotypes, while reporting what they saw.[1]

When most people process political information, they tend to do so haphazardly and automatically (Citrin and Green 1990; Lodge and Taber 2013; Sears 1983; Westen et al. 2006)—not in carefully controlled information environments guiding their processing, and usually handling more (and more complex) information than what we can convey with 50 avatars. Thus, we suspect that party stereotyping is not only an automatic process, but also a foremost reasoning device.

[1] One implication that might be tested in future work is whether respondents would perform better at this admittedly tough task if they were presented two sides separated by race or gender, with shirt color varying within each side. This would allow researchers to assess how effectively people can recognize $p(\text{party} \mid \text{group})$, the quantity that is more often presented in news and polling reports and which we argue contributes to the use of representativeness heuristics.

References

Ahler, Douglas J., Carolyn E. Roush and Gaurav Sood. 2019. The Micro-Task Market for Lemons: Data Quality on Amazon's Mechanical Turk. In *12th Annual NYU-CESS Experimental Political Science Conference*. New York: .

Bai, Hui. 2018. "Evidence that a Large Amount of Low Quality Responses on MTurk Can be Detected with Repeated GPS Coordinates." Available at https://www.maxhuibai.com/blog/evidence-that-responses-from-repeating-gps-are-random.

Citrin, Jack and Donald Philip Green. 1990. "The Self-Interest Motive in American Public Opinion." *Research in Micropolitics* 3(1):1–28.

Laohaprapanon, Suriyan and Gaurav Sood. 2018. "Know Your IP.". Available at https://github.com/themains/know_your_ip.

Lodge, Milton and Charles S. Taber. 2013. *The Rationalizing Voter*. New York: Cambridge University Press.

Lopez, Jesse and D. Sunshine Hillygus. 2018. Why So Serious? Survey Trolls and Misinformation. In *Annual Meeting of the Midwest Political Science Association*. Chicago: .

Ryan, Timothy J. 2018. "Data Contamination on MTurk." Available at http://timryan.web.unc.edu/2018/08/12/data-contamination-on-mturk/.

Sears, David O. 1983. "The Person-Positivity Bias." *Journal of Personality and Social Psychology* 44(2):233–250.

Weller, Joshua A., Nathan F. Dieckmann, Martin Tusler, C.K. Mertz, William J. Burns and Ellen Peters. 2013. "Development and Testing of an Abbreviated Numeracy Scale: A Rasch Analysis Approach." *Journal of Behavioral Decision Making* 26:198–212.

Westen, Drew, Pavel S. Blagov, Keith Harenski, Clint Kilts and Stephan Hamann. 2006. "Neural Bases of Motivated Reasoning: An fMRI Study of Emotional Constraints on Partisan Political Judgment in the 2004 U.S. Presidential Election." *Journal of Cognitive Neuroscience* 18(11):1947–1958.