

Online Supplemental Materials

A Meta-Analysis of Lineup Size Effects on Eyewitness Identification

Stefana Juncu & Ryan J. Fitzgerald

Psychology, Public Policy, and Law

2021

Receiver Operating Characteristics Analyses

Many previous analyses of eyewitness identifications used partial Receiver Operating Characteristic curve (partial ROCs). However, these include information about suspect identifications only and ignore filler identifications and rejections. Smith et al (2020) proposed the use of full ROC curves of investigator discriminability which incorporate data from all identification outcomes. We only have the necessary data to create these graphs from three studies (Akan et al., 2020; Juncu et al., in prep; Seale-Carlisle, et al., 2019). We categorised confidence as low (0-69%), medium (70-89%) and high (90-100%). Following Smith et al.'s (2020) example, we ordered each operating point based on a priori order: high-confidence suspect identifications, medium-confidence suspect identifications, low-confidence suspect identifications, low-confidence rejections, low-confidence filler identifications, medium-confidence rejections, medium-confidence filler identifications, high-confidence rejections, high-confidence filler identifications. If one condition dominates the other condition over the full range of false-positive rates, we know that the dominating procedure will always be superior. However, if the two curves intersect, expected utility analyses should be used as a follow-up (Smith et al., 2020).

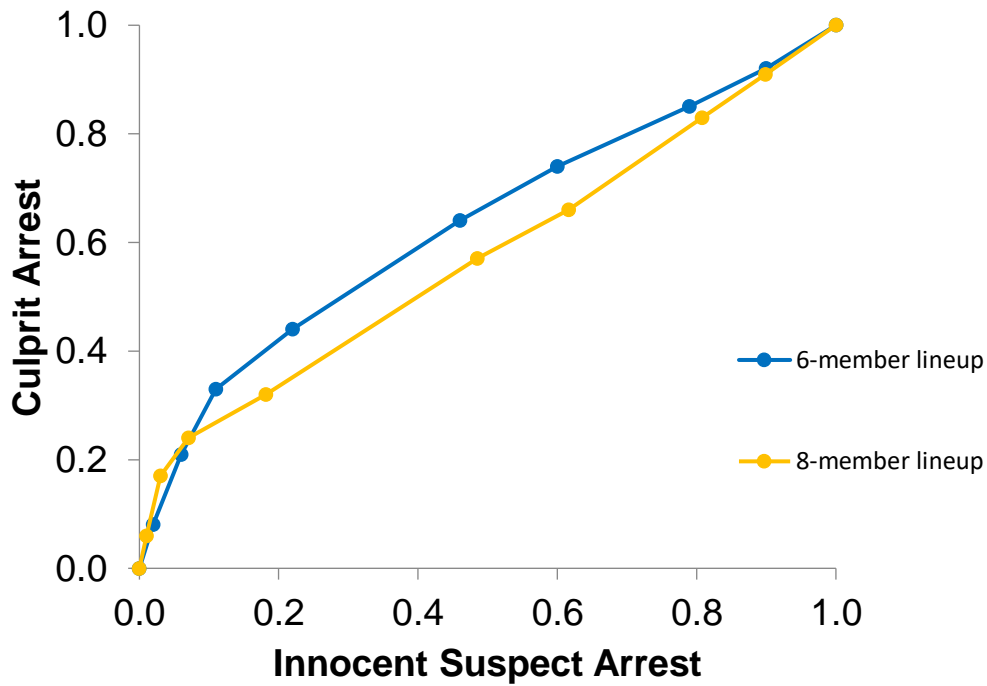


Figure S1. Full receiver operating characteristic (ROC) comparison of the 6-member and 8-member lineups from Seal-Carlisle et al. (2019).

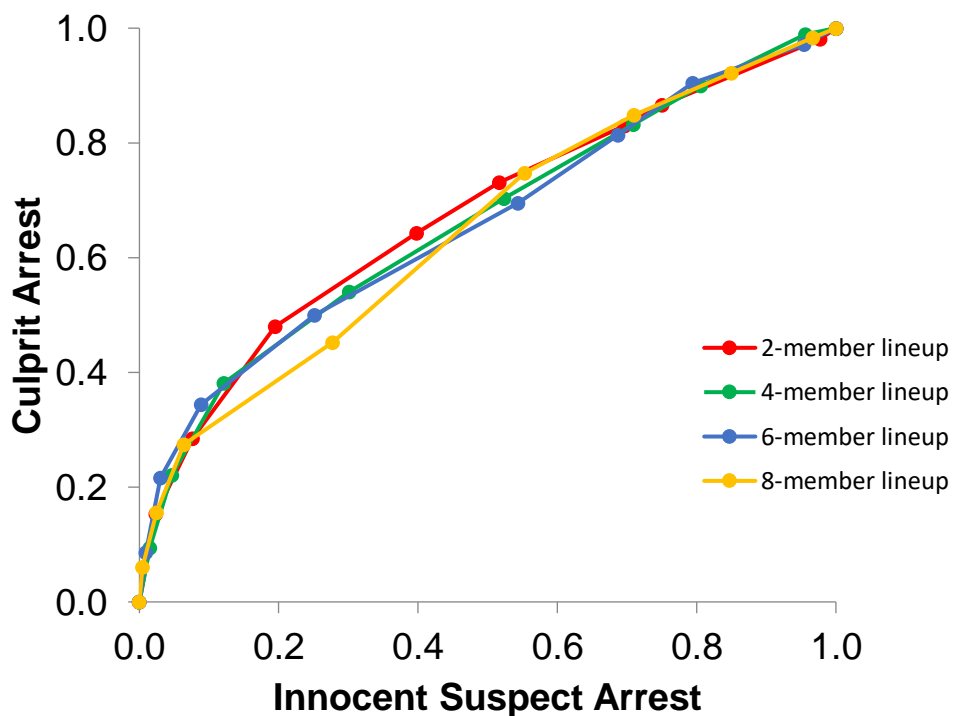


Figure S2. Full receiver operating characteristic (ROC) comparison of the 2-member, 4-member, 6-member and 8-member lineups from Akan et al. (2020).

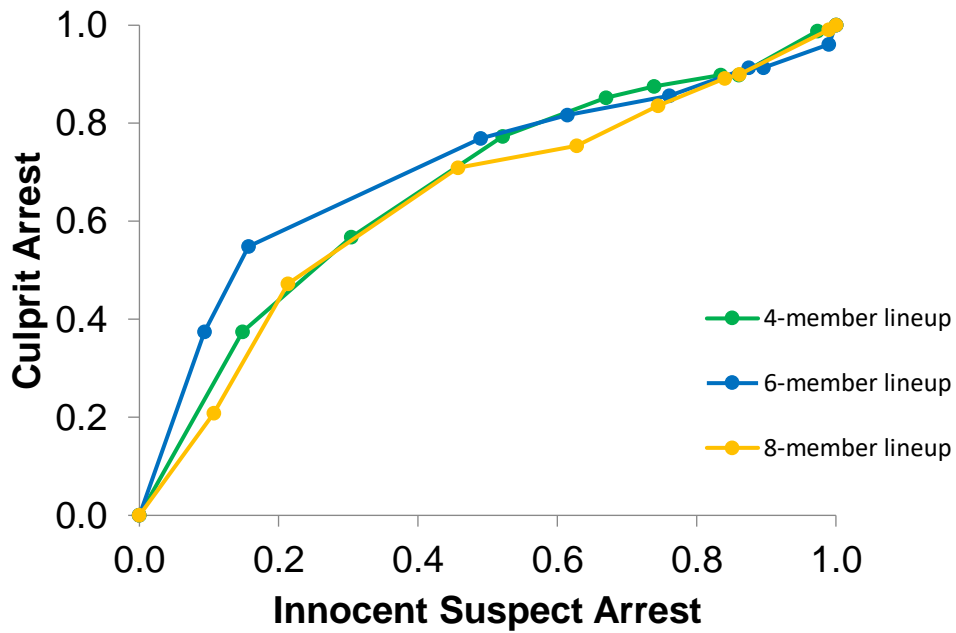


Figure S3. Full receiver operating characteristic (ROC) comparison of the 4-member, 6-member and 8-member low similarity lineups from Juncu and Fitzgerald (in prep).

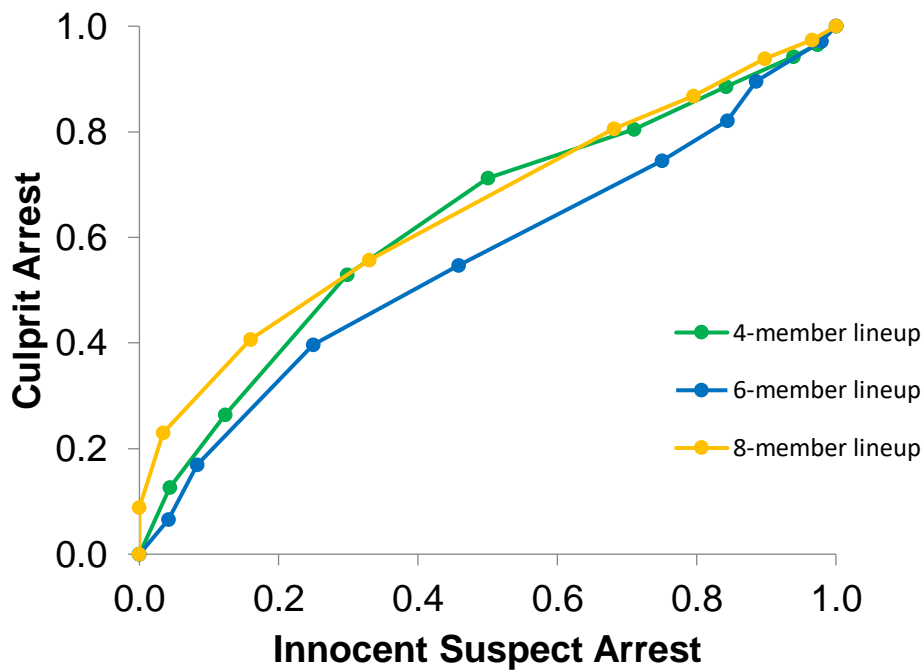


Figure S4. Full receiver operating characteristic (ROC) comparison of the 4-member, 6-member and 8-member high similarity lineups from Juncu and Fitzgerald (in prep).

Confidence-Accuracy Characteristics curves



Figure S5. Confidence-Accuracy Characteristics curves (CACs) comparing suspect identification accuracy when using 6-member and 8-member lineups at three levels of confidence. We categorised confidence as low (0-69%), medium (70-89%) and high (90-100%). Data from Seal-Carlisle et al. (2019).



Figure S6. Confidence-Accuracy Characteristics curves (CACs) comparing suspect identification accuracy at three levels of confidence. We categorised confidence as low (0-69%), medium (70-89%) and high (90-100%). Data from Akan et al. (2020).

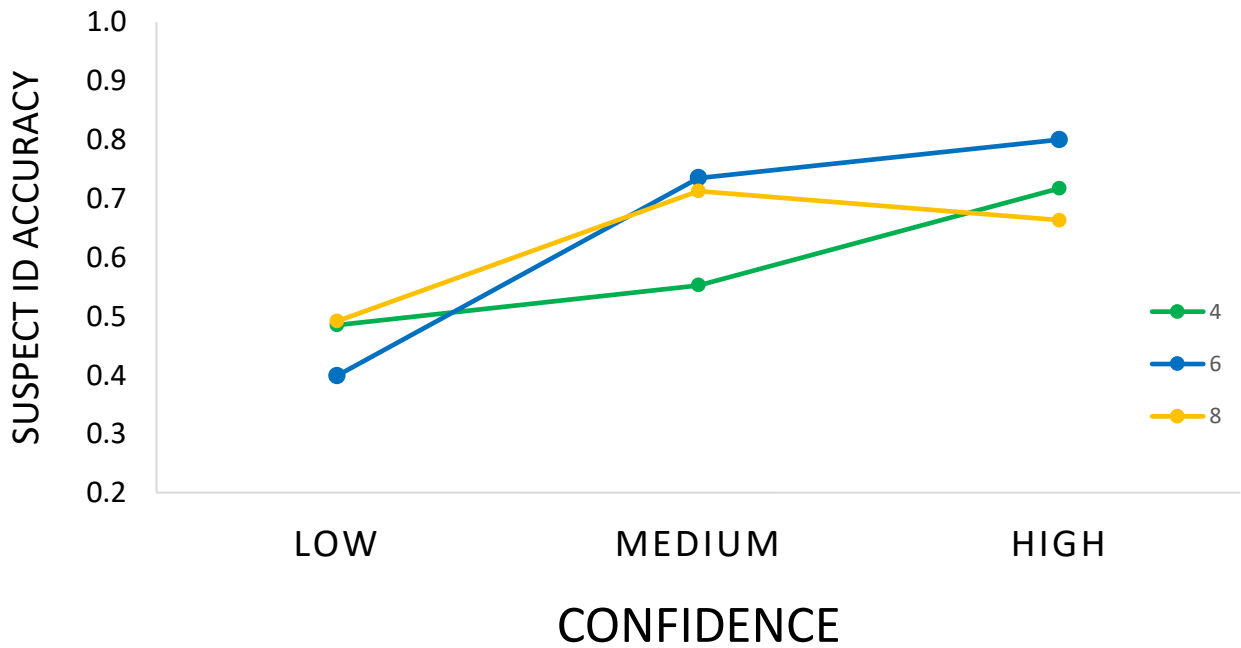


Figure S7. Confidence-Accuracy Characteristics curves (CACs) comparing suspect identification accuracy at three levels of confidence. We categorised confidence as low (0-69%), medium (70-89%) and high (90-100%). Data from Juncu & Fitzgerald (in prep), unfair lineup condition.

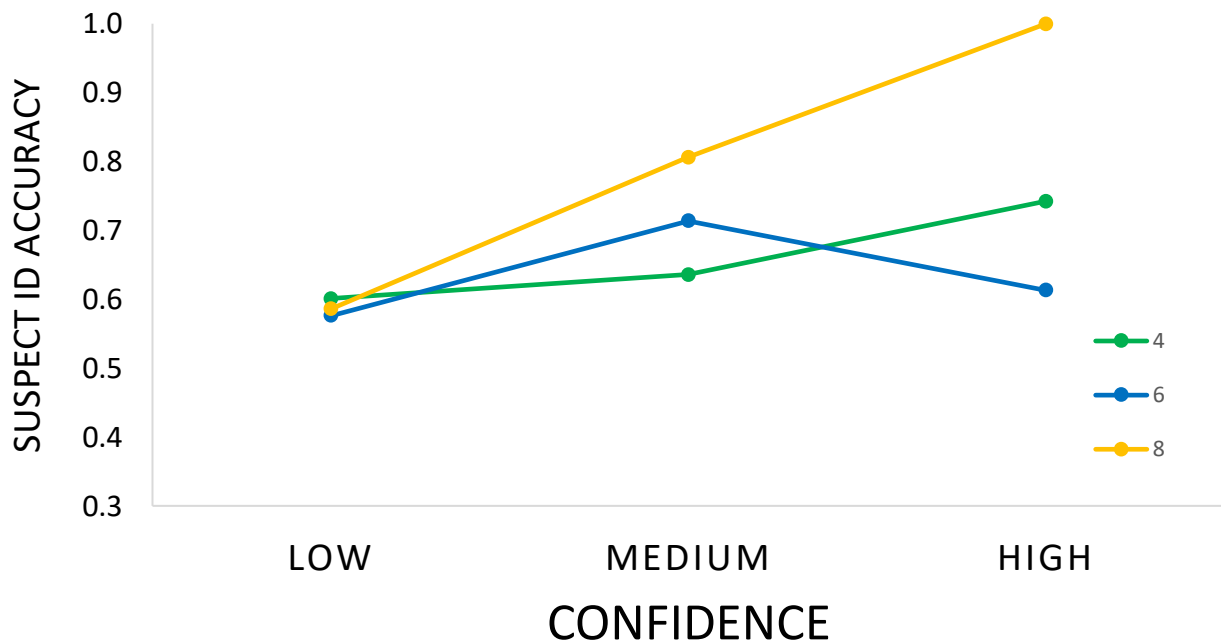


Figure S8. Confidence-Accuracy Characteristics curves (CACs) comparing suspect identification accuracy at three levels of confidence. We categorised confidence as low (0-

69%), medium (70-89%) and high (90-100%). Data from Juncu & Fitzgerald (in prep), fair lineup condition.

Expected cost analysis

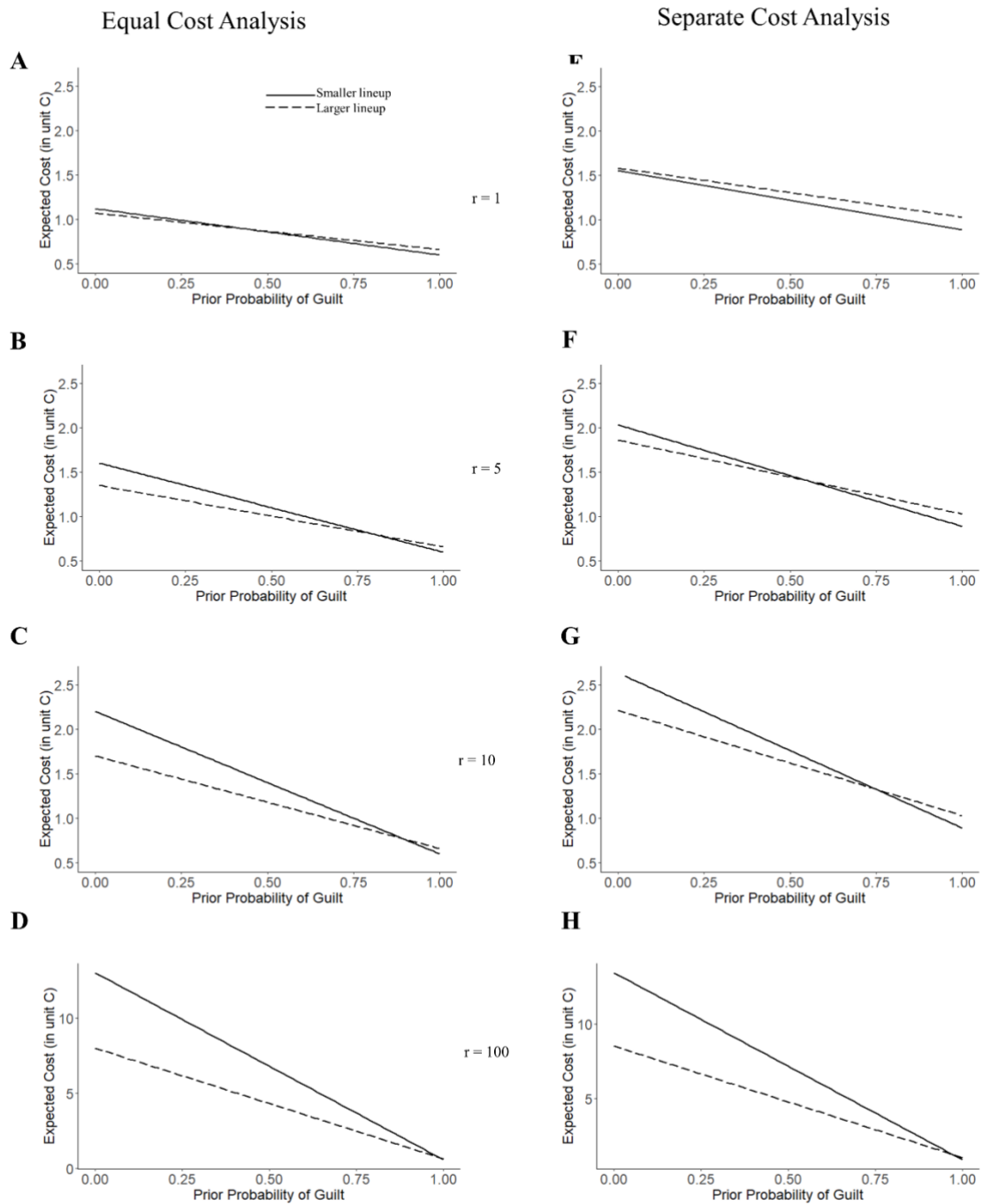


Figure S9. The expected costs of increasing lineup size, assuming equal costs for filler identifications and rejections (left

panels: A-D) or assuming separate costs for filler identifications and rejections (right panel: E-H). The two slanted lines represent the expected costs of smaller lineups (solid line) and larger lineups (dashed line). r represents the ratio between the cost of incriminating an innocent suspect and the cost of failing to incriminate a guilty suspect.

Funnel plots

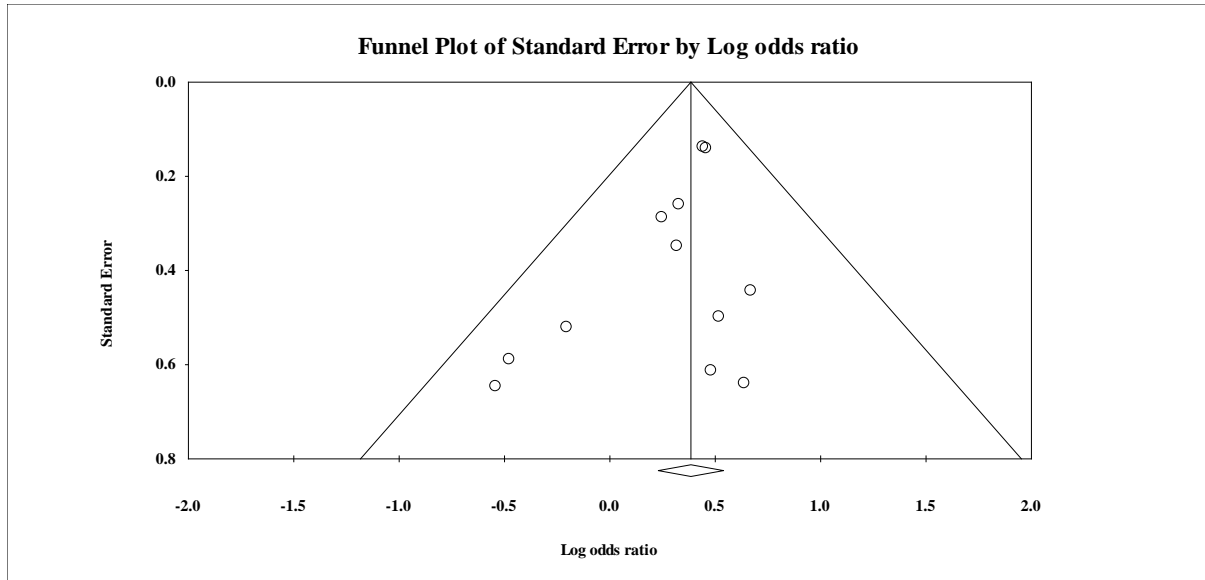


Figure S10. Funnel plot of the meta-analysis of correct suspect identifications. Each plotted point represents the standard error and log odds ratios of the difference between smaller and larger lineups for a single experiment.

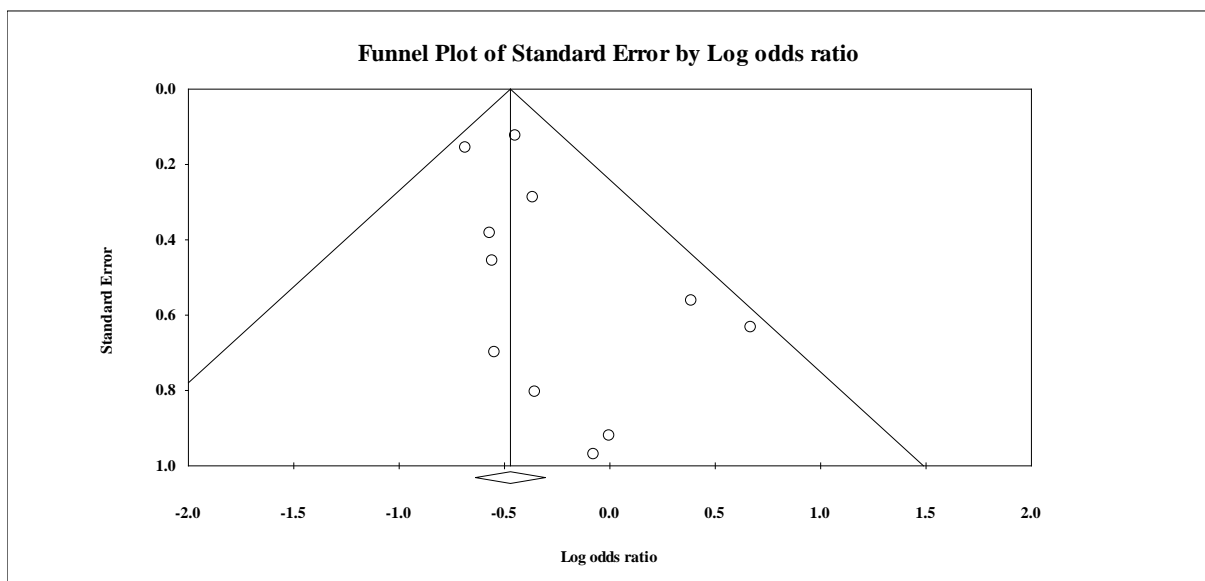


Figure S11. Funnel plot of the meta-analysis of filler identifications from target present lineups. Each plotted point represents the standard error and log odds ratios of the difference between smaller and larger lineups for a single experiment.

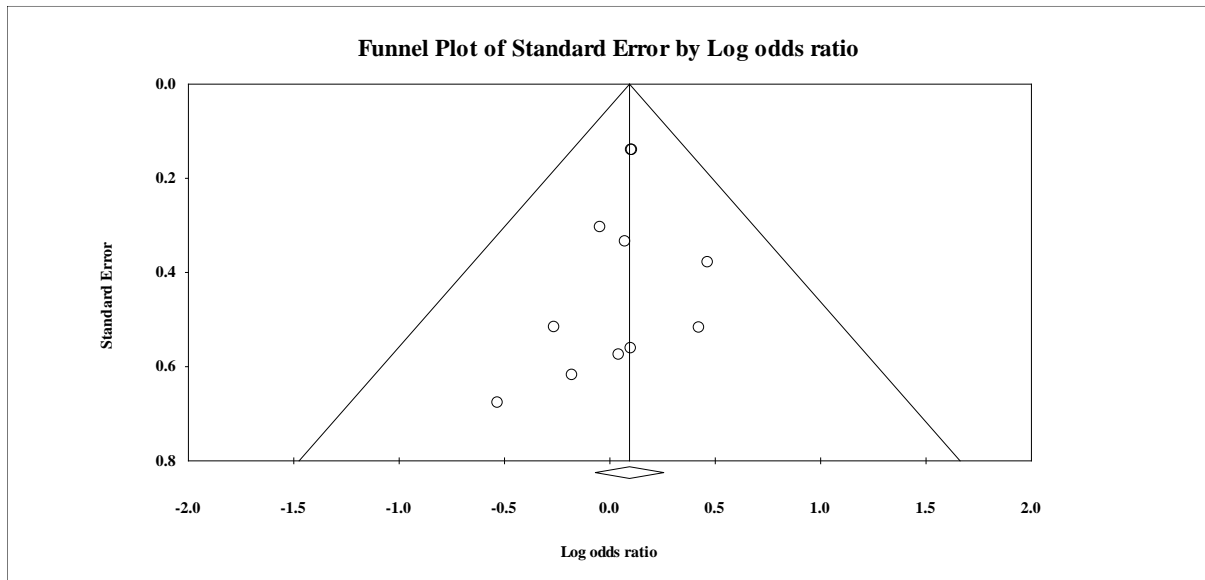


Figure S12. Funnel plot of the meta-analysis of incorrect rejection from target present lineups. Each plotted point represents the standard error and log odds ratios of the difference between smaller and larger lineups for a single experiment.

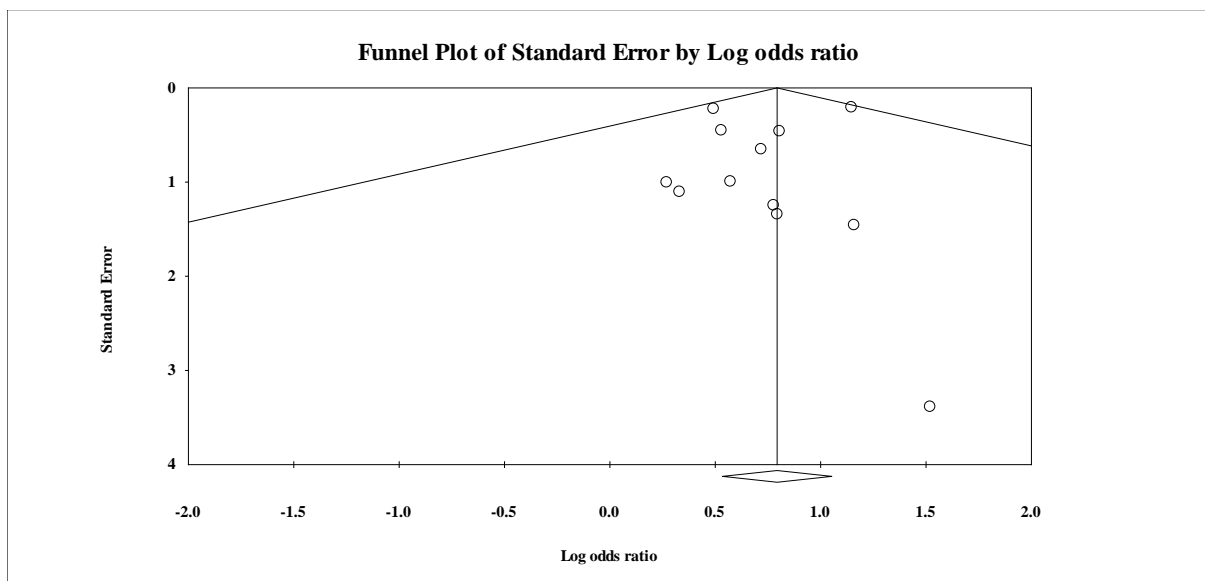


Figure S13. Funnel plot of the meta-analysis of incorrect suspect identifications. Each plotted point represents the standard error and log odds ratios of the difference between smaller and larger lineups for a single experiment.

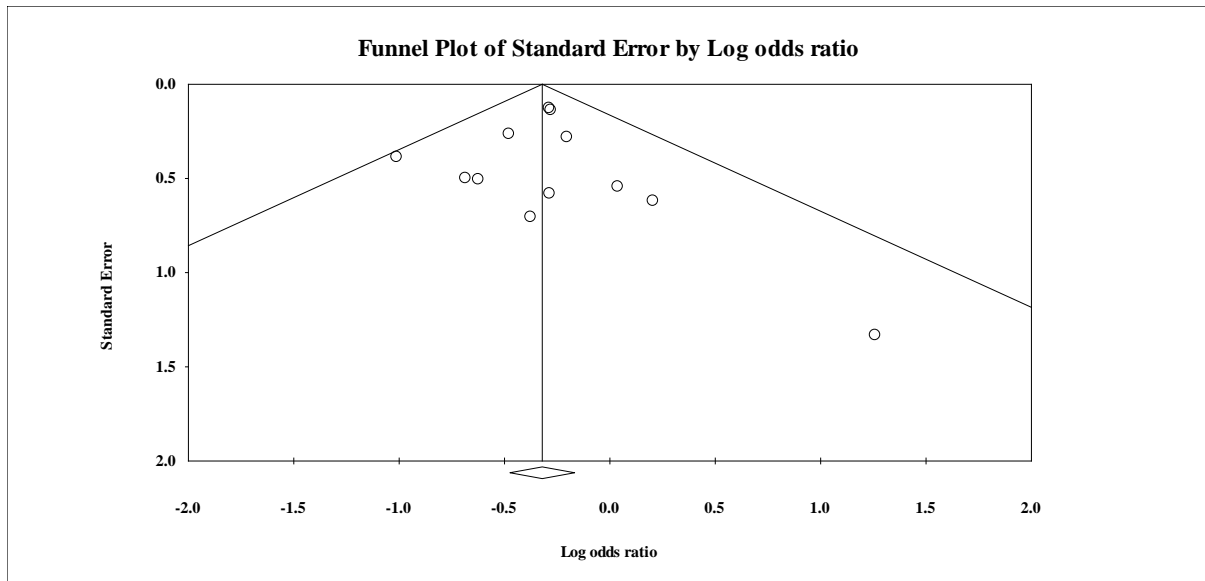


Figure S14. Funnel plot of the meta-analysis of filler identifications from target absent lineups. Each plotted point represents the standard error and log odds ratios of the difference between smaller and larger lineups for a single experiment.

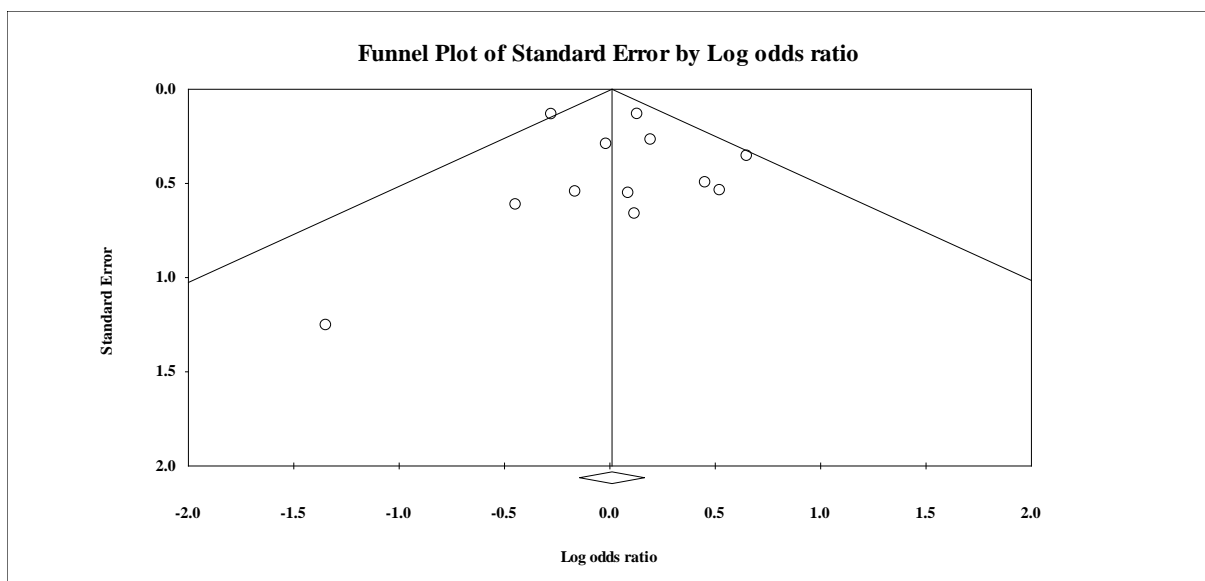


Figure S15. Funnel plot of the meta-analysis of correct rejection of target absent lineups. Each plotted point represents the standard error and log odds ratios of the difference between smaller and larger lineups for a single experiment.

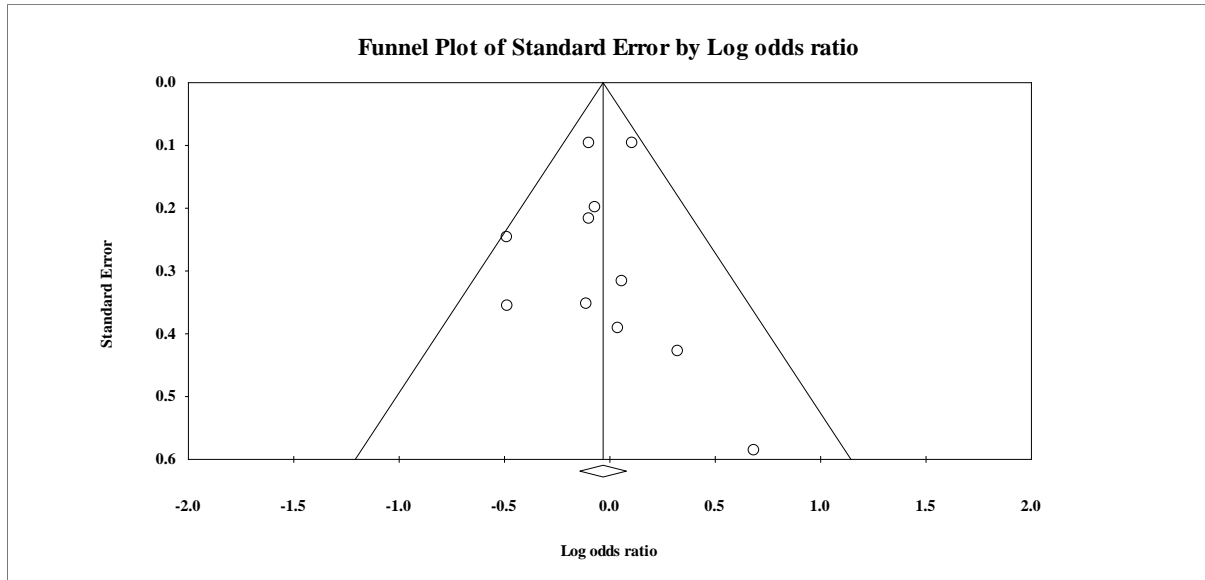


Figure S16. Funnel plot of the meta-analysis of choosing from both target present and target absent lineups. Each plotted point represents the standard error and log odds ratios of the difference between smaller and larger lineups for a single experiment.

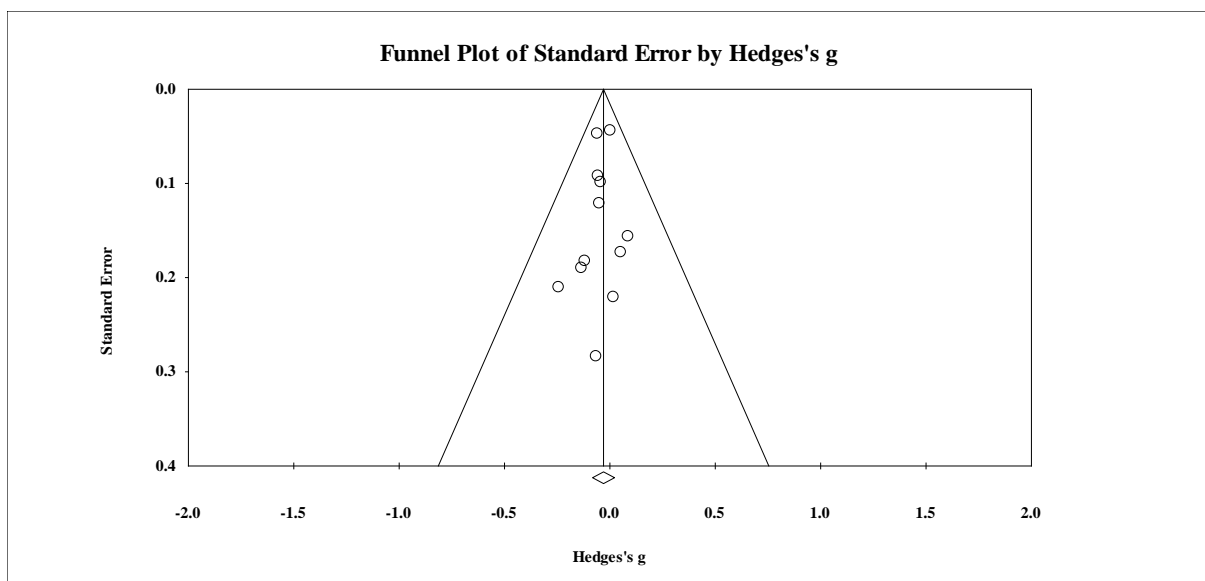


Figure S17. Funnel plot of the meta-analysis of discriminability. Each plotted point represents the standard error and standardized mean difference between smaller and larger lineups for a single experiment.

Sensitivity analyses

Table S1. *The Effects of Lineup Size on Identification Choices based on study design variable*

Culprit	Lineup Outcome	Sample	<i>m</i>	<i>k</i>	Effect Size & 95% CIs			Test of the Null			
					<i>ES</i>	<i>LL</i>	<i>UL</i>	<i>t</i>	<i>df</i>	<i>p</i>	
Present	Suspect	Entire sample	12	104	1.44	1.27	1.52	6.27	11	<.001	
		Counterbalance	6	37	1.54	1.46	1.64	19.80	5	<.001	
		Simultaneous	10	87	1.36	1.12	1.65	3.58	9	.006	
		Excluding Meissner et al.	11	89	1.42	1.24	1.63	5.64	10	<.001	
	Filler	Entire sample	11	86	0.64	0.47	0.86	-3.32	10	.008	
		Counterbalance	5	22	0.50	0.35	0.72	-5.41	4	.006	
		Simultaneous	10	84	0.68	0.45	1.04	-2.05	9	.071	
	Rejection	Entire sample	11	86	1.10	1.03	1.17	3.33	10	.008	
		Counterbalance	5	22	1.12	1.03	1.21	3.99	4	.016	
		Simultaneous	11	84	1.10	1.02	1.18	3.02	10	.013	
	Absent	Suspect	Entire sample	12	99	1.85	1.61	2.15	9.38	11	<.001
			Counterbalance	6	37	1.84	1.51	2.25	7.90	5	<.001
Simultaneous			11	81	1.99	1.80	2.19	15.60	9	<.001	
Excluding Meissner et al.			11	84	1.86	1.60	2.15	9.24	10	<.001	
Filler		Entire sample	12	99	0.64	0.53	0.81	-4.27	11	.001	
		Counterbalance	6	37	0.59	0.41	0.84	-3.81	5	.013	
		Simultaneous	10	82	0.65	0.49	0.86	-3.46	9	.007	
		Excluding Meissner et al.	11	84	0.65	0.52	0.83	-4.06	10	.002	
Rejection		Entire sample	12	99	1.21	1.07	1.38	3.33	11	.007	
		Counterbalance	6	37	1.26	1.04	1.53	3.13	5	.026	
		Simultaneous	10	82	1.23	1.02	1.48	2.48	9	.035	
		Excluding Meissner et al.	11	84	1.21	1.06	1.39	3.09	10	.011	
Both	Choosing	Entire sample	11	81	0.88	0.81	0.96	-3.20	10	.009	
		Counterbalance	5	22	0.85	0.74	0.96	-3.57	4	.023	

	Simultaneous	10	79	0.90	0.77	1.05	-1.52	9	.162
Discriminability	Entire sample	12	99	-0.03	-0.07	0.01	-1.77	11	.105
	Counterbalance	6	37	-0.02	-0.06	0.02	-1.16	5	.299
	Simultaneous	10	82	-0.04	-0.06	-0.03	-5.90	9	<.001
	Excluding Meissner et al.	11	84	-0.03	-0.07	0.01	-1.76	10	.108

Note. *m* = number of studies, *k* = number of effect sizes, *ES* = Effect Size; *CI* = Confidence Interval; *LL* = Lower Limit; *UL* = Upper Limit. For discriminability the effect size is Hedge's *g*. For all other outcomes, the effect sizes are odds ratio

Publication status

Table S2

Regression Coefficients for Moderator Analysis of Publication Status and Year

Culprit	Lineup choice	Effect type	Effect size & 95% CIs			Test of the NULL		
			ES	LL	UL	<i>t</i>	<i>df</i>	<i>p</i>
Present	Suspect	Publication status	0.08	-0.36	0.52	0.41	10	.69
	Filler	Publication status	-0.22	-1.04	0.60	-0.62	9	.55
	Rejection	Publication status	0.09	-0.10	0.27	1.09	9	.31
Absent	Suspect	Publication status	0.17	-0.05	0.40	1.72	10	.12
	Filler	Publication status	-0.30	-0.78	0.18	-1.38	10	.20
	Rejection	Publication status	0.23	-0.10	0.55	1.55	10	.15
Both	Choosing	Publication status	-0.18	-0.41	0.05	-1.78	9	.11
	Discriminability	Publication status	0.01	-0.11	0.11	0.01	10	.98

Note. *ES* = Effect Size; *CI* = Confidence Interval; *LL* = Lower Limit; *UL* = Upper Limit