

Supplemental Material: Stimuli and Lineup Construction

Table S1

Stimulus Sets 1–6. Video and Target Descriptions

Stimulus set	Video content	Video length	Target description
1	Male target breaking into an apartment and stealing a laptop	28 s; target's face in view for 8 s (mostly close-up)	Caucasian male; brown hair; average build, middle age, moustache
2	Male target kicking over a rubbish bin	14 s; target's face in view for 10 s (close-up for approx. 5 s)	Caucasian male; short brown hair, tall, medium build, 20s
3	Male target breaking into and stealing a car	24 s; target's face in view for 13 s (close-up for approx. 5 s)	Caucasian male; balding; grey hair; 60s; moustache
4	Female target stealing unknown object (not in view) from an open window	14 s; target's face in view for 4 s (2 side-on and 2 front-on – all close-up)	Caucasian female; shoulder length dark brown hair with a fringe; slim; 20s
5	Female target stealing a mobile phone from a café table	23 s; target's face in view for 10 s (close-up for approx. 4 s)	Caucasian female; long blonde hair; 20s
6	Male target stealing item from a clothing store	24 s; target in view for 14s (4 s close-up)	Caucasian male; short brown hair; medium build; 20s; round face

Table S2

Stimulus Set 1. Mean (and Standard Deviation) Similarity Ratings of the Designated High and Low Similarity Fillers to the Target (T) and Innocent Suspect (IS)

Filler similarity category	Lineup member					
	T or IS	Filler 1	Filler 2	Filler 3	Filler 4	Filler 5
Rating to target						
High	2.80 (1.84)	3.23 (1.93)	3.20 (1.80)	2.72 (2.07)	2.98 (1.83)	3.05 (1.81)
Low	-	1.40 (1.42)	1.58 (1.62)	1.47 (1.36)	1.52 (1.37)	1.50 (1.49)
Rating to innocent suspect						
High	3.86 (1.59)	3.00 (1.74)	4.32 (1.56)	3.77 (1.53)	3.74 (1.79)	4.39 (1.52)
Low	-	2.65 (1.69)	3.07 (1.85)	2.43 (1.70)	3.45 (1.94)	2.24 (1.81)

Notes: Rating scale 0–7 (0 = *very low similarity*; 7 = *very high similarity*)

Table S3

Stimulus Set 2. Mean (and Standard Deviation) Similarity Ratings of the Designated High and Low Similarity Fillers to the Target (T) and Innocent Suspect (IS)

Filler similarity category	Lineup member					
	T or IS	Filler 1	Filler 2	Filler 3	Filler 4	Filler 5
Rating to target						
High	2.25 (1.56)	2.13 (1.77)	2.12 (1.82)	2.18 (1.73)	2.40 (1.68)	2.43 (1.89)
Low	-	0.98 (1.07)	0.82 (1.20)	0.92 (1.33)	1.05 (1.47)	0.93 (1.16)
Rating to innocent suspect						
High	2.51 (1.85)	2.65 (1.75)	3.30 (1.62)	4.44 (1.63)	5.10 (1.26)	2.91 (1.95)
Low	-	2.72 (1.78)	2.66 (1.76)	2.32 (1.64)	1.87 (1.76)	3.07 (1.69)

Notes: Rating scale 0–7 (0 = *very low similarity*; 7 = *very high similarity*)

Table S4

Stimulus Set 3. Mean (and Standard Deviation) Similarity Ratings of the Designated High and Low Similarity Fillers to the Target (T) and Innocent Suspect (IS)

Filler similarity category	Lineup member					
	T or IS	Filler 1	Filler 2	Filler 3	Filler 4	Filler 5
Rating to target						
High	2.80 (1.82)	2.95 (1.93)	2.80 (1.90)	2.85 (1.74)	3.02 (1.92)	2.87 (1.98)
Low	-	1.25 (1.39)	1.22 (1.39)	1.17 (1.30)	1.38 (1.34)	1.18 (1.42)
Rating to innocent suspect						
High	3.44 (1.68)	3.76 (1.83)	3.48 (1.69)	4.1 (1.74)	3.80 (1.80)	2.91 (1.91)
Low	-	2.34 (1.72)	2.77 (1.68)	3.35 (1.78)	2.82 (1.92)	2.45 (1.71)

Notes: Rating scale 0–7 (0 = *very low similarity*; 7 = *very high similarity*)

Table S5

Stimulus Set 4. Mean (and Standard Deviation) Similarity Ratings of the Designated High and Low Similarity Fillers to the Target (T) and Innocent Suspect (IS)

Filler similarity category	Lineup member					
	T or IS	Filler 1	Filler 2	Filler 3	Filler 4	Filler 5
Rating to target						
High	3.37 (1.85)	3.37 (1.81)	3.55 (2.03)	3.67 (1.93)	3.57 (1.78)	3.43 (2.01)
Low	-	2.05 (1.60)	2.13 (1.81)	2.28 (1.68)	2.20 (1.54)	2.03 (1.54)
Rating to innocent suspect						
High	4.86 (1.50)	3.77 (1.87)	3.34 (1.66)	4.13 (1.76)	4.49 (1.66)	4.57 (1.72)
Low	-	3.82 (1.90)	3.16 (1.75)	2.55 (1.77)	4.72 (1.45)	3.68 (1.80)

Notes: Rating scale 0–7 (0 = *very low similarity*; 7 = *very high similarity*)

Table S6

Stimulus Set 5. Mean (and Standard Deviation) Similarity Ratings of the Designated High and Low Similarity Fillers to the Target (T) and Innocent Suspect (IS)

Filler similarity category	Lineup member					
	T or IS	Filler 1	Filler 2	Filler 3	Filler 4	Filler 5
Rating to target						
High	3.17 (2.03)	3.58 (2.06)	3.13 (1.98)	3.32 (1.99)	3.47 (1.96)	3.55 (1.84)
Low	-	1.72 (1.55)	1.87 (1.52)	1.82 (1.60)	1.92 (1.69)	1.63 (1.29)
Rating to innocent suspect						
High	3.74 (1.68)	4.20 (1.65)	3.97 (1.64)	4.10 (1.54)	3.87 (1.64)	4.63 (1.64)
Low	-	2.41 (1.65)	3.43 (1.82)	2.98 (1.91)	3.43 (1.65)	3.14 (1.63)

Notes: Rating scale 0–7 (0 = *very low similarity*; 7 = *very high similarity*)

Table S7

Stimulus Set 6. Mean (and Standard Deviation) Similarity Ratings of the Designated High and Low Similarity Fillers to the Target (T) and Innocent Suspect (IS)

Filler similarity category	Lineup member					
	T or IS	Filler 1	Filler 2	Filler 3	Filler 4	Filler 5
Rating to target						
High	2.72 (1.75)	2.72 (1.89)	2.77 (1.66)	2.75 (1.83)	3.03 (1.89)	2.95 (1.93)
Low	-	1.58 (1.76)	1.72 (1.46)	1.55 (1.49)	1.63 (1.43)	1.78 (1.50)
Rating to innocent suspect						
High	3.79 (1.80)	3.78 (1.95)	2.72 (1.75)	3.58 (1.58)	3.71 (1.83)	4.30 (1.69)
Low	-	2.51 (1.80)	3.79 (1.74)	4.26 (1.71)	2.99 (1.87)	3.61 (1.68)

Notes: Rating scale 0–7 (0 = *very low similarity*; 7 = *very high similarity*)

Table S8

Stimulus Set 1. Proportion of Mock-Witness Suspect Picks, Suspect Pick 95% Confidence

Interval (CI) and Difference from Chance Estimate for each Lineup

High similarity fillers	N	Suspect bias stats		
		Suspect picks	Suspect picks 95% CI	Difference from chance
Target-present				
0	58	.43	.30 – .58	4.07**
1	41	.27	.13 – .40	1.47
2	49	.29	.16 – .41	1.85
3	51	.33	.20 – .46	2.53*
4	54	.28	.16 – .40	1.82
5	52	.21	.10 – .32	0.79
Target-absent				
0	45	.18	.07 – .29	0.20
1	49	.31	.19 – .45	2.17*
2	47	.32	.19 – .45	2.24*
3	55	.24	.12 – .35	1.22
4	57	.18	.08 – .27	0.17
5	41	.10	.01 – .19	-1.49

*significantly different from chance at $p < .05$

**significantly different from chance at $p < .01$

Table S9

Stimulus Set 2. Proportion of Mock-Witness Suspect Picks, Suspect Pick 95% Confidence

Interval (CI) and Difference from Chance Estimate for each Lineup

High similarity fillers	N	Suspect bias stats		
		Suspect picks	Suspect picks 95% CI	Difference from chance
Target-present				
0	39	.23	.10 – .36	0.95
1	55	.27	.16 – .39	1.77
2	51	.06	-.01 – .12	-3.27**
3	51	.18	.07 – .28	0.18
4	49	.06	-.01 – .13	-3.08**
5	47	.11	.02 – .20	-1.34
Target-absent				
0	60	.25	.14 – .36	1.49
1	45	.09	.01 – .17	-1.83
2	48	.19	.08 – .30	0.37
3	53	.17	.07 – .28	0.06
4	46	.04	-0.02 – .10	-4.10**
5	55	.11	.03 – .19	-1.37

*significantly different from chance at $p < .05$

**significantly different from chance at $p < .01$

Table S10

Stimulus Set 3. Proportion of Mock-Witness Suspect Picks, Suspect Pick 95% Confidence

Interval (CI) and Difference from Chance Estimate for each Lineup

High similarity fillers	N	Suspect bias stats		
		Suspect picks	Suspect picks 95% CI	Difference from chance
Target-present				
0	37	.11	.01 – .21	-1.15
1	64	.02	-.01 – .07	-6.23**
2	66	.05	-.01 – .10	-4.73**
3	50	.02	-.02 – .06	-7.41**
4	39	.03	-.02 – .08	-5.57**
5	53	.04	-.01 – .09	-4.93**
Target-absent				
0	38	.26	.12 – .40	1.35
1	58	.22	.12 – .33	1.05
2	53	.15	.06 – .25	-0.32
3	48	.17	.06 – .27	0
4	45	.27	.14 – .40	1.52
5	48	.44	.30 – .58	3.78**

*significantly different from chance at $p < .05$

**significantly different from chance at $p < .01$

Table S11

Stimulus Set 4. Proportion of Mock-Witness Suspect Picks, Suspect Pick 95% Confidence

Interval (CI) and Difference from Chance Estimate for each Lineup

High similarity fillers	N	Suspect bias stats		
		Suspect picks	Suspect Picks 95% CI	Difference from chance
Target-present				
0	47	.36	.22 – .50	2.78**
1	53	.19	.08 – .29	0.41
2	48	.17	.06 – .27	0
3	36	.33	.18 – .49	2.12*
4	57	.28	.16 – .40	1.92
5	55	.20	.09 – .31	0.62
Target-absent				
0	45	.24	.12 – .37	1.21
1	47	.13	.03 – .22	-0.80
2	49	.06	-.01 – .13	-3.08**
3	42	.10	.01 – .18	-1.58
4	67	.06	0 – .12	-3.70**
5	53	.04	-.01 – .09	-4.93**

*significantly different from chance at $p < .05$

**significantly different from chance at $p < .01$

Table S12

Stimulus Set 5. Proportion of Mock-Witness Suspect Picks, Suspect Pick 95% Confidence

Interval (CI) and Difference from Chance Estimate for each Lineup

High similarity fillers	N	Suspect bias stats		
		Suspect picks	Suspect picks 95% CI	Difference from chance
Target-present				
0	49	.04	-.02 – .10	-4.45**
1	61	.03	-.01 – .08	-5.87**
2	44	.07	-.01 – .14	-2.59**
3	46	.04	-.02 – .10	-4.10**
4	52	.12	.03 – .20	-1.16
5	50	.02	-.02 – .06	7.41**
Target-absent				
0	38	.21	.08 – .34	0.66
1	54	.13	.04 – .22	-0.81
2	63	.14	.06 – .23	-0.54
3	42	.29	.15 – .42	1.71
4	49	.47	.33 – .61	4.25**
5	51	.29	.17 – .42	2.00*

*significantly different from chance at $p < .05$

**significantly different from chance at $p < .01$

Table S13

Stimulus Set 6. Proportion of Mock-Witness Suspect Picks, Suspect Pick 95% Confidence

Interval (CI) and Difference from Chance Estimate for each Lineup

High similarity fillers	N	Suspect bias stats		
		Suspect picks	Suspect picks 95% CI	Difference from chance
Target-present				
0	39	.13	.02 – .23	-0.72
1	42	.24	.11 – .37	1.09
2	43	.21	.09 – .33	0.69
3	50	.32	.19 – .45	2.32*
4	55	.22	.11 – .33	0.93
5	44	.11	.02 – .21	-1.11
Target-absent				
0	50	.18	.07 – .29	0.25
1	46	.13	.03 – .23	-0.73
2	70	.20	.11 – .29	0.70
3	63	.33	.22 – .45	2.81**
4	46	.20	.08 – .39	0.50
5	51	.06	-.01 – .12	-3.27**

*significantly different from chance at $p < .05$

**significantly different from chance at $p < .01$

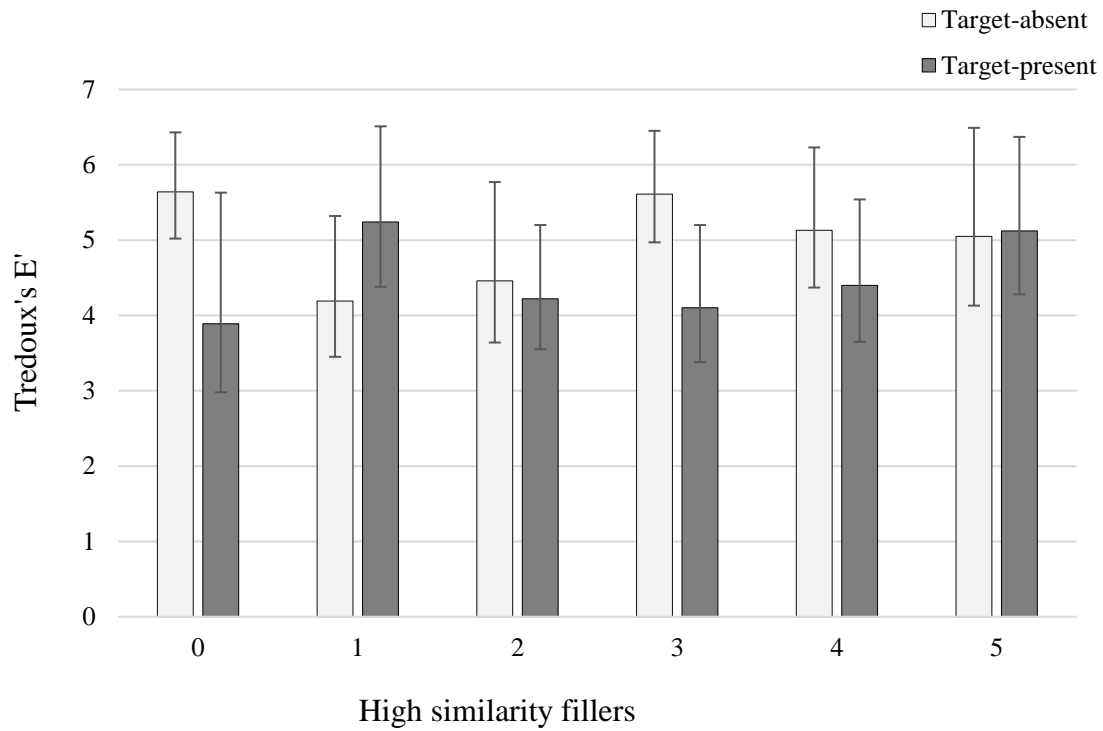


Figure S1. Stimulus set 1. Lineup functional size, with 95% CIs, for each lineup

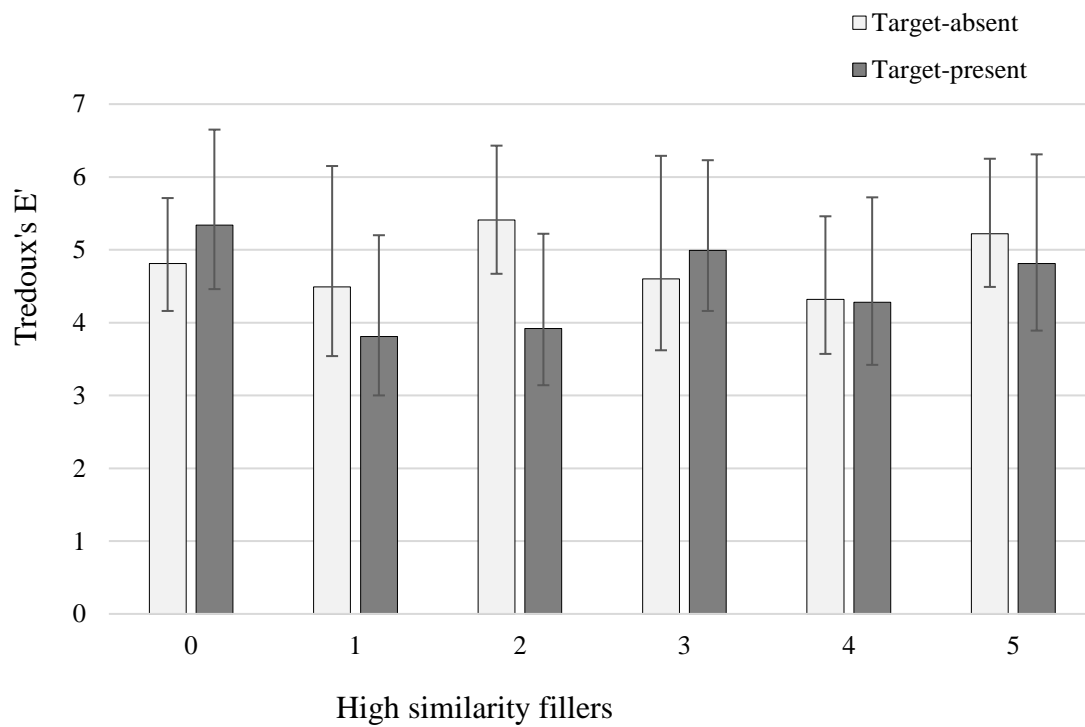


Figure S2. Stimulus set 2. Lineup functional size, with 95% CIs, for each lineup

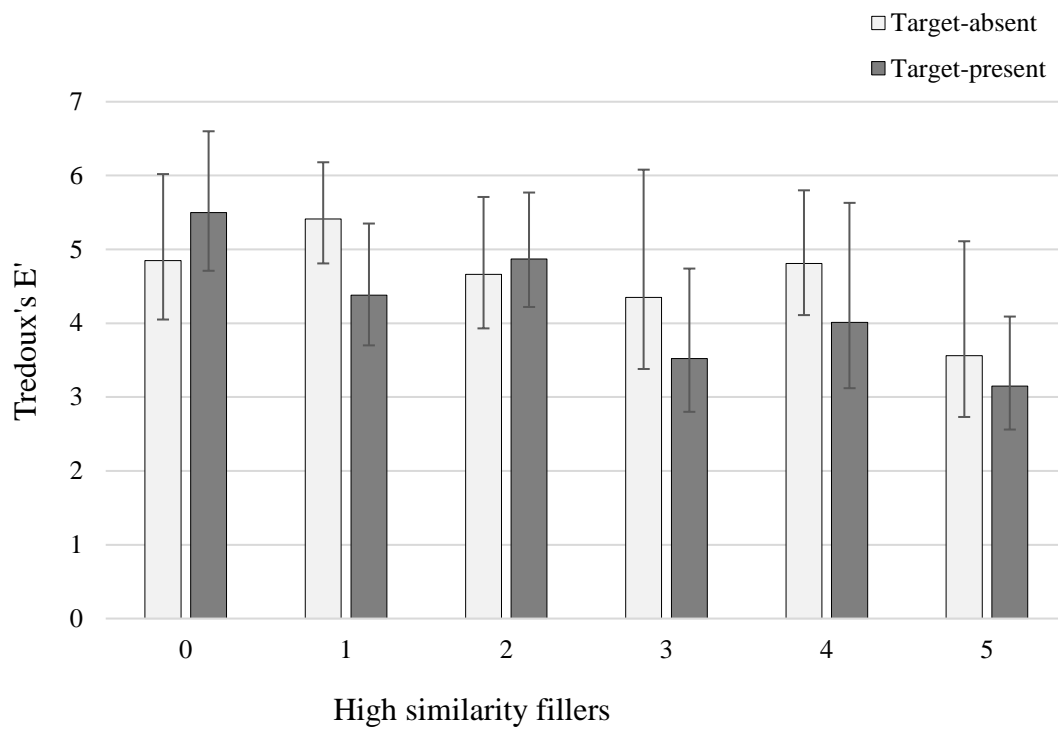


Figure S3. Stimulus set 3. Lineup functional size, with 95% CIs, for each lineup

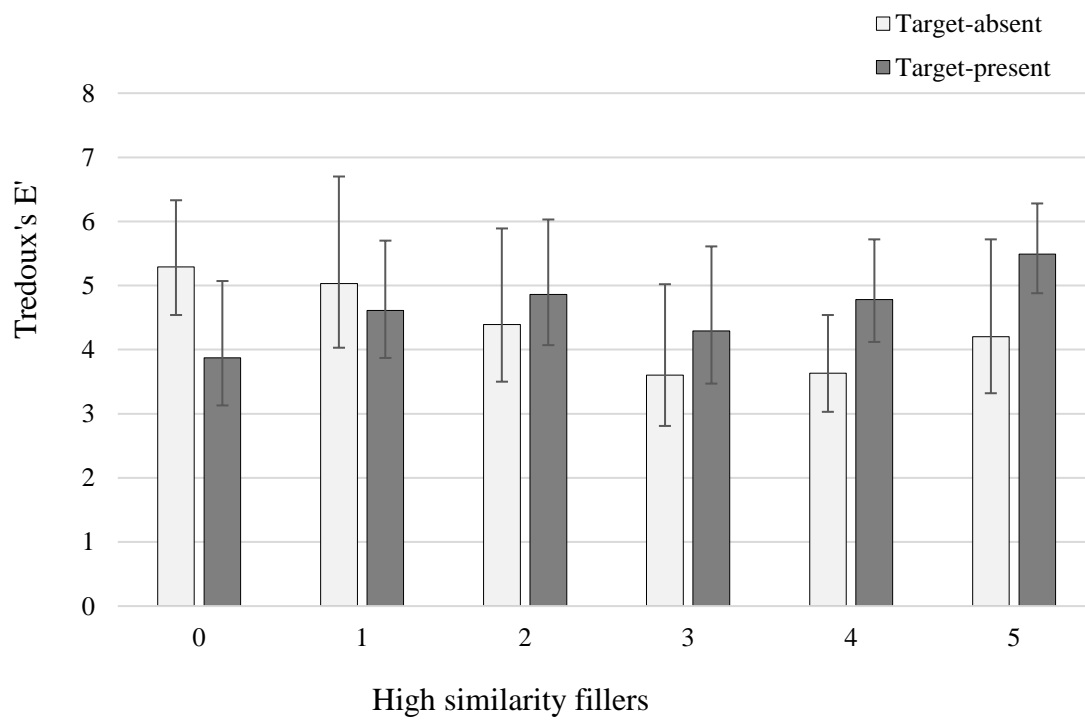


Figure S4. Stimulus set 4. Lineup functional size, with 95% CIs, for each lineup

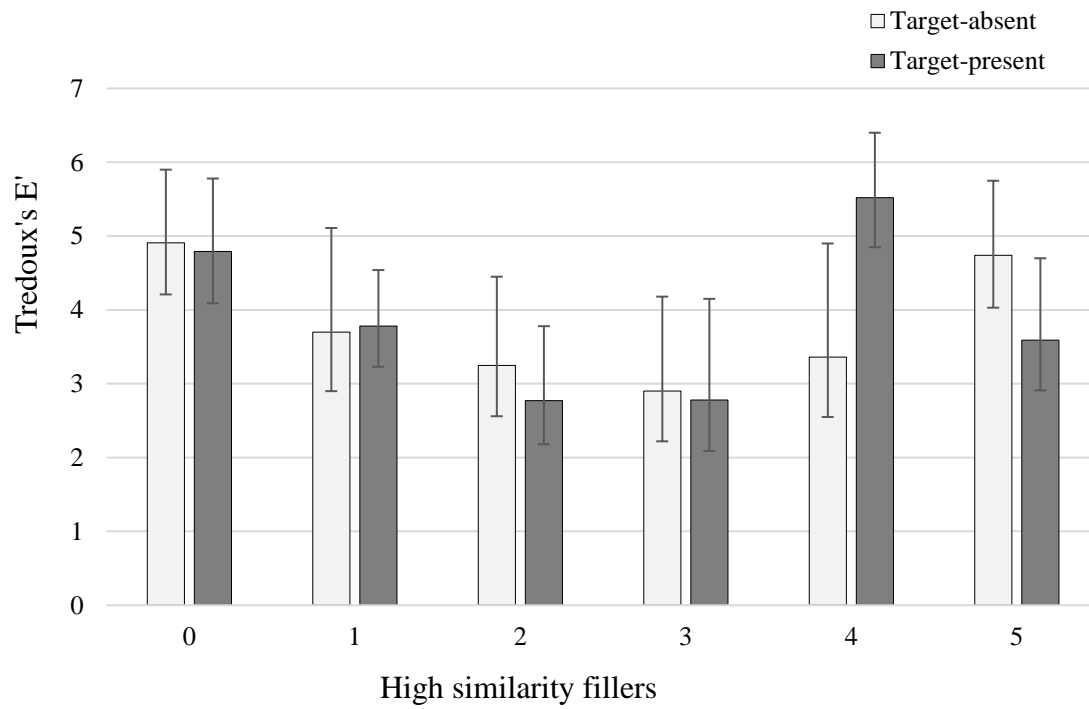


Figure S5. Stimulus set 5. Lineup functional size, with 95% CIs, for each lineup

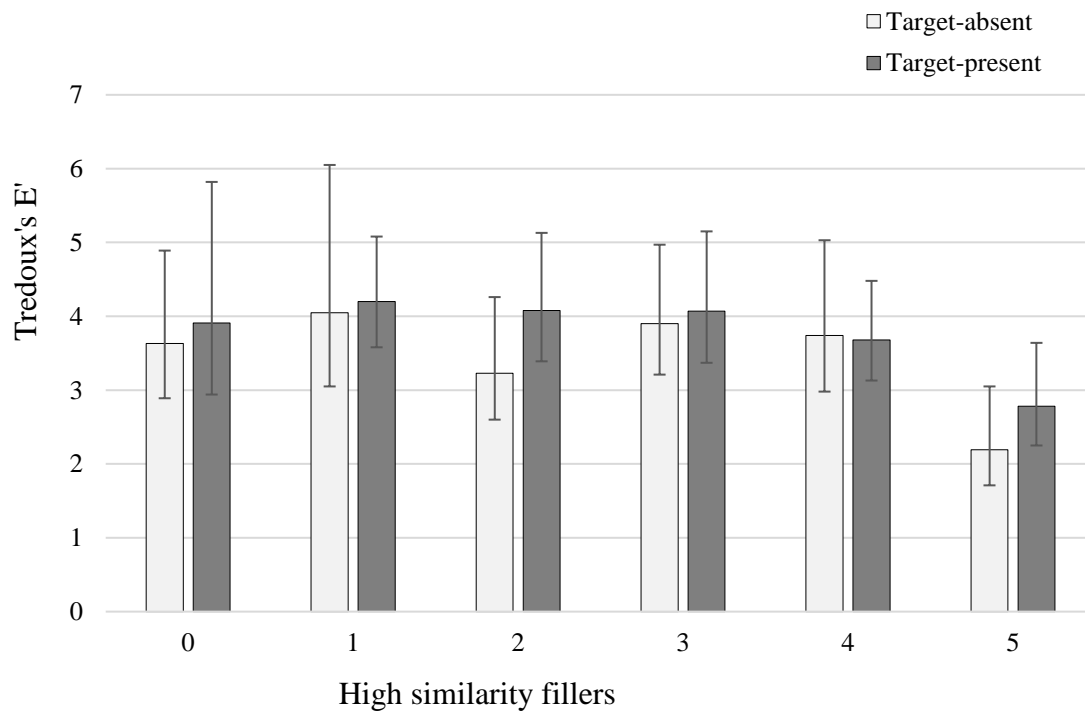


Figure S6. Stimulus set 6. Lineup functional size, with 95% CIs, for each lineup

Supplemental Material: Identification Patterns, Accuracy and Confidence

Table S14

Experiment 1 Stimulus Set 1. Identification Decision Patterns [with 95% CIs] Across Number of High Similarity Fillers and Target Presence

High similarity fillers	Target-present				Target-absent			
	<i>N</i>	Choosing	Suspect pick	Filler pick	<i>N</i>	Choosing	Suspect pick	Filler pick
0	53	.87 [.77, .95]	.81 [.70, .91]	.06 [.00, .11]	51	.55 [.40, .68]	.39 [.25, .52]	.16 [.05, .25]
1	53	.91 [.82, .97]	.89 [.79, .96]	.02 [.00, .05]	53	.66 [.52, .78]	.49 [.35, .62]	.17 [.06, .26]
2	51	.92 [.84, .99]	.86 [.76, .95]	.06 [.00, .11]	50	.62 [.48, .74]	.32 [.18, .44]	.30 [.16, .42]
3	52	.87 [.76, .95]	.83 [.71, .92]	.04 [.00, .08]	53	.60 [.46, .73]	.40 [.26, .52]	.21 [.09, .31]
4	53	.96 [.90, .100]	.91 [.82, .97]	.06 [.00, .11]	52	.58 [.43, .70]	.23 [.11, .34]	.35 [.21, .47]
5	51	100 -	.92 [.84, .99]	.08 [.00, .14]	51	.73 [.59, .84]	.27 [.14, .39]	.45 [.30, .58]

Table S15

Experiment 1 Stimulus Set 2. Identification Decision Patterns [with 95% CIs] Across Number of High Similarity Fillers and Target Presence

High similarity fillers	Target-present				Target-absent			
	<i>N</i>	Choosing	Suspect pick	Filler pick	<i>N</i>	Choosing	Suspect pick	Filler pick
0	52	.38 [.24, .51]	.33 [.19, .44]	.06 [.00, .11]	53	.28 [.15, .39]	.15 [.05, .24]	.13 [.03, .21]
1	53	.34 [.20, .46]	.17 [.06, .26]	.17 [.06, .26]	53	.21 [.09, .31]	.19 [.07, .28]	.02 [.00, .05]
2	52	.35 [.21, .47]	.33 [.19, .44]	.02 [.00, .05]	52	.25 [.12, .36]	.21 [.09, .31]	.04 [.00, .08]
3	52	.56 [.41, .68]	.27 [.14, .38]	.29 [.16, .40]	51	.27 [.14, .39]	.08 [.00, .14]	.20 [.08, .30]
4	51	.45 [.30, .58]	.22 [.09, .32]	.24 [.11, .34]	51	.49 [.34, .62]	.22 [.09, .32]	.27 [.14, .39]
5	52	.38 [.24, .51]	.27 [.14, .38]	.12 [.02, .19]	51	.33 [.19, .45]	.04 [.00, .08]	.29 [.16, .41]

Table S16

Experiment 1 Stimulus Set 3. Identification Decision Patterns [with 95% CIs] Across Number of High Similarity Fillers and Target Presence

High similarity fillers	Target-present				Target-absent			
	<i>N</i>	Choosing	Suspect pick	Filler pick	<i>N</i>	Choosing	Suspect pick	Filler pick
0	51	.63 [.49, .75]	.37 [.23, .50]	.25 [.13, .36]	53	.40 [.26, .52]	.28 [.15, .39]	.11 [.02, .19]
1	53	.70 [.57, .81]	.28 [.15, .39]	.42 [.27, .54]	52	.69 [.56, .81]	.29 [.16, .40]	.40 [.26, .53]
2	52	.71 [.58, .83]	.19 [.08, .29]	.52 [.37, .65]	51	.65 [.51, .77]	.27 [.14, .39]	.37 [.23, .50]
3	53	.70 [.57, .81]	.08 [.00, .14]	.62 [.48, .74]	53	.68 [.54, .80]	.15 [.05, .24]	.53 [.38, .65]
4	51	.71 [.57, .82]	.12 [.02, .20]	.59 [.44, .71]	53	.77 [.65, .88]	.13 [.03, .21]	.64 [.50, .76]
5	52	.79 [.67, .89]	.15 [.05, .24]	.63 [.49, .76]	49	.78 [.65, .88]	.14 [.03, .23]	.63 [.49, .76]

Table S17

Experiment 1 Stimulus Set 4. Identification Decision Patterns [with 95% CIs] Across Number of High Similarity Fillers and Target Presence

High similarity fillers	Target-present				Target-absent			
	<i>N</i>	Choosing	Suspect pick	Filler pick	<i>N</i>	Choosing	Suspect pick	Filler pick
0	50	.74 [.61, .85]	.56 [.41, .69]	.18 [.06, .28]	52	.62 [.47, .74]	.13 [.03, .22]	.48 [.34, .61]
1	52	.71 [.58, .83]	.42 [.28, .55]	.29 [.16, .40]	51	.51 [.36, .64]	.08 [.00, .14]	.43 [.29, .56]
2	51	.71 [.57, .82]	.51 [.36, .64]	.20 [.08, .30]	53	.58 [.44, .71]	.04 [.00, .08]	.55 [.40, .67]
3	53	.79 [.67, .89]	.55 [.40, .67]	.25 [.12, .35]	52	.71 [.58, .83]	.08 [.00, .14]	.63 [.49, .76]
4	52	.79 [.67, .89]	.46 [.32, .59]	.33 [.19, .44]	52	.69 [.56, .81]	.04 [.00, .08]	.65 [.51, .77]
5	52	.83 [.71, .92]	.38 [.24, .51]	.44 [.30, .57]	53	.62 [.48, .74]	.06 [.00, .11]	.57 [.42, .69]

Table S18

Experiment 1 Stimulus Set 5. Identification Decision Patterns [with 95% CIs] Across Number of High Similarity Fillers and Target Presence

High similarity fillers	Target-present				Target-absent			
	<i>N</i>	Choosing	Suspect pick	Filler pick	<i>N</i>	Choosing	Suspect pick	Filler pick
0	53	.74 [.61, .85]	.64 [.50, .76]	.09 [.01, .16]	52	.27 [.14, .38]	.13 [.03, .22]	.13 [.03, .22]
1	52	.75 [.62, .86]	.63 [.49, .76]	.12 [.02, .19]	51	.35 [.21, .47]	.16 [.05, .25]	.20 [.08, .30]
2	52	.65 [.51, .77]	.60 [.45, .72]	.06 [.00, .11]	53	.30 [.17, .42]	.06 [.00, .11]	.25 [.12, .35]
3	52	.69 [.56, .81]	.54 [.39, .66]	.15 [.05, .24]	50	.42 [.27, .55]	.16 [.05, .25]	.26 [.13, .37]
4	52	.69 [.56, .81]	.56 [.41, .68]	.13 [.03, .22]	51	.39 [.25, .52]	.06 [.00, .11]	.33 [.19, .45]
5	52	.75 [.62, .86]	.62 [.47, .74]	.13 [.03, .22]	53	.47 [.33, .60]	.08 [.00, .14]	.40 [.26, .52]

Table S19

Experiment 1 Stimulus Set 6. Identification Decision Patterns [with 95% CIs] Across Number of High Similarity Fillers and Target Presence

High similarity fillers	Target-present				Target-absent			
	<i>N</i>	Choosing	Suspect pick	Filler pick	<i>N</i>	Choosing	Suspect pick	Filler pick
0	51	.75 [.61, .85]	.73 [.59, .84]	.02 [.00, .05]	52	.19 [.08, .29]	.12 [.02, .19]	.08 [.00, .14]
1	51	.75 [.62, .85]	.75 [.62, .85]	.00 -	52	.17 [.06, .27]	.08 [.00, .14]	.10 [.01, .17]
2	53	.83 [.72, .92]	.64 [.50, .76]	.19 [.07, .28]	52	.48 [.34, .61]	.00 -	.48 [.34, .61]
3	53	.81 [.70, .91]	.60 [.46, .73]	.21 [.09, .31]	50	.62 [.48, .74]	.04 [.00, .08]	.58 [.43, .71]
4	52	.85 [.74, .93]	.52 [.37, .65]	.33 [.19, .44]	53	.51 [.37, .63]	.06 [.00, .11]	.45 [.31, .58]
5	52	.87 [.76, .95]	.63 [.49, .76]	.23 [.11, .34]	52	.52 [.37, .65]	.02 [.00, .05]	.50 [.35, .63]

Table S20

Experiment 1 Across Stimulus Sets. Accuracy Proportions [with 95% CIs] and Mean Confidence (with Standard Deviations) Across Number of High Similarity Fillers and Target Presence

High similarity fillers	Accuracy			Confidence	
	<i>Trials</i>	<i>Prop</i>	<i>[95% CI]</i>	<i>M</i>	<i>(SD)</i>
Target-absent					
0	313	.62	[.56, .67]	67.16	(24.81)
1	312	.57	[.51, .62]	65.58	(24.89)
2	311	.52	[.46, .57]	62.70	(24.42)
3	309	.45	[.39, .50]	61.55	(24.03)
4	312	.43	[.37, .48]	58.88	(24.11)
5	309	.43	[.37, .48]	59.09	(24.94)
Overall	1,866	.50	[.48, .52]	62.50	(24.70)
Target-present					
0	310	.57	[.51, .62]	68.90	(23.92)
1	314	.52	[.46, .57]	66.31	(25.08)
2	311	.52	[.46, .57]	64.63	(25.72)
3	315	.48	[.42, .53]	60.89	(25.20)
4	311	.47	[.41, .52]	63.02	(24.74)
5	311	.50	[.44, .55]	63.12	(24.16)
Overall	1,872	.51	[.49, .53]	64.47	(24.91)
All cases					
0	623	.60	[.56, .64]	68.03	(24.37)
1	626	.54	[.50, .58]	65.94	(24.97)
2	622	.52	[.48, .56]	63.67	(25.08)

3	624	.46	[.42, .50]	61.22	(24.61)
4	623	.45	[.41, .49]	60.95	(24.49)
5	620	.46	[.42, .50]	61.11	(24.61)
Overall	3,738	.51	[.49, .53]	63.49	(24.82)

Table S21

Experiment 1 Stimulus Set 1. Accuracy Proportions [with 95% CIs] and Mean Confidence (with Standard Deviations) Across Number of High Similarity Fillers and Target Presence

High similarity fillers	Target-present			Target-absent		
	<i>N</i>	Accuracy	Confidence	<i>N</i>	Accuracy	Confidence
0	53	.81 [.70, .91]	78.30 (23.10)	51	.45 [.30, .58]	64.51 (21.10)
1	53	.89 [.79, .96]	80.00 (20.38)	53	.34 [.20, .46]	68.11 (22.37)
2	51	.86 [.76, .95]	79.61 (24.08)	50	.38 [.24, .50]	64.20 (23.22)
3	52	.83 [.71, .92]	75.19 (25.40)	53	.40 [.26, .52]	67.55 (21.39)
4	53	.91 [.82, .97]	79.06 (18.94)	52	.42 [.28, .55]	62.88 (19.44)
5	51	.92 [.84, .99]	73.33 (21.88)	51	.27 [.14, .39]	65.49 (24.03)

Table S22

*Experiment 1 Stimulus Set 2. Accuracy Proportions [with 95% CIs] and Mean Confidence
(with Standard Deviations) Across Number of High Similarity Fillers and Target Presence*

High similarity fillers	Target-present			Target-absent		
	<i>N</i>	Accuracy	Confidence	<i>N</i>	Accuracy	Confidence
0	52	.33 [.19, .44]	73.85 (20.88)	53	.72 [.59, .83]	69.81 (27.35)
1	53	.17 [.06, .26]	69.62 (24.02)	53	.79 [.67, .89]	77.74 (21.18)
2	52	.33 [.19, .44]	71.92 (21.70)	52	.75 [.62, .86]	68.65 (24.58)
3	52	.27 [.14, .38]	59.81 (23.47)	51	.73 [.60, .84]	68.43 (24.03)
4	51	.22 [.10, .32]	61.96 (19.90)	51	.51 [.37, .63]	63.92 (25.38)
5	52	.27 [.14, .38]	67.50 (24.08)	51	.67 [.53, .78]	69.02 (20.81)

Table S23

Experiment 1 Stimulus Set 3. Accuracy Proportions [with 95% CIs] and Mean Confidence (with Standard Deviations) Across Number of High Similarity Fillers and Target Presence

High similarity fillers	Target-present			Target-absent		
	<i>N</i>	Accuracy	Confidence	<i>N</i>	Accuracy	Confidence
0	51	.37 [.23, .50]	61.76 (25.04)	53	.60 [.46, .73]	65.09 (23.83)
1	53	.28 [.15, .39]	60.38 (23.94)	52	.31 [.17, .42]	56.73 (22.03)
2	52	.19 [.08, .29]	55.58 (23.80)	51	.35 [.21, .47]	57.84 (24.03)
3	53	.08 [.00, .14]	56.60 (23.77)	53	.32 [.19, .44]	53.58 (22.11)
4	51	.12 [.02, .20]	55.69 (25.87)	53	.23 [.10, .33]	53.21 (22.43)
5	52	.15 [.05, .24]	58.27 (25.95)	49	.22 [.10, .33]	50.61 (24.70)

Table S24

Experiment 1 Stimulus Set 4. Accuracy Proportions [with 95% CIs] and Mean Confidence (with Standard Deviations) Across Number of High Similarity Fillers and Target Presence

High similarity fillers	Target-present			Target-absent		
	<i>N</i>	Accuracy	Confidence	<i>N</i>	Accuracy	Confidence
0	50	56.00 [.41, .69]	54.60 (23.49)	52	38.46 [.24, .51]	53.65 (27.44)
1	52	42.31 [.28, .55]	48.65 (24.74)	51	49.02 [.34, .62]	46.67 (22.69)
2	51	50.98 [.36, .64]	44.51 (25.79)	53	41.51 [.27, .54]	50.19 (24.38)
3	53	54.72 [.40, .67]	48.87 (23.34)	52	28.5 [.15, .40]	48.27 (22.81)
4	52	46.15 [.32, .59]	49.23 (25.12)	52	30.77 [.17, .42]	43.65 (21.51)
5	52	38.46 [.24, .51]	49.04 (22.07)	53	37.74 [.24, .50]	45.47 (25.69)

Table S25

*Experiment 1 Stimulus Set 5. Accuracy Proportions [with 95% CIs] and Mean Confidence
(with Standard Deviations) Across Number of High Similarity Fillers and Target Presence*

High similarity fillers	Target-present			Target-absent		
	<i>N</i>	Accuracy	Confidence	<i>N</i>	Accuracy	Confidence
0	53	.64 [.50, .76]	74.72 (19.77)	52	.73 [.60, .84]	74.23 (19.84)
1	52	.63 [.50, .75]	67.31 (24.90)	51	.65 [.51, .77]	70.78 (22.26)
2	52	.60 [.45, .72]	65.96 (20.41)	53	.70 [.56, .81]	67.17 (24.05)
3	52	.54 [.39, .66]	63.08 (25.63)	50	.58 [.43, .71]	63.40 (24.63)
4	52	.56 [.41, .68]	62.69 (22.68)	51	.61 [.46, .73]	66.47 (23.05)
5	52	.62 [.47, .74]	63.08 (22.93)	53	.53 [.34, .65]	60.94 (23.31)

Table S26

*Experiment 1 Stimulus Set 6. Accuracy Proportions [with 95% CIs] and Mean Confidence
(with Standard Deviations) Across Number of High Similarity Fillers and Target Presence*

High similarity fillers	Target-present			Target-absent		
	<i>N</i>	Accuracy	Confidence	<i>N</i>	Accuracy	Confidence
0	51	.73 [.59, .84]	69.22 (23.48)	52	.81 [.69, .91]	75.58 (22.61)
1	51	.75 [.62, .85]	71.76 (21.23)	52	.83 [.71, .92]	72.88 (25.62)
2	53	.64 [.50, .76]	70.00 (23.20)	52	.52 [.37, .65]	68.27 (21.58)
3	53	.60 [.46, .73]	62.08 (22.98)	50	.38 [.24, .50]	68.60 (22.13)
4	52	.52 [.50, .76]	69.04 (24.75)	53	.49 [.34, .62]	63.40 (25.19)
5	52	.63 [.49, .76]	67.69 (21.57)	52	.48 [.34, .61]	63.08 (23.31)

Experiment 2 Laptop Theft Stimulus Set¹

Identification decision patterns. The 4-way association between the number of high similarity fillers, retention interval, target presence and identification decisions was significant, $k = 4$, LR $\chi^2(4) = 10.43$, $p = .034$. Separate examinations of each decision type revealed a 4-way association between the predictors and choosing, $k = 4$, LR $\chi^2(2) = 7.21$, $p = .027$ (note, however, that this association is not significant with an adjusted alpha of $.05/3 = .017$). Descriptive statistics, shown in Table S25, suggest that the trend for choosing to increase in lineups containing larger numbers of high similarity fillers was only evident for target-present cases in the short retention interval condition and target-absent cases in the long retention interval condition. Further, note that suspect and filler identifications appeared to be similarly affected across conditions (i.e., in line with the effects observed when collapsing across stimuli, increasing the number of high similarity fillers decreased suspect identifications and increased filler identifications).

Accuracy. The association between the number of high similarity fillers, retention interval, target presence and accuracy was non-significant, $k = 4$, LR $\chi^2(2) = 1.17$, $p = .572$. The test for 3-way associations was significant, $k = 3$, LR $\chi^2(7) = 42.92$, $p < .001$, with partial associations showing a significant interaction between retention interval and target presence on accuracy, $\chi^2(1) = 38.04$, $p < .001$. In target-present cases, accuracy decreased from .81, 95% CI [.75, .86], in the short retention interval condition to .39, 95% CI [.33, .44], in the

¹ Experiment 1, Stimulus Set 1.

long retention interval condition, $OR = 6.65$. In target-absent cases, accuracy only decreased from .49, 95% CI [.42, .56] in the short retention interval condition to, to .44, 95% CI [.38, .49], in the long retention interval condition, $OR = 1.22$. The test of 2-way associations was also significant, $k = 2$, $LR \chi^2(9) = 65.67, p < .001$; however, partial associations indicated that the significant effects were for variables already associated with a higher order effect (and, therefore, considered confounded with the higher order effect and ignored). Thus, there was no evidence of the number of high similarity fillers affecting accuracy for this stimulus set.

Confidence. A 3 (number of high similarity fillers: zero, two, five) $\times 2$ (retention interval: short, long) $\times 2$ (target presence: present, absent) ANOVA showed a main effect of the number of high similarity fillers on confidence, $F(2, 994) = 7.69, p < .001$, with descriptive statistics indicating a marginal decrease in confidence from the lineup including no high similarity fillers ($M = 61.07, SD = 25.86$) to those including 2 ($M = 59.00, SD = 26.78$) or 5 ($M = 59.00, SD = 26.78$) (d 0 vs. 2 & 5 = 0.08; d 2 vs. 5 = 0). A main effect of target presence, $F(1, 994) = 25.53, p < .001$, reflected lower confidence in target-absent cases ($M = 54.90, SD = 25.50$) compared with -present cases ($M = 61.43, SD = 27.48$), $d = 0.25$. There was also a main effect of retention interval, $F(1, 994) = 286.12, p < .001$, with confidence notably higher for identifications made after the short ($M = 73.31, SD = 21.31$) than the long ($M = 48.08, SD = 25.12$) retention interval, $d = 1.07$. There was an interaction between retention interval and target presence on confidence, $F(1, 994) = 9.89, p = .002, f = 0.10$. Descriptive statistics suggest the decrease in mean confidence over time was greater in target-

present (29.95, $d = 1.29$) than -absent (20.57, $d = 0.88$) cases. None of the other interactions were significant ($F_s < 0.91$, $p_s > .407$, $f_s < 0.05$).

Table S27

Experiment 2 Laptop Theft Stimulus Set. Identification Decision Patterns [with 95% CIs] Across all Experimental Conditions

	Short retention interval				Long retention interval			
	<i>N</i>	Choosing	Suspect pick	Filler pick	<i>N</i>	Choosing	Suspect pick	Filler pick
Target-absent								
0	67	.49 [.37, .60]	.45 [.32, .56]	.04 [.00, .09]	102	.58 [.48, .67]	.35 [.26, .44]	.23 [.14, .30]
2	69	.52 [.40, .63]	.36 [.24, .47]	.16 [.07, .24]	100	.47 [.37, .56]	.28 [.19, .36]	.19 [.11, .26]
5	66	.51 [.39, .63]	.20 [.09, .29]	.32 [.20, .42]	100	.63 [.53, .72]	.18 [.10, .25]	.45 [.35, .54]
Target-present								
0	66	.82 [.72, .90]	.80 [.70, .89]	.02 [.00, .04]	102	.59 [.49, .68]	.44 [.34, .53]	.15 [.07, .21]
2	67	.81 [.70, .89]	.79 [.69, .88]	.01 [.00, .04]	103	.59 [.49, .68]	.39 [.29, .48]	.20 [.12, .28]
5	67	.96 [.90, 1.00]	.84 [.74, .92]	.12 [.03, .19]	97	.63 [.53, .72]	.34 [.24, .43]	.29 [.19, .37]

Experiment 2 Car Theft Stimulus Set²

Identification decision patterns. The 4-way association between the number of high similarity fillers, retention interval, target presence and identification decisions, was non-significant, $k = 4$, LR $\chi^2(4) = 3.95$, $p = .412$, as was the test for any 3-way associations, $k = 3$, LR $\chi^2(12) = 10.33$, $p = .587$. The test for 2-way associations was significant, $k = 2$, LR $\chi^2(13) = 134.96$, $p < .001$, with partial associations indicating significant effects of the number of high similarity fillers $\chi^2(4) = 77.83$, $p < .001$, target presence, $\chi^2(2) = 6.95$, $p = .031$, and retention interval, $\chi^2(2) = 52.56$, $p < .001$, on identification decisions. Separate examinations of each identification decision type indicated significant effects of the number of high similarity fillers on, choosing, $\chi^2(2) = 11.35$, $p = .003$, suspect identifications, $\chi^2(2) = 39.71$, $p < .001$, and filler identifications $\chi^2(2) = 60.93$, $p < .001$. As the number of high similarity fillers increased from zero to two to five, choosing increased from .55, 95% CI [.49, .60], to .60, 95% CI [.54, .65], to .67, 95% CI [.62, .72], (OR 0 vs. 2 = 1.23; OR 0 vs. 5 = 1.62; OR 2 vs. 5 = 1.31). Suspect identifications decreased from .25, 95% CI [.20, .29], to .14, 95% CI [.10, .17], to .07, 95% [.04, .10] (OR 0 vs. 2 = 2.05; OR 0 vs. 5 = 4.10; OR 2 vs. 5 = 1.99), while filler identifications increased from .30, 95% CI [.25, .35], to .46, 95% CI [.40, .51], to .60, 95% CI [.54, .65] (OR 0 vs. 2 = 2.00; OR 0 vs. 5 = 3.50; OR 2 vs. 5 = 1.75). There was a significant effect of target presence on suspect identifications, $\chi^2(1) = 6.15$, $p = .013$, with more suspects being identified in target-absent .18, 95% CI [.15, .21] than -present .13, 95% CI [.10, .17], cases, OR = 1.55. Retention interval affected both choosing and filler identification rates, $\chi^2(1) = 36.64$, $p < .001$ and $\chi^2(1) = 47.41$, $p < .001$, respectively. Choosing was higher when identifications were made after the long retention interval, .68, 95% CI [.64, .72], than the short retention interval, .49, 95% CI [.44, .54], retention interval, OR = 362.

² Experiment 1, Stimulus Set 3.

Filler identifications occurred at a rate of .54, 95% CI [.50, .58] at the long retention interval and .32, 95% CI [.27, .36], at the short retention interval, OR = 2.51.

Accuracy. The 4-way association between the predictors and accuracy was non-significant $k = 4$, LR $\chi^2(2) = 1.43$, $p = .490$, as was the test for 3-way associations, $k = 3$, LR $\chi^2(7) = 8.02$, $p = .331$. The test for 2-way associations was significant, $k = 2$, LR $\chi^2(9) = 137.87$, $p < .001$, with significant effects on accuracy of the number of high similarity fillers, $\chi^2(2) = 18.64$, $p < .001$, retention interval, $\chi^2(1) = 33.85$, $p < .001$, and target presence, $\chi^2(1) = 90.71$, $p < .001$. Accuracy decreased as the number of high similarity fillers increased from zero, .32, 95% CI [.27, .37], to two, .24, 95% CI [.19, .26], to five, .19, 95% CI [.14, .23] (OR 0 vs. 2 = 1.49; OR 0 vs. 5 = 2.05; OR 2 vs. 5 = 1.38). Accuracy was higher for decisions made after the short, .34, 95% CI [.30, .39], compared with the long, .19, 95% CI [.16, .22], retention interval, OR = 2.27 and higher in target-absent, .37, 95% CI [.33, .41] than -present, .13, 95% CI [.10, .15], cases, OR = 4.14.

Confidence. A 3 (number of high similarity fillers: zero, two, five) \times 2 (retention interval: short, long) \times 2 (target presence: present, absent) ANOVA showed a main effect of the number of high similarity fillers on confidence, $F(2, 992) = 5.95$, $p = .003$, with descriptive statistics indicating a marginal decrease in confidence as the number of high similarity fillers increased from 0 ($M = 55.31$, $SD = 25.44$), to 2 ($M = 52.49$, $SD = 25.02$) to 5 ($M = 49.32$, $SD = 24.05$) (d 0 vs. 2 = 0.11; d 0 vs. 5 = 0.24; d 2 vs. 5 = 0.13). The main effect of target presence was non-significant, $F(1, 992) = 0.91$, $p = .340$, $d = 0.06$. There was a main effect of retention interval, $F(1, 992) = 187.04$, $p < .001$, with higher confidence for identifications made after the short ($M = 64.62$, $SD = 21.78$) than the long ($M = 44.39$, $SD = 23.60$) retention interval, $d = 0.88$. None of the interactions were significant (F s < 1.89 , p s $> .152$, f s < 0.07).

Table S28

Experiment 2 Car Theft Stimulus Set. Identification Decision Patterns [with 95% CIs] Across all Experimental Conditions

High similarity fillers	Short retention interval				Long retention interval			
	<i>N</i>	Choosing	Suspect pick	Filler pick	<i>N</i>	Choosing	Suspect pick	Filler pick
Target-absent								
0	67	.40 [.28, .51]	.24 [.13, .33]	.16 [.07, .25]	99	.68 [.58, .76]	.32 [.23, .41]	.35 [.25, .44]
2	64	.45 [.32, .57]	.14 [.05, .22]	.31 [.19, .42]	99	.74 [.65, .82]	.18 [.10, .25]	.56 [.45, .65]
5	67	.61 [.49, .72]	.12 [.03, .19]	.49 [.37, .60]	106	.74 [.65, .82]	.08 [.02, .12]	.66 [.57, .75]
Target-present								
0	66	.52 [.39, .63]	.32 [.20, .42]	.20 [.09, .29]	102	.54 [.44, .63]	.14 [.07, .20]	.40 [.30, .49]
2	67	.43 [.31, .54]	.13 [.05, .21]	.30 [.18, .40]	103	.66 [.56, .75]	.10 [.04, .15]	.56 [.46, .65]
5	65	.52 [.39, .64]	.08 [.00, .13]	.45 [.32, .56]	99	.75 [.66, .83]	.04 [.00, .07]	.71 [.61, .79]

Experiment 2 Mobile Phone Theft Stimulus Set³

Identification decision patterns. The 4-way association between the number of high similarity fillers, retention interval, target presence and identification decisions was significant, $k = 4$, $LR \chi^2(4) = 13.83$, $p = .008$. Separate examinations of each identification decisions showed 3-way interactions between the predictors on both suspect, $k = 4$, $LR \chi^2(2) = 7.94$, $p = .019$ and filler identifications, $k = 4$, $LR \chi^2(2) = 7.48$, $p = .024$ (note, however, that none of the associations were significant with an adjusted alpha of $.05/3 = .017$). Descriptive statistics are shown in Table S26.

Accuracy. The 4-way association between the predictors and accuracy was non-significant $k = 4$, $LR \chi^2(2) = 3.62$, $p = .164$, as was the test for 3-way associations, $k = 3$, $LR \chi^2(7) = 11.40$, $p = .122$. The test for 2-way associations was significant, $k = 2$, $LR \chi^2(9) = 131.59$, $p < .001$, with significant effects of the number of high similarity fillers, $\chi^2(2) = 20.40$, $p < .001$, retention interval, $\chi^2(1) = 70.51$, $p < .001$, and target presence, $\chi^2(1) = 44.67$, $p < .001$, on accuracy. As the number of high similarity fillers increased from zero, to two, to five, accuracy decreased from .54, 95% CI [.49, .60], to .40, 95% CI [.35, .45], to .39, 95% CI [.34, .45] (OR 0 vs 2 = 1.77; OR 0 vs 5 = 1.83; OR 2 vs. 5 = 1.03). Accuracy was higher in the short retention interval .60, 95% CI [.56, .65], than the long retention interval, .34, 95% CI [.30, .38], condition, OR = 2.93. Also, accuracy was higher for target-absent .55, 95% CI [.50, .59], than target-present .35, 95% CI [.30, .39] cases, OR = 2.28.

Confidence. A 3 (number of high similarity fillers: zero, two, five) \times 2 (retention interval: short, long) \times 2 (target presence: present, absent) ANOVA showed a main effect of the number of high similarity fillers on confidence, $F(2, 989) = 3.77$, $p = .023$, with descriptive statistics indicating a marginal decrease in confidence as the number of high similarity fillers increased from 0 ($M = 54.81$, $SD = 26.94$), to 2 ($M = 50.60$, $SD = 26.10$) to 5

³ Experiment 1, Stimulus Set 5.

($M = 49.91$, $SD = 27.10$) (d 0 vs. 2 = 0.16; d 0 vs. 5 = 0.18; d 2 vs. 5 = 0.03). The main effect of target presence was non-significant, $F(1, 989) = 1.85$, $p = .174$, $d = 0.05$. There was a main effect of retention interval, $F(1, 989) = 317.78$, $p < .001$, with higher confidence for identifications made after the short ($M = 67.98$, $SD = 21.66$) than the long ($M = 41.13$, $SD = 24.38$) retention interval, $d = 1.15$. There was an interaction between retention interval and target presence on confidence, $F(1, 989) = 4.97$, $p = .026$, $f = 0.07$. Descriptive statistics suggest the decrease in mean confidence over time was greater in target-present (30.19, $d = 1.28$) than -absent (23.15, $d = 1.02$) cases. None of the other interactions were significant (F s < 1.13 , $ps > .326$, $fs < 0.06$).

Table S29

Experiment 2 Mobile Phone Theft Stimulus Set. Identification Decision Patterns [with 95% CIs] Across all Experimental Conditions

High similarity fillers	Short retention interval				Long retention interval			
	<i>N</i>	Choosing	Suspect pick	Filler pick	<i>N</i>	Choosing	Suspect pick	Filler pick
Target-absent								
0	68	.28 [.17, .38]	.19 [.09, .28]	.09 [.01, .15]	100	.41 [.31, .50]	.12 [.05, .18]	.29 [.20, .37]
2	66	.33 [.21, .44]	.06 [.00, .11]	.27 [.16, .37]	98	.62 [.52, .71]	.17 [.06, .24]	.45 [.35, .54]
5	64	.42 [.29, .54]	.06 [.00, .11]	.36 [.23, .47]	103	.54 [.44, .64]	.10 [.04, .15]	.45 [.35, .54]
Target-present								
0	67	.76 [.65, .86]	.61 [.49, .72]	.15 [.06, .23]	100	.53 [.43, .62]	.33 [.23, .42]	.25 [.16, .32]
2	68	.68 [.56, .78]	.51 [.39, .63]	.16 [.07, .24]	104	.65 [.56, .74]	.18 [.10, .25]	.46 [.36, .55]
5	64	.72 [.60, .82]	.53 [.40, .65]	.19 [.08, .28]	99	.60 [.49, .69]	.12 [.05, .18]	.46 [.36, .55]

Supplemental Material: SDT-CD Analyses

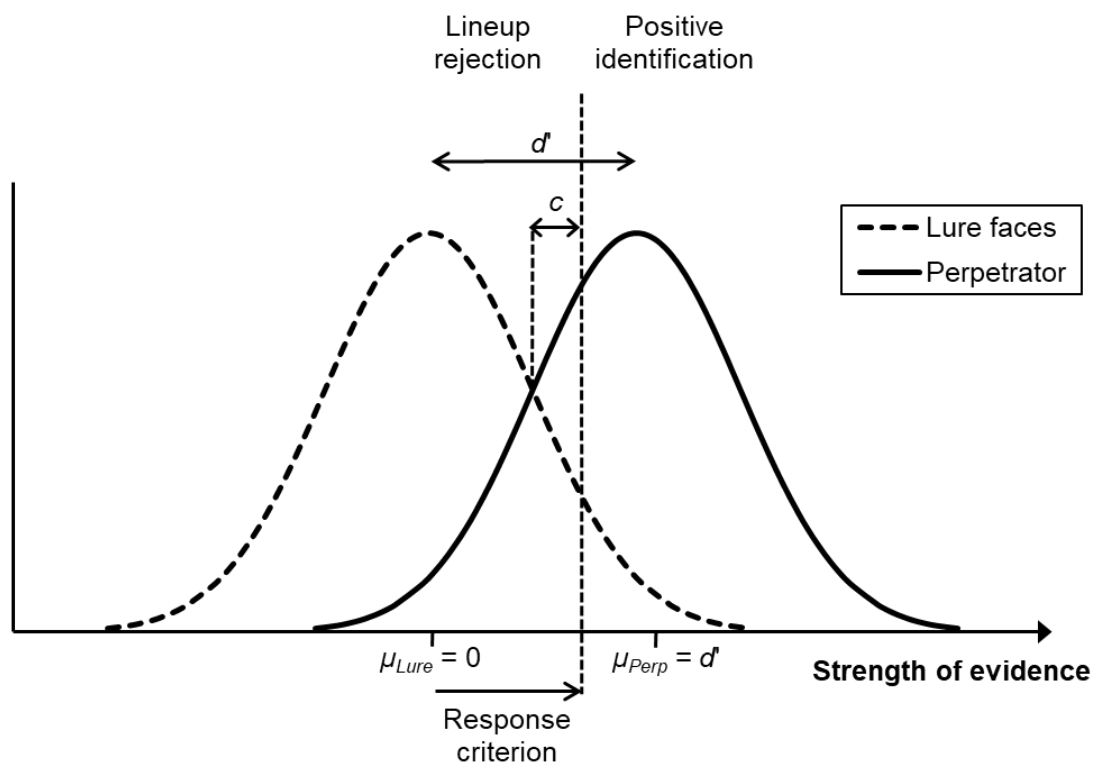


Figure S7. Conceptual representation of the *detection* component of an eyewitness identification decision. Discriminability (d') indexes the separation between strength of evidence distributions for perpetrator and lure faces. The response criterion indexes the amount of evidence required for a positive identification, relative to the mean of the lure distribution (which has a mean of zero and standard deviation of one, as is common in signal detection analysis). Response bias (c) indexes the placement of the response criterion relative to the point at which the perpetrator and lure distributions overlap.

Model Fitting

For each condition, the best-fitting SDT-CD independent was identified by using likelihood ratio tests to find the combination of d' and response criterion values that produced model-generated expected response frequencies that most closely matched observed response frequencies. Tables S28 (Experiment 1) and S29 (Experiment 2) show model fit statistics and observed and model-generated response frequencies for correct identifications and filler identifications from target-present lineups, and incorrect identifications from target-absent lineups.

Table S30

Experiment 1. Observed and Model-generated Response Frequencies, and Model Fit

Statistics for each Condition

Number of high similarity fillers	Observed			Model			G_{total}	p
	FA	CID	FID	FA	CID	FID		
0	0.38	0.57	0.11	0.32	0.51	0.18	23.82	< .001
1	0.43	0.52	0.17	0.38	0.47	0.22	13.42	.004
2	0.48	0.52	0.17	0.42	0.46	0.23	18.44	< .001
3	0.55	0.48	0.26	0.51	0.44	0.30	7.51	.057
4	0.57	0.47	0.28	0.53	0.43	0.31	6.98	.073
5	0.57	0.50	0.27	0.54	0.47	0.30	5.05	.168

Table S31

*Experiment 2. Observed and Model-generated Response Frequencies, and Model Fit**Statistics for each Condition*

Number of high similarity fillers	Observed			Model			G_{total}	p	
	FA	CID	FID	FA	CID	FID			
Short retention interval									
0	0.39	0.58	0.12	0.33	0.52	0.18	12.52	.006	
2	0.44	0.48	0.16	0.37	0.42	0.22	14.34	.002	
5	0.52	0.48	0.25	0.48	0.45	0.28	3.43	.330	
Long retention interval									
0	0.55	0.30	0.25	0.47	0.24	0.31	26.26	< .001	
2	0.61	0.22	0.41	0.58	0.21	0.43	3.17	.367	
5	0.64	0.17	0.49	0.63	0.16	0.50	0.38	.945	

Calculation of 95% Inferential Confidence Intervals (ICIs)

To facilitate inferential comparisons, we calculated 95% inferential confidence intervals (ICIs; Tryon, 2001) around d' , response criterion, and c values. We first used a modified jackknife procedure to generate estimates of standard error for each parameter value (Mosteller & Tukey, 1968; Koriat, Lichtenstein, & Fischhoff, 1980; Weber & Brewer, 2006). For each condition, d' , response criterion, and c were calculated as many times as there were participants, with one participant omitted each time. For example, in a condition with 300 participants, each parameter would be calculated 300 times, with each calculation based on data from 299 participants. This produced distributions of d' , response criterion, and c values that were used to calculate a jackknife estimate of the standard error and standard deviation for d and c in each condition. Jackknife standard errors were then used to calculate inferential confidence intervals (Tryon, 2001).

In calculating ICIs, we used an average E value (which indexes the equivalence of standard errors across conditions being compared). ICIs typically enable inferential comparisons for only a specific pair of conditions; however, the use of an average E value enables comparisons between multiple pairs of conditions (Tryon, 2001). Thus, for a given parameter (d' ; response criterion; c) any pair of conditions can be compared within Experiment 1 or Experiment 2.

Supplemental Material: Full ROC Curves

Table S32

Experiment 1. Hit Rate (HR), False Alarm Rate (FAR) and Diagnosticity Ratio (DR) for Each Confidence Category in the Small, Medium and Large Number of High Similarity Filler Conditions

Confidence categories	N	Hits	False alarms	HR	FAR	DR
Small number of high similarity fillers						
Suspect pick 80-100	193	162	31	0.26	0.05	5.23
Suspect pick 50-70	196	130	66	0.21	0.11	1.97
Suspect pick 0-40	83	50	33	0.08	0.05	1.52
Filler pick 0-40	79	34	45	0.05	0.07	0.76
Rejection 0-40	82	36	46	0.07	0.07	1.00
Filler pick 50-70	107	40	67	0.06	0.11	0.60
Rejection 50-70	194	73	121	0.12	0.19	0.60
Filler pick 80-100	26	13	13	0.02	0.02	1.00
Rejection 80-100	289	86	203	0.14	0.32	0.42
Medium number of high similarity fillers						
Suspect pick 80-100	150	128	22	0.20	0.04	5.76
Suspect pick 50-70	172	130	42	0.21	0.07	3.07
Suspect pick 0-40	83	54	29	0.09	0.05	1.84
Filler pick 0-40	131	63	68	0.10	0.11	0.92
Rejection 0-40	89	36	53	0.06	0.09	0.67
Filler pick 50-70	170	55	115	0.09	0.19	0.47
Rejection 50-70	185	72	113	0.12	0.18	0.63
Filler pick 80-100	62	18	44	0.03	0.07	0.41
Rejection 80-100	204	70	134	0.11	0.22	0.52
Large number of high similarity fillers						
Suspect pick 80-100	142	130	12	0.21	0.02	10.82
Suspect pick 50-70	149	120	29	0.19	0.05	4.13
Suspect pick 0-40	77	49	28	0.08	0.05	1.75
Filler pick 0-40	160	60	100	0.10	0.16	0.60
Rejection 0-40	77	35	42	0.06	0.07	0.83
Filler pick 50-70	205	76	129	0.12	0.21	0.59
Rejection 50-70	167	61	106	0.10	0.17	0.57
Filler pick 80-100	93	35	58	0.06	0.09	0.60
Rejection 80-100	173	56	117	0.09	0.19	0.48

Note: Alternative ROC curve operating point order (i.e., when not ordering by diagnosticity ratio) as shown above

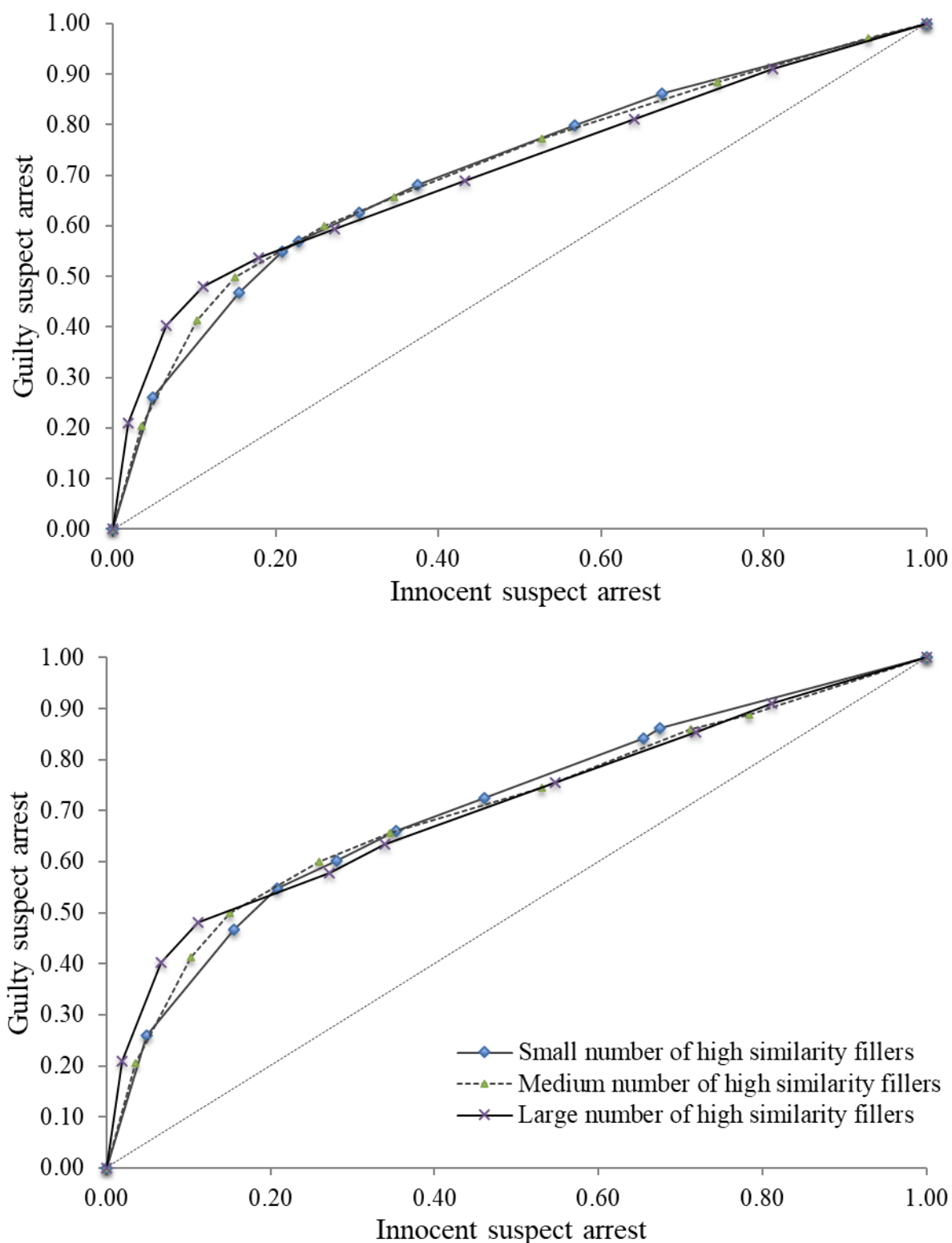


Figure S8. Experiment 1. Full ROCs with operating points ordered by diagnosticity ratio (top panel) and a priori order (bottom panel). Area under the curve values when operating points ordered by diagnosticity ratio = .709, 95% CI [.681, .737], .711, 95% CI [.681, .739] and .710, 95% CI [.680, .738] in the small, medium and large number of high similarity filler

conditions, respectively ($ps > .934$, $Ds < 0.09$). Area under the curve values when operating points ordered by a priori order = .706, 95% CI [.677, .734], .702, 95% CI [.672, .730] and .706, 95% CI [.676, .734] in the small medium and large number of high similarity filler conditions, respectively ($ps > .837$, $Ds < 0.19$).

Table S33

Experiment 2. Hit Rate (HR), False Alarm Rate (FAR) and Diagnosticity Ratio (DR) for Each Confidence Category in the 0, 2 and 5 High Similarity Filler Conditions

Confidence categories	N	Hits	False alarms	HR	FAR	DR
0 high similarity fillers						
Suspect pick 80-100	118	86	32	0.17	0.06	2.68
Suspect pick 50-70	145	74	71	0.15	0.14	1.04
Suspect pick 0-40	83	47	36	0.09	0.07	1.30
Filler pick 0-40	103	50	53	0.10	0.11	0.94
Rejection 0-40	128	69	59	0.14	0.12	1.17
Filler pick 50-70	78	36	42	0.07	0.08	0.86
Rejection 50-70	201	85	116	0.17	0.23	0.73
Filler pick 80-100	26	14	12	0.03	0.02	1.16
Rejection 80-100	125	43	82	0.09	0.16	0.52
2 high similarity fillers						
Suspect pick 80-100	81	68	13	0.13	0.03	5.07
Suspect pick 50-70	121	66	55	0.13	0.11	1.16
Suspect pick 0-40	65	32	33	0.06	0.07	0.94
Filler pick 0-40	154	74	80	0.14	0.16	0.90
Rejection 0-40	134	71	63	0.14	0.13	1.09
Filler pick 50-70	141	66	75	0.13	0.15	0.85
Rejection 50-70	156	64	92	0.13	0.19	0.67
Filler pick 80-100	32	20	12	0.04	0.02	1.61
Rejection 80-100	124	51	73	0.10	0.15	0.68
5 high similarity fillers						
Suspect pick 80-100	70	60	10	0.12	0.02	6.17
Suspect pick 50-70	90	61	29	0.12	0.06	2.16
Suspect pick 0-40	45	23	22	0.05	0.04	1.08
Filler pick 0-40	216	99	117	0.20	0.23	0.87
Rejection 0-40	132	55	77	0.11	0.15	0.73
Filler pick 50-70	166	73	93	0.15	0.18	0.81
Rejection 50-70	155	66	89	0.13	0.18	0.76
Filler pick 80-100	49	22	27	0.04	0.05	0.84
Rejection 80-100	73	32	41	0.07	0.08	0.80

Note: Alternative ROC curve operating point order (i.e., when not ordering by diagnosticity ratio) as shown above

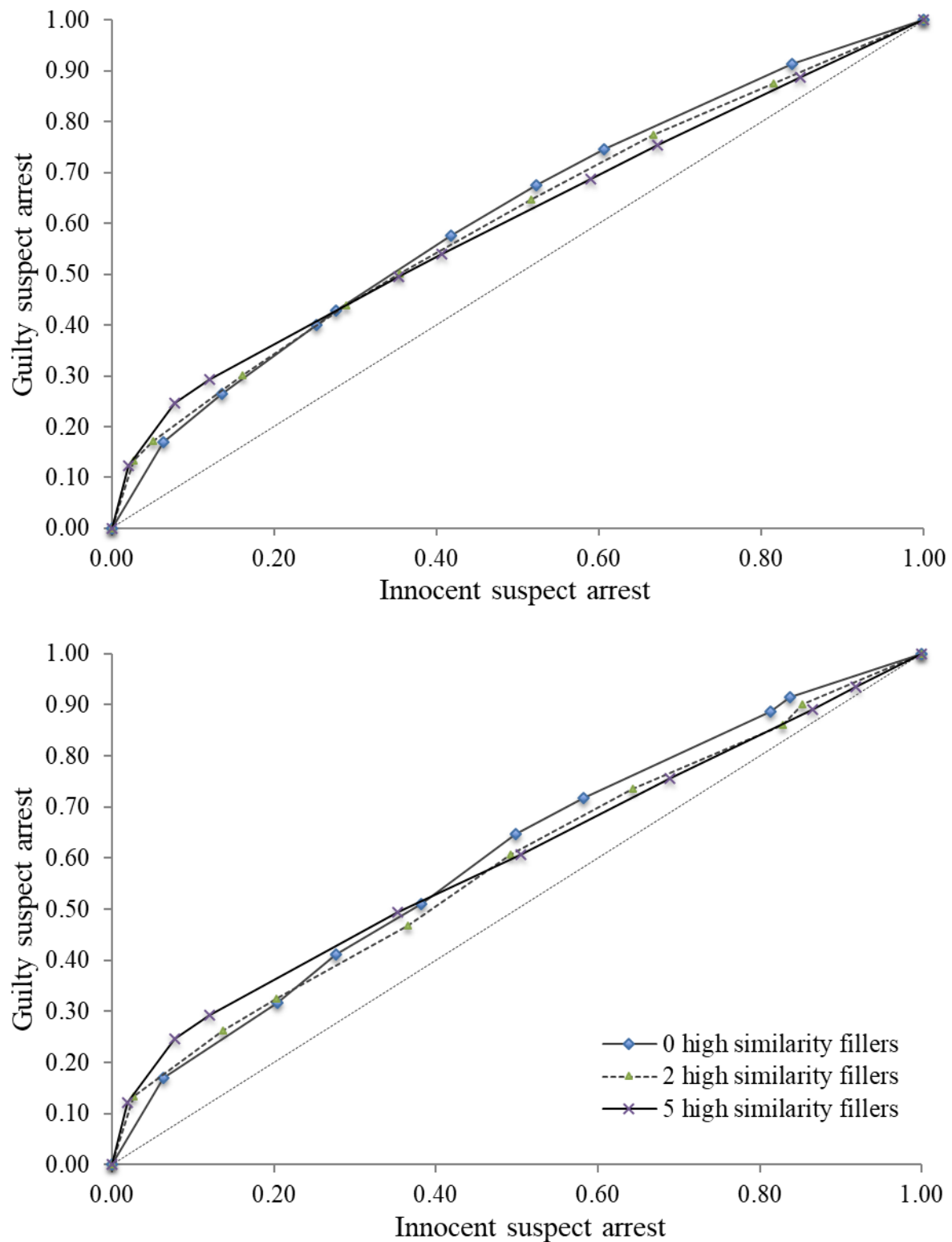


Figure S9. Experiment 2. Full ROCs with operating points ordered by diagnosticity ratio (top panel) and a priori order (bottom panel). Area under the curve values when operating points ordered by diagnosticity ratio = .613, 95% CI [.579, .647], .605, 95% CI [.571, .640] and .603, 95% CI [.567, .637] in the 0, 2 and 5 high similarity filler conditions, respectively (p s >

.666, $D_s < 0.44$). Area under the curve values when operating points ordered by a priori order = .602, 95% CI [.568, .636], .586, 95% CI [.551, .621] and .597, 95% CI [.562, .631] in the 0, 2 and 5 high similarity filler conditions, respectively ($p_s > .535$, $D_s < 0.63$).

Table S34

Experiment 2. Hit Rate (HR), False Alarm Rate (FAR) and Diagnosticity Ratio (DR) for Each Confidence Category in the Short and Long Retention Interval Conditions

Confidence categories	N	Hits	False alarms	HR	FAR	DR
Short retention interval						
Suspect pick 80-100	227	185	42	0.31	0.07	4.41
Suspect pick 50-70	168	101	67	0.17	0.11	1.51
Suspect pick 0-40	34	21	13	0.04	0.02	1.62
Filler pick 0-40	69	27	42	0.05	0.07	0.64
Rejection 0-40	74	30	44	0.05	0.07	0.68
Filler pick 50-70	128	50	78	0.08	0.13	0.64
Rejection 50-70	216	75	141	0.13	0.24	0.53
Filler pick 80-100	54	28	26	0.05	0.04	1.08
Rejection 80-100	225	80	145	0.13	0.24	0.55
Long retention interval						
Suspect pick 80-100	42	29	13	0.03	0.01	2.22
Suspect pick 50-70	188	100	88	0.11	0.10	1.13
Suspect pick 0-40	159	81	78	0.09	0.09	1.03
Filler pick 0-40	404	196	208	0.22	0.23	0.94
Rejection 0-40	320	165	155	0.18	0.17	1.06
Filler pick 50-70	257	125	132	0.14	0.15	0.94
Rejection 50-70	296	140	156	0.15	0.17	0.89
Filler pick 80-100	53	28	25	0.03	0.03	1.12
Rejection 80-100	97	46	51	0.05	0.06	0.90

Note: Alternative ROC curve operating point order (i.e., when not ordering by diagnosticity ratio) as shown above

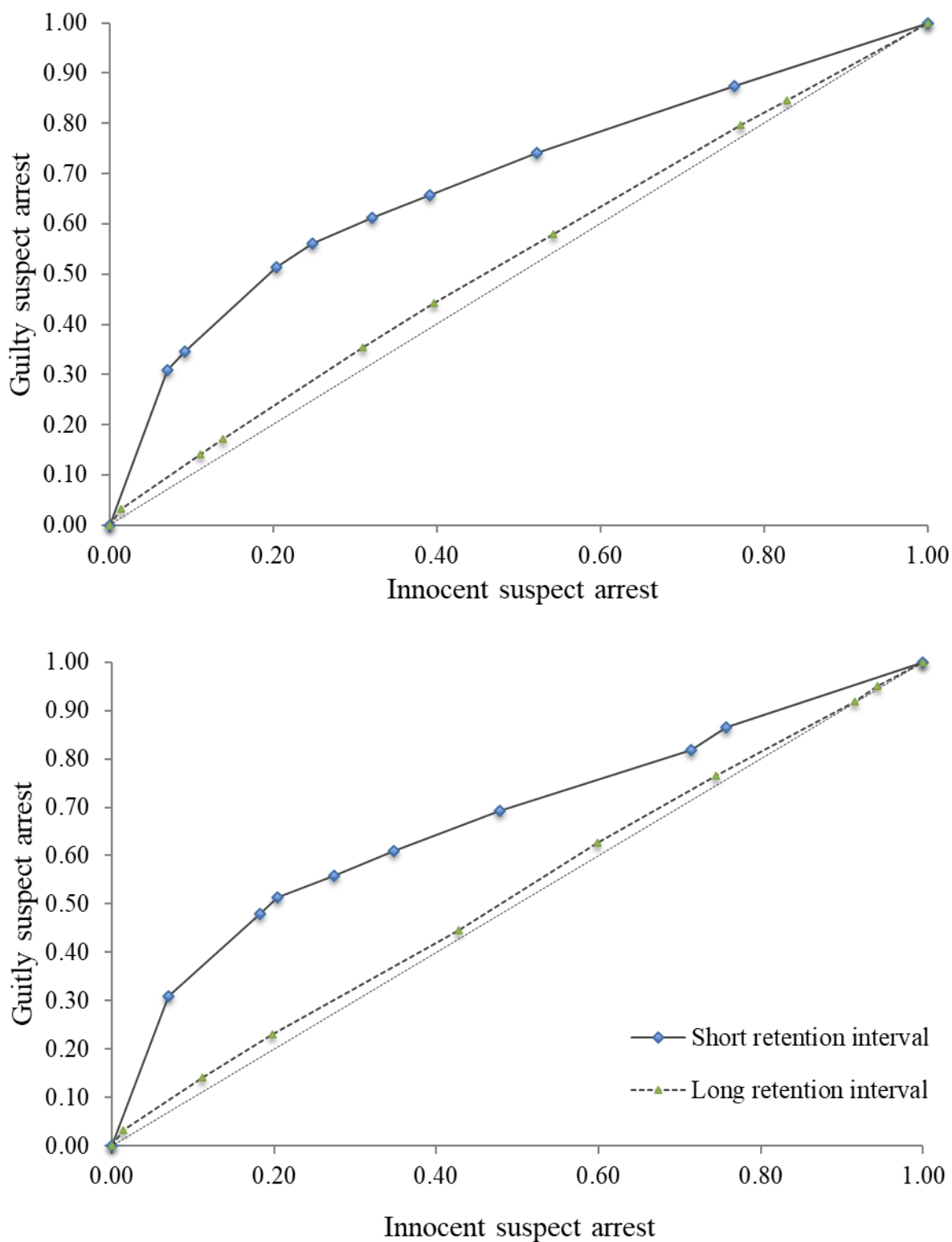


Figure S10. Experiment 2. Full ROCs with operating points ordered by diagnosticity ratio (top panel) and a priori order (bottom panel). Area under the curve values when operating points ordered by diagnosticity ratio = .685, 95% CI [.655, .715] and .530, 95% CI [.504, .556] in the short and long retention interval conditions, respectively. Area under the curve

values when operating points ordered by a priori order = .673, 95% CI [.643, .702] and .521, 95% CI [.495, .548] in the short and long retention interval conditions, respectively.

Supplemental Material: Partial ROC Curves

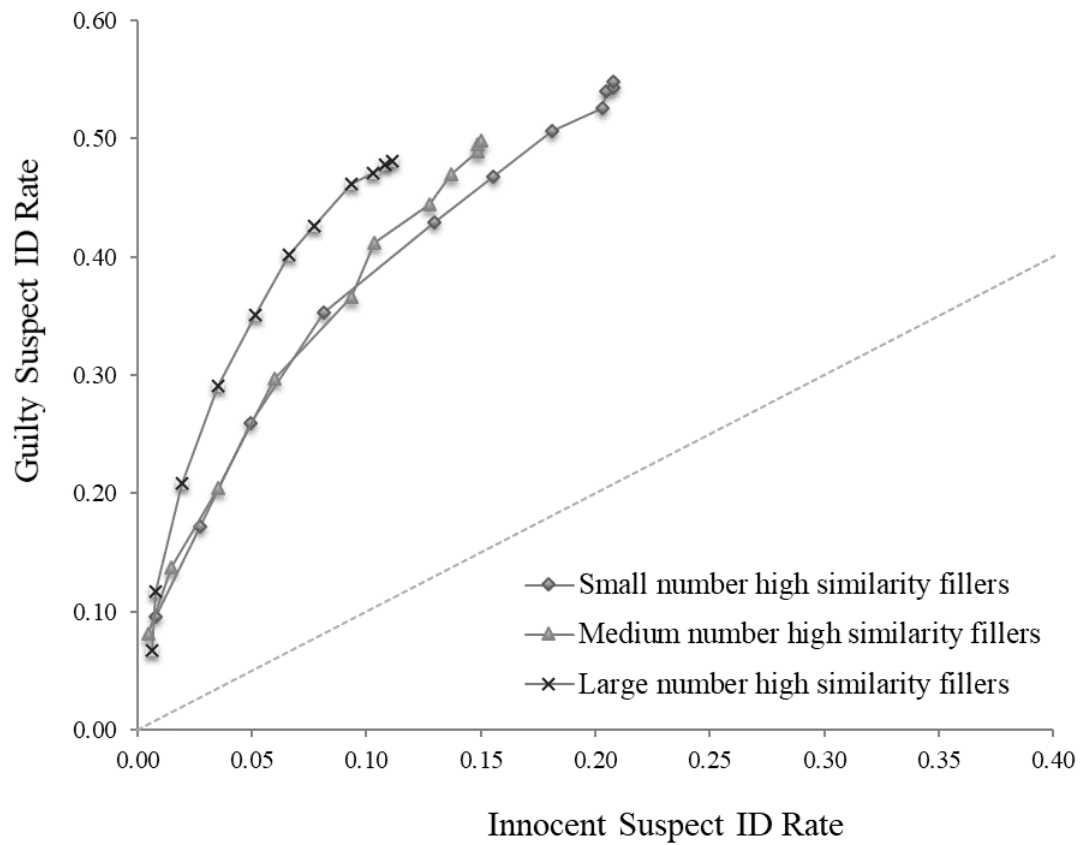


Figure 11. Experiment 1. Partial ROCs across lineups containing small, medium or large numbers of high similar fillers. Partial under the curve values = .070, 95% CI [.061, .080], .073, 95% CI [.063, .082], and .082, 95% CI [.073, .091] in the small, medium and large number of high similarity fillers conditions, respectively ($p_s > .070$, $D_s < 1.46$).

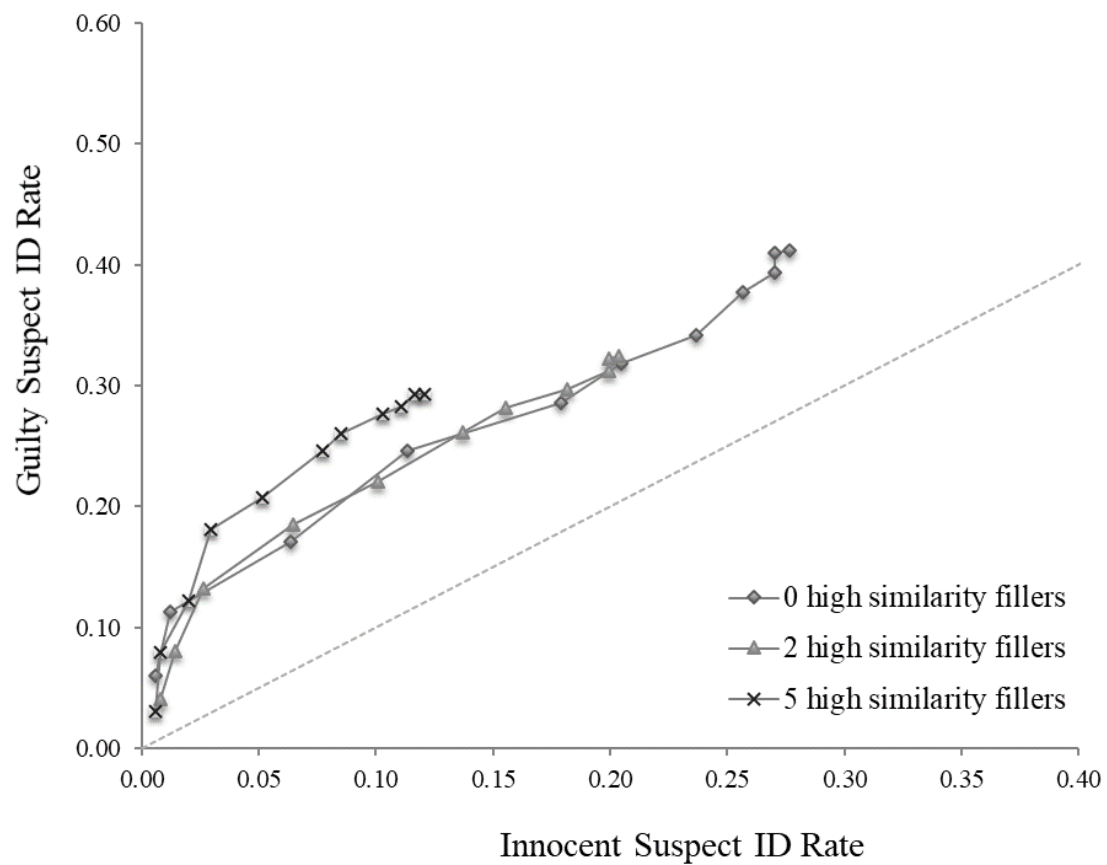


Figure S12. Experiment 2. Partial ROCs across decisions made from lineups containing 0, 2 or 5 high similarity fillers. Partial area under the curve values = .071, 95% CI [.060, .083], .071, 95% CI [.060, .082] and .082, 95% CI [.071, .094] in the 0, 2 and 5 high similarity filler conditions, respectively ($ps > .164$, $Ds < 1.39$).

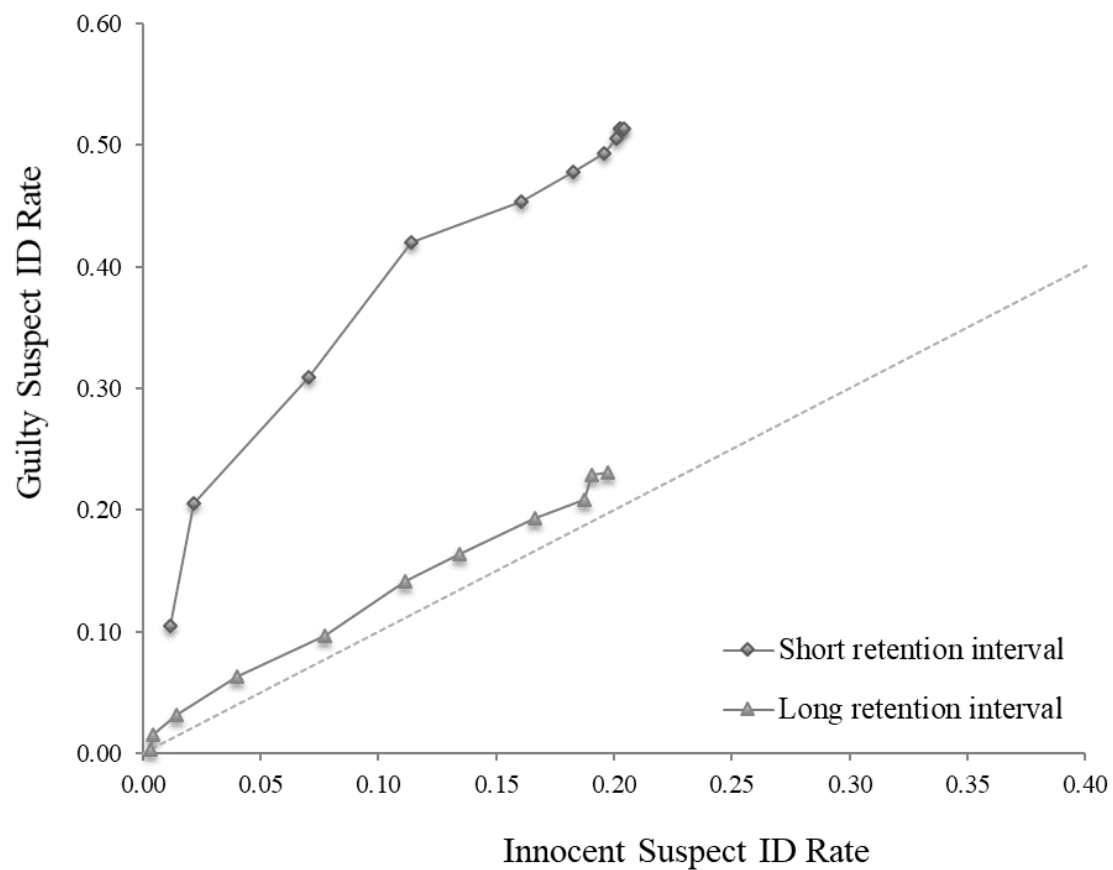


Figure S13. Experiment 2. Partial ROCs across decisions made after the short and long retention intervals. Partial area under the curve values = .070, 95% CI [.061, .079] and .025, 95% CI [.020, .030], in the short and long retention interval conditions, respectively ($p < .001$, $D = 8.59$).

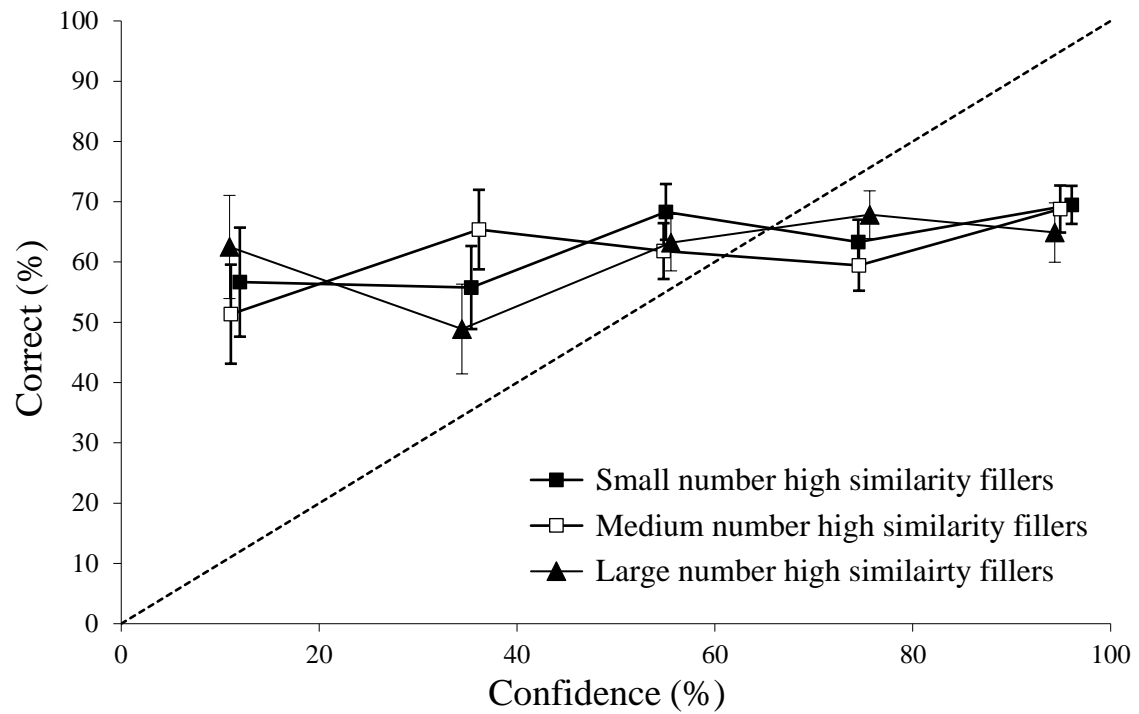
Supplemental Material: Calibration Curves

Figure S15. Experiment 1. Calibration curves for non-choosers across lineups containing small, medium and large numbers of high similarity fillers.

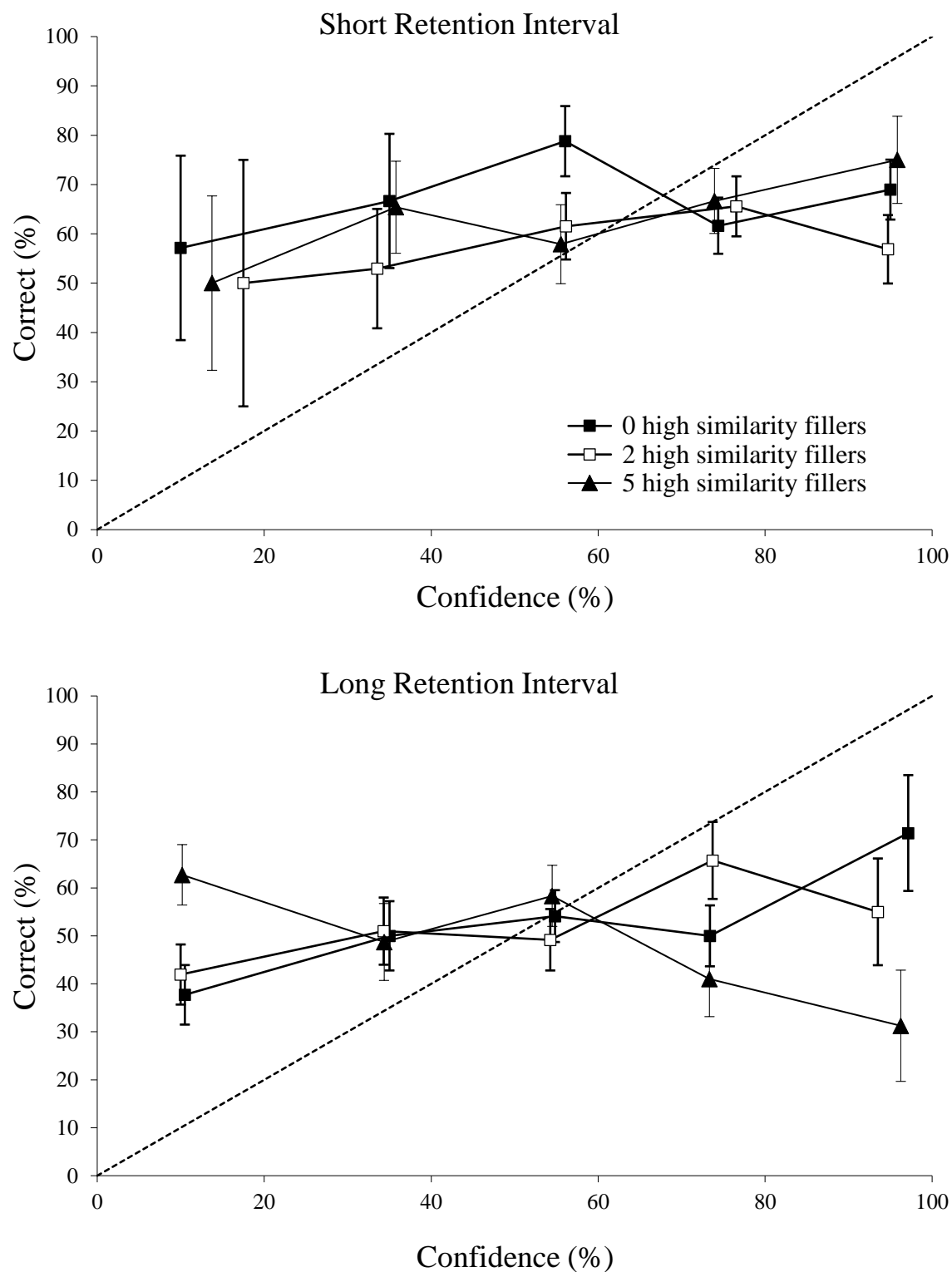


Figure S16. Experiment 2. Calibration curves for non-choosers across lineups containing 0, 2 and 5 high similarity fillers, show separately for identifications made after a short (top panel) or long (bottom panel) retention interval