Table 1a
Between- and Within-cluster variance, and Intraclass Correlations for Judgments and Justifications
in the Peer and Home Contexts

Outcomes	$S_b{}^2$	$S_w{}^2$	ρ	ESS
Peer Context Judgments				
Undifferentiated Exclusion	0.37	2.27	0.14	303
Group-based Exclusion (GBE)	0.49	2.57	0.16	304
Parent-sanctioned GBE	0.73	3.28	0.18	304
Peer-sanctioned GBE	0.62	3.25	0.16	304
Peer Context Justifications				
Societal	0.002	0.18	0.01	300
Victim Blaming	0.02	0.22	0.07	302
Personal Choice	0.0003	0.24	0.001	300
Moral	0.01	0.11	0.07	302
Excluder Empathy	0.01	0.15	0.05	301
Home Context Judgments				
Undifferentiated Exclusion	0.73	2.53	0.22	305
Group-based Exclusion	0.78	2.89	0.21	305
Parent-sanctioned GBE	0.82	2.71	0.23	306
Peer-sanctioned GBE	0.73	2.90	0.20	305
Home Context Justifications				
Societal	0.01	0.18	0.04	301
Victim Blaming	0.01	0.24	0.02	301
Personal Choice	$1.437_E{}^{-15}$	0.23	0.00	300
Moral	0.01	0.10	0.05	301
Excluder Empathy	0.73	2.90	0.20	305

Note. $S_b{}^2$ = between-cluster variance, $S_w{}^2$ = within-cluster variance, ρ = intraclass correlation; ESS = effective
sample size. Variance components were calculated using the lmer function of the lme4 package in r. Intraclass
correlations were calculated using the formula $\rho = S_b{}^2/(S_b{}^2 + S_w{}^2)$ (see Killip, Mahfoud, & Pearce, 2004).

Table 2a

Wild Cluster Bootstrapping for Intervention Effects on Undifferentiated Exclusion Judgments

Group	Pre-test		Post-test		6-month follow-up	
	Pre-Bootstrap	Wild Cluster Bootstrap	Pre-Bootstrap	Wild Cluster Bootstrap	Pre-Bootstrap	Wild Cluster Bootstrap
	B (SE) CI	*B (SE)* CI	*B (SE)* CI	*B (SE)* CI	*B (SE)* CI	*B (SE)* CI
Peer Context						
Skills	-.47 (.21)* [-.88, -.06]	-.47 (.12)*** [-.71, -.23]	-.14 (.19) [-.51, .23]	-.14 (.11) [-.36, .08]	.11 (.18) [-.25, .47]	.11 (.14) [-.17, .39]
Skills + Contact	-.07 (.21) [-.48, .34]	-.07 (.23) [-.52, .39]	.23 (.19) [-.14, .60]	.23 (.20) [-.16, .62]	.50 (.18)** [.14, .86]	.50 (.16)** [.19, .82]
Home Context						
Skills	-.43 (.22)+ [-.87, .01]	-.43 (.22)+ [-.87, .01]	.17 (.20) [-.24, .57]	.17 (.20) [-.24, .57]	.35 (.19)+ [-.02, .72]	.35 (.17)* [.02, .69]
Skills + Contact	.11 (.22) [-.34, .55]	.11 (.26) [-.40, .62]	.63 (.21)** [.23, 1.04]	.63 (.26)* [.12, 1.14]	.90 (.19)*** [.53, 1.27]	.90 (.19)*** [.53, 1.27]

Note. Regression models for treatment group effects before and after wild cluster bootstrapping (Cameron, Gelbach, & Miller, 2008), at pre-test, post-test, and 6-month follow-up. Ethnicity and gender were controlled for in the models. Referent group = control group. All confidence intervals are calculated as 95% confidence intervals. *B* = treatment effect slope, *SE* = standard error, CI = confidence interval.

+p < .10 *p < .05. **p < .01. ***p < .001.

Table 3a
Wild Cluster Bootstrapping for Intervention Effects on Undefined Group-based Exclusion Judgments

	Pre-test		Post-test		6-month follow-up	
	Pre-Bootstrap	Wild Cluster Bootstrap	Pre-Bootstrap	Wild Cluster Bootstrap	Pre-Bootstrap	Wild Cluster Bootstrap
Group	B (SE) CI	B (SE) CI	B (SE) CI	B (SE) CI	B (SE) CI	B (SE) CI
Peer Context						
Skills	-.50 (.22)* [-.94, -.06]	-.50 (.23)* [-.95, -.05]	.40 (.18)* [.04, .76]	.40 (.16)* [.09, .72]	.74 (.19)*** [.37, 1.10]	.74 (.15)*** [.45, 1.04]
Skills + Contact	-.36 (.22) [-.80, .08]	-.36 (.35) [-1.05, .33]	.41 (.18)* [.05, .77]	.41 (.20)* [.02, .80]	.76 (.19)*** [.40, 1.13]	.76 (.19)*** [.40, 1.13]
Home Context						
Skills	-.47 (.24)* [-.94, -.001]	-.47 (.20)* [-.86, -.08]	.43 (.21)* [.01, .84]	.43 (.15)** [.14, .73]	.64 (.20)** [.24, 1.04]	.64 (.14)*** [.37, .92]
Skills + Contact	-.12 (.24) [-.59, .35]	-.12 (.27) [-.65, .41]	.63 (.21)** [.22, 1.04]	.63 (.20)** [.24, 1.02]	.97 (.20)*** [.57, 1.37]	.97 (.17)*** [.64, 1.31]

Note. Regression models for treatment group effects before and after wild cluster bootstrapping (Cameron, Gelbach, & Miller, 2008), at pre-test, post-test, and 6-month follow-up. Ethnicity and gender were controlled for in the models. Referent group = control group. All confidence intervals are calculated as 95% confidence intervals. *B* = treatment effect slope, *SE* = standard error, CI = confidence interval.
[+]p < .10 *p < .05. **p < .01. ***p < .001.

Table 4a

Wild Cluster Bootstrapping for Intervention Effects on Peer-sanctioned Group-based Exclusion Judgments

	Pre-test		Post-test		6-month follow-up	
	Pre-Bootstrap	Wild Cluster Bootstrap	Pre-Bootstrap	Wild Cluster Bootstrap	Pre-Bootstrap	Wild Cluster Bootstrap
Group	B (SE) CI	B (SE) CI	B (SE) CI	B (SE) CI	B (SE) CI	B (SE) CI
Peer Context						
Skills	-.49 (.25)* [-.98, -.01]	-.49 (.14)*** [-.77, -.22]	.39 (.22)+ [-.04, .82]	.39 (.17)* [.06, .73]	.81 (.21)*** [.40, 1.22]	.81 (.17)*** [.48, 1.15]
Skills + Contact	-.01 (.25) [-.49, .48]	-.01 (.17) [-.35, .33]	.61 (.22)** [.18, 1.04]	.61 (.20)** [.22, 1]	.99 (.21)*** [.58, 1.40]	.99 (.13)*** [.73, 1.25]
Home Context						
Skills	-.34 (.24) [-.81, .13]	-.34 (.19)+ [-.71, .03]	.37 (.22)+ [-.06, .79]	.37 (.17)* [.04, .71]	.60 (.20)** [.20, .99]	.60 (.16)*** [.29, .92]
Skills + Contact	.15 (.24) [-.32, .62]	.15 (.19) [-.22, .52]	.65 (.22)** [.23, 1.08]	.65 (.21)** [.24, 1.06]	1.06 (.20)*** [.66, 1.45]	1.06 (.17)*** [.73, 1.40]

Note. Regression models for treatment group effects before and after wild cluster bootstrapping (Cameron, Gelbach, & Miller, 2008), at pre-test, post-test, and 6-month follow-up. Ethnicity and gender were controlled for in the models. Referent group = control group. All confidence intervals are calculated as 95% confidence intervals. *B* = treatment effect slope, *SE* = standard error, CI = confidence interval.

+p < .10 *p < .05. **p < .01. ***p < .001.

Table 5a

Wild Cluster Bootstrapping for Intervention Effects on Parent-sanctioned Group-based Exclusion Judgments

Group	Pre-test		Post-test		6-month follow-up	
	Pre-Bootstrap	Wild Cluster Bootstrap	Pre-Bootstrap	Wild Cluster Bootstrap	Pre-Bootstrap	Wild Cluster Bootstrap
	B (SE) CI	*B (SE)* CI	*B (SE)* CI	*B (SE)* CI	*B (SE)* CI	*B (SE)* CI
Peer Context						
Skills	-.47 (.25)+ [-.97, .03]	-.47 (.20)* [-.86, -.08]	.28 (.23) [-.18, .73]	.28 (.17)+ [-.06, .62]	.63 (.22)** [.19, 1.07]	.63 (.16)*** [.32, .95]
Skills + Contact	-.09 (.26) [-.59, .41]	-.09 (.28) [-.64, .46]	.32 (.23) [-.13, .78]	.32 (.26) [-.19, .83]	.77 (.22)*** [.33, 1.21]	.77 (.20)*** [.38, 1.16]
Home Context						
Skills	-.21 (.23) [-.67, .24]	-.21 (.20) [-.60, .18]	.42 (.21)* [.01, .84]	.42 (.17)* [.09, .76]	.46 (.20)* [.06, .86]	.46 (.13)*** [.20, .72]
Skills + Contact	.12 (.23) [-.34, .58]	.12 (.25) [-.54, .12]	.65 (.21)** [.23, 1.07]	.65 (.21)** [.23, 1.07]	.83 (.20)*** [.43, 1.24]	.83 (.18)*** [.48, 1.18]

Note. Regression models for treatment group effects before and after wild cluster bootstrapping (Cameron, Gelbach, & Miller, 2008), at pre-test, post-test, and 6-month follow-up. Ethnicity and gender were controlled for in the models. Referent group = control group. All confidence intervals are calculated as 95% confidence intervals. *B* = treatment effect slope, *SE* = standard error, CI = confidence interval.

+p < .10 *p < .05. **p < .01. ***p < .001.

Table 6a

Means and Standard Deviations for Baseline Equivalency Assessment for Exclusion Judgments and Justifications by Ethnicity and Treatment Group

	Undifferentiated Exclusion		Judgments by Context Group-Based Exclusion		Parent Sanctioned Group-Based Exclusion		Peer Sanctioned Group-Based Exclusion	
	Peer *M(SD)*	Home *M(SD)*	Peer *M(SD)*	Home *M(SD)*	Peer *M(SD)*	Home *M(SD)*	Peer *M(SD)*	Home *M(SD)*
Ethnicity	***	***	***	***	***	***	***	***
Palestinian-Israeli/P-I	2.85(1.48)	2.20(1.38)	4.07(1.77)	3.12(1.71)	2.92(1.79)	2.18(1.38)	3.24(1.84)	2.53(1.55)
Jewish-Israeli/J-I	3.99(1.55)	3.78(1.80)	5.33(1.46)	4.75(1.72)	4.48(1.87)	3.85(1.90)	4.75(1.77)	4.14(1.86)
Treatment Group	-	-	-	-	-	-	-	-
Control	3.62(1.56)	3.11(1.77)	5.01(1.60)	4.15(1.77)	3.91(1.97)	3.06(1.69)	4.19(1.98)	3.41(1.79)
Skills	3.17(1.57)	2.71(1.68)	4.52(1.70)	3.71(1.86)	3.46(1.87)	2.87(1.79)	3.71(1.85)	3.10(1.78)
Skills+Contact	3.53(1.69)	3.20(1.90)	4.63(1.87)	4.01(2.05)	3.79(2.11)	3.16(2.08)	4.15(2.02)	3.54(2.08)

	Social-Conventional		Stereotype		Justification by Context Personal Choice		Moral		Excluder Empathy	
	Peer	Home	Peer	Home	Peer	Home	Peer	Home	Peer	Home
Ethnicity	*	***	***	**	-	-	***	***	*	**
Palestinian-Israeli/P-I	0.57(0.43)	0.71(0.40)	0.53(0.50)	0.50(0.35)	0.40(0.49)	0.39(0.49)	0.17(0.30)	0.12(0.26)	0.15(0.36)	0.12(0.32)
Jewish-Israeli/J-I	0.45(0.41)	0.54(0.44)	0.27(0.44)	.35(0.48)	0.37(0.48)	0.34(0.48)	0.32(0.36)	0.28(0.36)	0.24(0.43)	0.23(0.42)
Treatment Group	-	-	-	-	-	-	-	-	-	*
Control	0.51(0.42)	0.60(0.43)	0.43(0.50)	0.39(0.49)	0.40(0.49)	0.40(0.49)	0.24(0.32)	0.22(0.34)	0.25(0.43)	0.24(0.43)
Skills	0.56(0.42)	0.69(0.43)	0.39(0.49)	0.45(0.50)	0.37(0.49)	0.33(0.47)	0.21(0.32)	0.17(0.30)	0.18(0.38)	0.19(0.39)
Skills+Contact	0.47(0.42)	0.57(0.43)	0.37(0.49)	0.43(0.50)	0.38(0.49)	0.37(0.49)	0.28(0.37)	0.22(0.34)	0.17(0.38)	0.10(0.30)

Note. Significance values for each follow-up comparisons is represented above each pair of means and standard deviations.
* = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$; - = not significant.

Table 7a

Pairwise Comparisons of Means and Standard Error for Justifications by Gender, Ethnicity, and Scenario

	Social-Conventional	Stereotype	Personal Choice	Moral	Excluder Empathy
	M(SE)	*M(SE)*	*M(SE)*	*M(SE)*	*M(SE)*
Overall use of Justification	$0.49(0.02)^{a}$***	$0.29(0.02)^{b}$ bc, bd*	$0.38(0.02)^{c}$ cb*; cd***	$0.30(0.15)^{bc}$	$0.20(0.16)^{d}$ da, dc***; db*
Gender					
Female	0.43(0.03)	0.25(0.03)	0.37(0.03)	0.35(0.02)	0.23(0.02)
Male	0.55(0.03)	0.33(0.03)	0.38(0.03)	0.26(0.02)	0.18(0.02)
Ethnicity	***	***	*	***	**
Palestinian-Israeli	0.57(0.03)	0.37(0.03)	0.42(0.03)	0.22(0.02)	0.16(0.02)
Jewish-Israeli	0.42(0.03)	0.22(0.03)	0.33(0.03)	0.39(0.03)	0.25(0.02)
Scenario	***	-	-	***	-
Peer	0.44(0.02)	0.29(0.02)	0.36(0.02)	0.34(0.02)	0.20(0.02)
Family	0.54(0.02)	0.29(0.02)	0.38(0.02)	0.27(0.02)	0.20(0.02)

Note. Significance values for each pairwise comparison is represented above each pair of Means and Standard Errors; Different superscript letters for the "Overall use of Justification" row indicate significant mean differences across justifications; $* = p < 0.05$; $** = p < 0.01$; $*** = p < 0.001$; - = not significant.