Temporal effects of ostracism:

A meta-analysis of Cyberball studies.

Hartgerink, Chris H.J.1

Van Beest, Ilja1

Wicherts, Jelte M.1

Williams, Kipling D.2

1 Tilburg University, the Netherlands

2 Purdue University, United States of America

**Word count (main text): 8290 words**

**Word count (abstract): 249 words**

**Word count (references, tables, footnotes, supplement): 5761 words**

**Abstract**

We present a meta-analysis of randomized experiments that involved the use of Cyberball to determine whether the impact of ostracism decreases over time. In addition, we study whether moderation of the ostracism effect by theoretically hypothesized factors (e.g., social categorization of Cyberball players) increases over time. We considered both the first and last outcome measures in 120 experiments (combined N = 11,869) and submitted the data to random- and mixed effects models. First, the main analyses showed that (1) the average ostracism effect is large and heterogeneous, (2) the effect and heterogeneity diminishes from the first to the last measure, (3) both the first and last measure are moderated by a medium interaction effect, and (4) time passed since being ostracized does not predict the effect sizes of the last measure. Second, the structural aspects of the Cyberball game (i.e., number of players, gender composition of sample, origin of study, sample age, duration of ostracism, type of needs scale) did not predict the average ostracism effect. Third, sensitivity analyses using different dependent variables revealed that the average effect was most pronounced for immediate fundamental needs and least pronounced for interpersonal measures. Finally, an analysis of 19 studies using a three-player cyberball game with 30 ball tosses on immediate need satisfaction showed considerable heterogeneity in the average ostracism effect (I2 = 83%). Results corroborate theorizing that ostracism has a large psychological impact. However, the fact that first measures can be moderated and that the average effect has a large heterogeneity adds that immediate responses are more flexible and perhaps less reflexive than previously assumed.

*Keywords: Cyberball, meta-analysis, temporal, ostracism*

Temporal Effects of Ostracism: a Meta-Analysis of Cyberball Studies.

Cyberball (Williams, Cheung, & Choi, 2000) is a virtual ball-tossing game that is used to manipulate the degree of social inclusion, or ostracism, in social psychological experiments. In this game the participant supposedly plays with two other participants, who are in fact part of the computer program. The program varies the degree to which participants are being passed the ball by the other players; ostracized participants are not passed the ball, while included participants are passed the ball. In the study of the psychological effects of ostracism and exclusion, this paradigm has been widely used next to other paradigms, such as the future life rejection (see Baumeister, Twenge, & Nuss, 2002) and the get-acquainted paradigm (Nezlek, Kowalski, Leary, Blevings, & Holgate, 1997). Our literature search showed that at least 200 published papers involved the use of the Cyberball paradigm to study ostracism, and that over 19,500 participants have played the game thus far. Thus, the Cyberball paradigm has received much traction in experimental work in the field of social ostracism.

**Historical background**

Since the introduction of Cyberball, social ostracism research has culminated in updated theory (i.e., Williams, 2009), has been topic of several meta-analyses (Blackhart, Nelson, Knowles, & Baumeister, 2009; Cacioppo, Frum, Asp, Weiss, Lewis, & Cacioppo, 2013; Gerber & Wheeler, 2009), and has received growing interest (Williams & Jarvis, 2006). Because everybody gets excluded sometimes, ostracism research is socially relevant (e.g., in the context of school shootings; Leary, Kowalski, Smith, & Phillips, 2003). Through experimental work, it has been repeatedly shown that being ostracized has an effect on people—either on their psychological functioning (e.g., decreases in mood; Lustenberger & Jagacinski, 2010) or on certain interpersonal behaviors (e.g., increases in aggressive behaviors; Van Beest, Carter-Sowell, Van Dijk, & Williams, 2012). These experiments have highlighted the (mostly negative) impact of ostracism on fundamental needs (e.g., belonging; Baumeister, & Leary, 1995), mood, physiology (e.g., body temperature; Ijzerman, Galucci, Pouw, Weiβgerber, Van Doesum, & Williams, 2012), and various other constructs, including those measured with behavioral measures. In the current paper, we refer to the general effect of being ostracized compared to being included in Cyberball as the ‘ostracism effect.’

Williams (2009) proposed a temporal model of ostracism, in which he suggested three stages in the ostracism effect, namely: (1) a reflexive stage, (2) a reflective stage, and (3) a resignation stage. In the reflexive stage, the response to the ostracism sequence is immediate and occurs like a reflex. This initial response is theorized to be socially painful and threatening (Baumeister & Leary, 1995), and easily detectable due to evolutionary over-sensitivity to cues of ostracism (Haselton, & Buss, 2000). Such a reflex would not take into account situational specifics, and provides no or little room for coping, and the reflex is proposed to affect primarily pain and different fundamental needs. The affected fundamental needs are (1) belonging, (2) self-esteem, (3) control, and (4) meaningful existence. The reflective (or ‘delayed’) stage, which follows this immediate response, is subject to more rational thought and coping with the threats. Part of such coping is need fortification of the affected fundamental needs. The resignation stage occurs after prolonged ostracism, causing prolonged periods of pain and more fundamental need threat. If one is not able to fortify the fundamental needs, a prolonged ostracism sequence leads to feelings of helplessness, alienation, depression, and unworthiness. In this paper we concern ourselves with the reflexive and reflective stages. For these stages, Williams’s temporal model implies (but does not claim) that as time passes, the effects of a *singular* occurrence of social ostracism will diminish. In addition, the model implies that moderation of such effects by other socially relevant factors (e.g., type of group from which one is excluded) will increase over time.

Previous meta-analyses on the topic of social exclusion have focused on estimating the effect of exclusion1 on different constructs (e.g., belonging, self-esteem, etc.; Gerber & Wheeler, 2009), moderation of effects by types of manipulation (Blackhart et al., 2009) and neurophysiological effects in fMRI studies (Cacioppo, et. al, 2013). Gerber and Wheeler (2009) found that rejection has a medium to large effect on mood and fundamental needs, which results in either anti-social responses if control *can* be restored, or pro-social responses if control *cannot* be restored. Blackhart and colleagues (2009) found significant negative effects of rejection and ostracism on emotion, mood, and self-esteem, and found indications of moderation by type of rejection manipulation used. Cacioppo et al. (2013) did not find support for the claim that pain due to ostracism is similar to physical pain, as suggested by Eisenberger, Lieberman, and Williams (2003).

Whereas these meta-analyses focused on the social ostracism effect within different constructs, in different exclusion paradigms, and in fMRI studies, the current meta-analysis limits the paradigm to that of Cyberball and looks to test more general ideas of social ostracism. Here we focus on the workings of the reflexive and reflective stages (rather than on the resignation stage). We included only randomized experiments with between-subjects designs that involved the use of the Cyberball paradigm. The main reasons for these inclusion criteria are (1) to encompass the typical Cyberball experiment, (2) to assure study quality, and (3) to remove the need for within-subjects correction of effect sizes. These criteria were also used to ensure broad enough criteria to include as many studies as possible to test our hypotheses.

**Hypotheses**

The current meta-analysis is primarily concerned with modeling the ostracism effect over time and the degree to which ostracism effects are moderated as time progresses. We also explored whether specific changes to the Cyberball manipulation affect the average effect sizes and whether the average effect sizes differ between different types of dependent variables. Below, we will outline the primary hypotheses and the secondary hypotheses. Confirmatory hypotheses were registered a priori on the Open Science Framework.2

**Primary hypotheses.** The first main question is: does the effect size of ostracism decrease over time? For this, we considered the first and last dependent measure of all selected studies. Since a formal test of time-based change is not possible in the current setting (reasons for this are further explained in methods section), we took two indirect approaches: (1) the average effect size on the last measure is estimated within a mixed-effects model with time passed since the ostracism manipulation as a predictor, and (2) the confidence intervals of the average effects are compared for overlap.

Second, we will inspect the moderation of ostracism over time. Williams’ (2009) temporal model suggests that the reflexive response becomes less as time progresses, after which coping plays a larger role. This implies that as time progresses, moderation of the ostracism effect would *increase*. Our operationalization of *moderation* of the ostracism effect is straightforward: many experimental designs included a second factor besides the factor related to ostracism, and so interactions in the ANOVA realm would indicate moderation of the ostracism effect. For example in a 2 (social status: ostracized vs. included) by 2 (group: in-group vs. out-group) design, the ostracism effect is expected to be larger for the in-group level than for the out-group level. Such moderation can be numerically seen as an interaction effect. For instance, the difference between simple effects of ostracism for the in- and out-group conditions reflects this interaction (specific calculations are reported in the methods section and formulae in the Appendix).

**Secondary hypotheses.** In addition, we assessed how structural aspects of Cyberball studies affect first and last measures. We ….

We consider structural aspects associated with Cyberball. We assessed whether it country of origin in which the study was conducted would impact cyberball experience. We assessed. (e.g., number of ball tosses, number of Cyberball players, gender of participants, etc.) and differences between dependent variables for moderation and robustness, respectively. These elements can provide nuances in the composition of the estimates.

We will use meta-regressions with study-level indicators to study effects of the composition of the manipulation and the samples. This is conceptually similar to a (multiple) linear regression, but with estimated average effects as the dependent variable. Our predictors for these meta-regressions are the study-level indicators. First, because collectivism might influence the degree to which belonging is important (see Hofstede, 1980), we included a variable which gives a crude categorization of continents (i.e., U.S., other western countries, Asian countries, and residual countries). Second, because social aspects may be less evolutionarily relevant for males than for females, we included proportion of male participants. Third, given that exclusion may be more relevant for younger people, we included mean age. Fourth, because it may matter by how many people one gets excluded, we included the number of players in the game. Fifth, as the length of the exclusion may matter, we included duration of exclusion. Sixth, we considered potential differences between types of scales used to measure fundamental needs.

Additionally, we inspected robustness of the findings across different subsets of dependent variables. Overall, the dependent variables included in the meta-analysis were only subject to the criterion that they were affected by ostracism, which does not limit the measures to just one type. In other words, we included multiple types of dependent variables with varying psychometric properties in the primary studies. The preponderance of measures used in Cyberball studies concern ‘fundamental needs’ (i.e., belonging, self-esteem, control and meaningful existence) questionnaires, which were developed within the Cyberball paradigm (see Van Beest & Williams, 2006; Williams et al., 2000; Zadro, Williams, & Richardson, 2004). Next to these fundamental need questionnaires, the measures vary widely and no clear grouping can be made besides intrapersonal measures (i.e., measures relating to the self only) and interpersonal measures (i.e., measures relating to others as well). Examples are self-reported hostile feelings and donation behavior, respectively. So the estimated effect size will be an overall estimate and not one for a specific measure, such as belonging. The types of measures are subdivided into fundamental needs, intrapersonal and interpersonal measures and these subsets are tested separately as a sensitivity analysis of the overall results.

Besides using different subsets of measures for sensitivity analyses, we also coded whether the first- and last measure included was immediate (i.e., questions relating to during the game) or delayed (i.e., questions relating to after the game). This addresses the possible confounding element in measure selection from a substantive perspective, where first measures need not be immediate, and last measures need not be delayed.

In sum, the hypotheses are subdivided into two primary and several secondary questions. The two main questions underlying the primary hypotheses are: (1) does the ostracism effect diminish over time?, and (2) does the moderation of the ostracism effect, due to crossed factors, increase over time? The question underlying the secondary hypotheses is: do study characteristics affect the estimated average effect? In other words, are the results robust across different subsets that substantively could yield different results? These questions will be answered with random and mixed-effects meta-analytic models applied to 120 studies.

**Method**

**Study inclusion criteria**

First, we only considered experiments that incorporated a between-subjects design with random assignment. Within-subject designs were excluded, because most within-subjects designs regard high-dimensional neurophysiological measurements such as fMRI that are beyond the scope of this meta-analysis (see Cacioppo, et. al, 2013). Also, meta-analyses of effects of within-subjects designs require the correlations between measures in primary studies, and we did not expect these to be reliably reported in the papers.

Second, experiments were required to have a factor that manipulated number of virtual ball tosses obtained by the participants. For this ostracism factor we only considered the condition in which participants were ostracized by all other participants and the condition in which participants were equally included by all other players. Studies that used other (between-subjects) factors alongside the ostracism factor were included as well. Limiting the study designs included is presumed to decrease variability due to design characteristics, which increases power for moderator analyses. Note that we collapsed effect sizes across irrelevant factors if primary authors expressed no expectations concerning the potential moderating effect of that crossed factor (i.e., non-moderating factors). Moreover, continuous variables that were dichotomized were also collapsed due to the many problems dichotomization can cause (e.g., underestimation of effect size, spurious effects; see Hunter & Schmidt, 1990; MacCallum, Zhang, Preacher, & Rucker, 2002). For example, when participants were grouped into high- and low neuroticism groups based on a continuous measure of neuroticism (Boyes, & French, 2009), we used pooled means and standard deviations across these two groups, reducing the design to an ostracism/inclusion design.

Reasons for these inclusion criteria are threefold. (1) Most Cyberball experiments take place in such a format, making it an encompassing criterion for the purposes of this meta-analysis. (2) The choice to limit the meta-analysis to between-subject designs rendered computational aspects more feasible based on reported statistics in papers. (3) These criteria were assumed to heighten primary study quality, which is preferable to subjective quality assessments of individual studies. For the dependent measures the criterion was that they were (expected to be) affected by the ostracism manipulation. We considered the measures that immediately followed the manipulation (first measure) and the measure at the end of the study (last measure), while excluding manipulation checks in this assessment.

**Literature search**

To have a comprehensive meta-analysis of Cyberball studies, we used seven search strategies in the period of November 2012 through April 2013. These search strategies included (1) database searches, (2) a call for data, (3) cross-reference with Kip Williams’ list of Cyberball studies, (4) Google Scholar alerts, (5) citation records, (6) SPSP conference abstracts, and (7) personal communication.

The databases searched included Web of Knowledge, PubMed, ScienceDirect and Worldcat using the sources from Tilburg University. The first three cover only published articles, whereas Worldcat also covers books and dissertations as well as the PsycINFO database. All these databases were searched with the keywords “cyberball”, “ball-tossing” and “ball AND ostraci\*”. Web of Knowledge was the first database searched. For this database, an additional search term (i.e., “ball AND exclu\*”) was used, but this yielded zero relevant hits on a total of 501 hits. Across all these searches, results included 1927 hits of which 109 were saved for coding. Within Web of Knowledge, we looked through all citation records of the seminal papers by Williams and colleagues (2000); Williams and Jarvis (2006). These papers were cited 332 times (as of 5th of November, 2012), of which 43 papers were saved for coding. This literature search provided 2259 initial hits (including possible duplicates across searches), of which 152 were selected to be included in the coding.

The call for data was put on the list servers or forums of Society for Personality and Social Psychology (SPSP), European Association of Social Psychology (EASP), and Social Psychology Network (SPN; all on 3rd of December, 2012). This resulted in nine replies, from which three useful studies, which were included in the coding procedure.

Kip Williams keeps a list of Cyberball studies on his personal website. This list was used to check for extra articles that did not turn up in the initial searches on November 15th, 2013.3 The list included 93 papers, of which 9 papers were included to be coded.

The final searches included Google Scholar alerts, SPSP conference abstracts and personal communication. The Google Scholar alerts were used to keep up to date with new literature. These alerts notify a user when new hits for a search term occur, and were used for “cyberball” and “ball-tossing”, which yielded 85 hits of which 25 were saved for coding. SPSP conference abstracts from 2006 through 2013 were searched for Cyberball studies. This led to personal communications with the authors of the conference abstracts, leading to additional studies. Pooled, the personal communication and the conference abstracts yielded 21 hits, of which 20 were saved for coding. The seminal paper by Williams and colleagues (2000) was added separately.

In sum, the literature search spanned 2,468 hits, resulting in 205 that were saved for coding. During coding, papers were assessed to fit the inclusion criteria. Of the 205 papers, 107 papers were excluded for a variety of reasons. Several involved the use of a within-subjects design (52 papers). Some papers could not be accessed (5 papers) or could not be included because we did not receive the required data on request (7 papers). Some were excluded for other reasons (43 papers), such as not involving new data (e.g., a dissertation study that was later published). All included papers were published between 2000 (after the introduction of Cyberball) and April 2013. This resulted in a final, fully coded sample of 98 papers containing 120 studies, with mean sample size 98.9 and median sample size 74.4 There were a total of 11,869 Cyberball participants.

**Coding procedure**

The first author coded all the studies and conducted all the analyes. The third author double-checked a subset of all studies, the second author double-checked all 52 studies that entailed a full two-by-two design. The and reran the R-code of all analyses Finally, an extensive account of all coding decisions is publicly available via Open Science Framework on a paper-by-paper basis (see Footnote 2 for the direct link).

Group means and standard deviations were retrieved for both the first and last relevant measure in each study for effect size calculation. Relevant measures were defined as constructs that were expected by primary authors to show an ostracism effect (e.g. fundamental needs, mood, pro-social helping behavior, etc.). Coding that was crucial for testing the confirmatory hypotheses concerned the amount of items from the first through last measure plus any additional time in between (e.g., rest period). This made up the estimation of time from the first to last measure, where each item was counted as lasting six seconds (the six-second rule was based on a longstanding practice used to estimate average completion time in the freshmen testing program of the University of Amsterdam; e.g., Smits, Dolan, Horst, Wicherts, & Timmerman, 2011). Any additional time reported in the procedure was also included. Note that some measures are variable on time (e.g., persistence tasks) and that these were arbitrarily estimated in a conservative manner to at least take these measures into account at some level.

The type of measure used was coded for in the following general terms: (1) fundamental needs, (2) intrapersonal, and (3) interpersonal. Intrapersonal measures were defined as measuring constructs that relate only to the self (e.g., ‘how angry do you feel?’, physiological measures, etc.). Interpersonal measures were defined as measuring constructs that relate to (the self and) others (e.g., ‘how angry do you feel towards person X?’, donations to charity, etc.). For the exploratory analyses, we coded sample characteristics (e.g., age, gender composition), Cyberball characteristics (e.g., amount of players, length of game), measure properties (e.g., intra- or interpersonal), and whether the first- and last measure fit the definition of immediate (i.e., during the game) or delayed (i.e., after the game/now), respectively.

Because relevant measures were defined broadly we included different kinds of measures that are expected to show different directions of an ostracism effect. For example, when compared to included participants belongingness scores are expected to be lower for ostracized participants, whereas retaliation scores are expected to be higher for ostracized participants. To counteract computational problems (i.e., cancellation of effects) being caused by this bidirectionality of ostracism effects, we coded the direction of the ostracism effect for each specific measure, such that negative effect sizes depict negative psychological effects. Moreover, in two-by-two designs in which the ostracism effect was crossed with another factor (i.e., a moderator), we coded for expected direction of that moderator. For example, in Table 1, we show hypothetical data for the four study designs that are possible when crossing direction of the effect and direction of the moderation. The relevant effect sizes should be corrected to attain comparable effect sizes across studies. Effect sizes for the simple ostracism effect (column wise) were corrected only for the type of measure. For instance, for panels (a) (involving, e.g., need threat) and (b) (involving, e.g., need satisfaction), the corrections entailed a multiplication with -1 or +1, respectively. Simple moderator effects (row wise comparisons) are interesting for understanding the effect of the moderator under either ostracism or inclusion. These simple moderator effects were corrected for both the type of measure *and* the expected moderation (i.e., exacerbation, -1, or minimization, +1). For example in panel (c), the 5 and 8 on the right are used to compute the “standard ostracism effect” (as in Williams et al., 2000), whereas the 3 and 8 in the left column represent an ostracism effect that is thought to be exacerbated. For example, in a given ostracism study with a two-by-two design, adolescents are expected to show stronger ostracism effects, compared to young adults (Pharo, Gross, Richardson, & Hayne, 2011). The 5 and 8 would subsequently represent the scores for the young adults, whereas the 3 and 8 would represent the scores for the young adolescents. In panel (d) we depict a study in which the “moderated” column is thought to lead to a minimal ostracism effect, as could be expected when Cyberball is played with members of a despised out-group (Gonsalkorale & Williams, 2007). The margins (greyed out) denote the simple effects, which are after correction comparable across all panels (a) through (d), indicating that this correction did what we intended it to.

Relevant information that was missing in the papers was requested from the authors via e-mail. In case of non-response, three follow-up e-mails were sent. All this communication was documented and can be found on the OSF page for this project. In case of non-response or non-willingness to send data, studies were either eliminated if the information was crucial (i.e., means and standard deviations of the measures per group), computed if possible (i.e., cell sizes), or assumed if deemed reasonable on the basis of additional information. For instance, when no information was given we considered the Cyberball manipulation characteristics to be similar to previous studies in the same paper or in earlier papers referred to in the paper.

**Statistical analyses.**

For the analyses, we used the “metafor” package (Viechtbauer, 2010) in the R statistical environment (R Core Team, 2013).

**Effect size metric.** We used Hedges’ g version of the standardized mean differences as the effect size. Hedges’ g corrects for the slightly biased estimate given by Cohen’s d (Hedges, 1981). Standardized simple effects were calculated across the ostracism factor and the interaction effect was calculated by taking the standardized difference between the unstandardized main effects (see the Appendix for the exact formulae used). This was done for both the first and last dependent variable in each experiment. For example, in a 2 (social status: ostracized vs. included) by 2 (moderator: present vs. absent) design with multiple measures, we calculated two simple ostracism effects (Hypothesis 1) and two interaction effects (Hypothesis 2). Non-factorial studies delivered only simple effects for the first and last measure and no interactions.

**Meta-analytic model.** We used random- and mixed-effects models, since heterogeneity in the effect sizes is expected due to the different measures included and additional unknown methodological and substantive factors. The meta-regression element in some of the analyses is the variable time as predictor of the ostracism effect. Analyses without this study-level predictor reduce to a random-effects model. We used Restricted Maximum Likelihood (REML) to estimate tau-squared (i.e., the (residual) variance), as recommended by Viechtbauer (2005). Note that when estimating a mixed- or random effects model, one does not estimate a single ‘true’ effect, but rather the mean and variance of underlying effects (Viechtbauer, 2005).

**Sensitivity analyses.** To test for robustness of the effects, we incorporated several sensitivity analyses. We flagged possibly problematic outliers on the basis of studentized deleted residuals, Q-Q plots, and Cook’s distance values. Subsequently, we inspected the effect of these outliers on substantial results in sensitivity analyses in which these outliers were excluded. Another sensitivity analysis entailed fitting of the mixed-effects model with tau-squared fit at the upper bound value of the 95% confidence interval.

**Funnel plot asymmetry.** A funnel plot depicts each study’s effect size against its standard error (Light & Pillemer, 1984). Larger studies have smaller standard errors, and vice versa for smaller studies. Following from a theoretical fluctuation of the population effect size due to sampling variance, a funnel plot should be symmetrical around the estimated mean effect size. If there are no methodological or substantive reasons to expect a link between effect sizes and standard errors, funnel plot *asymmetry* can indicate publication bias (Bakker et al, 2012). To test funnel plot asymmetry, we used Egger’s regression test (Egger, Smith, Schneider, & Minder, 1997) for mixed-effects models (Sterne & Egger, 2005). This tests whether the distribution of effect sizes is equal on both sides of the average effect, when accounting for true heterogeneity. Funnel plot asymmetry thus indicates bias in the estimated mean effect size, and possibly publication bias.

**Results**

In our reporting of the effect sizes, d indicates a main effect and Δd indicates an interaction effect. Even though Hedges’ g effect sizes were used, the notation of d was maintained, since it is only a minor correction to Cohen’s d. Sensitivity analyses are only reported if they showed different effects (all sensitivity analyses can be freely inspected via the files available on OSF).

**Primary hypotheses.** The two confirmatory hypotheses are tested in four meta-analyses, of which the study level effects are reported in Table 2. The table includes effect sizes used in the estimation of the average simple effect of ostracism on the first measure, the average simple effect on the last measure and the estimation of the average interaction effect on both the first and last measure.

**Simple ostracism effect (Hypothesis 1).** In a random-effects model on the main effect of ostracism (*k* = 120), heterogeneity was significant, Q (119) = 1395, p < .0001, I2 = 92.99% and estimated τ2 = 0.90, 95% CI [0.70, 1.24]. The analysis yielded an estimated average effect of d = -1.36, p < .0001, 95% CI [-1.54, -1.18]. A random-effects version of the Egger’s test (Egger et al., 1997; Sterne & Egger, 2005) indicated funnel plot asymmetry, z = -6.14, p < .0001. Due to the size of the average effect, and hence large power to acquire significant outcomes in primary studies, we do not suspect publication bias to explain this asymmetry. In other words, immediately after being ostracized, the average ostracism effect is estimated at -1.36 standard deviation units, which entails a large effect (Cohen, 1988).

Next, we fitted a mixed-effects regression model for the ostracism effect on the last measure (*k =* 95), including estimated time in seconds since completing the Cyberball game as predictor. Residual heterogeneity was significant, QE (93) = 803, p < .0001 and estimated τ2 = 0.38, 95% CI [0.27, 0.54]. The intercept was estimated to be: dintercept = -0.76, p < .0001, 95% CI [-0.91, -0.61]. Contrary to our expectation, the estimated time in seconds failed to moderate the average effect, b = 0.0001, p = .187, 95% CI [-0.0001, 0.0003]. However, we have to take into consideration the low power of the moderation analyses due to the large (residual) heterogeneity in effect sizes (Hedges & Pigott, 2004). A regression test for mixed-effects model with moderator (i.e., including both the time and SE as predictor) showed no funnel plot asymmetry, z = -0.72, p = .474. In short, long after ostracism has occurred (Mtime = 291.18 seconds), ostracized participants on average scored around -0.73 standard deviation units lower when compared with included participants, an effect that does not appear to be moderated further by time passed since the ostracism occurrence.

Thus, results show a clear effect of ostracism on both the first and last measures, of which the latter is *not* predicted by time as operationalized in our analysis. The ostracism effect over time can also be inspected via confidence intervals. Comparing the 95% confidence intervals for the average ostracism effect on the first measure (i.e., [-1.54, -1.18]) and on the last measure (i.e., [-0.91, -0.61]) showed no overlap. Although the difference in average effect sizes between first and last measure cannot be formally tested (because of a lack of information on the correlation between measures in the primary studies), the mean difference is sizeable and CIs suggest that the average ostracism effect is smaller for the last measure, as expected. Given the expected positive correlation between interaction effects for first and last measures, the comparison of CIs is likely to be conservative (Schenker & Gentleman, 2001), and hence we conclude that the average ostracism effects are different between first and last measures.

**Moderation of ostracism (Hypothesis 2).** To test moderation of the ostracism effect, we selected the experiments that manipulated ostracism and another independent variable in between-subjects designs.A random-effects model on the interaction effect (Δd) on the first measure (*k* = 52) showed heterogeneity in underlying effects, Q (51) = 103.24, p < .0001, I2 = 50.60% and an estimated τ2 = 0.19, 95% CI [0.07, 0.41]. The average interaction effect equaled Δd = -0.46, p < .0001, 95% CI [-0.64, -0.28], indicating a change in the ostracism effect due to the moderator level and vice versa (i.e., moderation of the ostracism effect). There was indication of funnel plot asymmetry in this analysis, z = -2.43, p = .015. Thus, the data indicate that the ostracism effect *can* be moderated on the first measure following the ostracism sequence, but it is possible that publication bias may have affected the interaction estimates.

On the last measure (*k* = 46), the mixed-effects model (with estimated time as predictor) for the interaction effect again showed residual heterogeneity, QE (44) = 100.82, p < .0001 and estimated τ2 = 0.21, 95% CI [0.10, 0.55]. The intercept of the interaction effect was estimated at Δdintercept­ = -0.20, p = .052, 95% CI [-0.402, 0.002] and no moderation of time was found, b = 0.0002, p = .159, 95% CI [-0.0001, 0.0004]. The regression test with the time and SE as predictors showed no funnel plot asymmetry, z = -0.68, p = .495. These results indicate that moderation of the average ostracism effect is *not* found at a later time-point in the included studies, and time itself does not moderate the computed interaction effects. However, sensitivity analyses showed that this interaction *was* significant when we removed three outliers based on studentized residuals, Δdintercept­ = -0.32, p = .029, 95% CI [-0.60, -0.03], whereas the regression coefficient time continued to be non-significant, b = 0.0002, p = .207, 95% CI [-0.0001, 0.0006]. This indicates that the non-significant interaction effect is sensitive to outliers in the data.

To see whether the interaction effects decrease from the first to the last measure, we again compared confidence intervals. On the first measure, the 95% CI was [-0.64, -0.28] whereas for the last measure, the 95% CI was [-0.402, 0.002]. Considering the overlap of these CIs, there is no indication for an average reduction in the moderation across the measures examined.

In light of these interaction effects, we inspected simple effects of both the ostracism- and moderator factors in the studies. Note that the simple effects for the standard ostracism effect (e.g., ostracism effect when no physical pain manipulation takes place; Riva, Wirth, & Williams, 2011) *represent* the results of the first hypothesis, which indicated that the effect was larger on the first measure (d = -1.36, 95% CI [-1.54, -1.18]) than on the last measure (d = -0.72, 95% CI [-0.86, -0.59]). Estimating the mean simple effect of ostracism within the moderated level (e.g., ostracism effect under a physical pain manipulation; Riva et al., 2011) showed d = -1.34 on the first measure (*k* = 52, p < .0001, 95% CI [-1.69, -0.998]) and d = -0.68 on the last measure (*k* = 46, p < .0001, 95% CI [-0.93, -0.43]). In short, these results indicate that the average effect of ostracism is similar within both levels of the moderator factor, and that the interaction is driven by differences in the simple moderator effects in the primary studies. That is, they are not visible on the basis of comparisons of means across the studies because of Simpson’s paradox (see Borenstein et al., 2009, Chapter 33).

An inspection of the moderator effect *within* the levels of the ostracism factor showed the following. Within the ostracism level, the mean simple moderator effect was estimated at d = 0.31 on the first measure (*k =* 52, p = .001, 95% CI [0.13, 0.48]), and at d = 0.02 on the last measure (*k* = 46, p = .86, 95% CI [-0.17, 0.20]). This indicates that there is an effect of the moderator on the first measure, which can be qualified as the moderated group scoring higher than the non-moderated group (e.g., not mentally visualizing the Cyberball game versus visualizing it; Kassner, Dongning, Law, & Williams, n.d.). For the inclusion level, the simple moderator effect is estimated at d = -0.16 on the first measure (*k* = 52, p = .038, 95% CI [-0.32, -0.09]), and d = -0.17 on the last measure (*k* = 46, p = .013, 95% CI [-0.31, -0.04]). This indicates that the (simple) effect of the moderator is itself moderated by the ostracism effect. Specifically, on the included level, the moderator has a negative mean effect (e.g., self-affirmation present or absent; Webb, Harris, & McAtamney, n.d.) as opposed to a mean positive effect for those who were ostracized. These analyses thus indicate the significant impact the moderator level has on the first measure for the ostracism level, and on both measures for the inclusion level.

**Secondary analyses.** In addition to the simple effects over all studies, we analyzed subsets of studies that differ in terms of measure type to study robustness of the effects. We also inspected whether sample composition, scale composition, and Cyberball specifics could predict the estimated effect size. Finally, we selected a homogeneous subset of studies to come to grips with the relatively large heterogeneity of simple main effects found in the confirmatory analyses.

**Measures.** To inspect the robustness of the estimates, we ran simple effects across several subsets of measures. These subsets encompassed (1) fundamental needs (single- and composite needs), (2) intrapersonal measures (i.e., measures that relate only to the self), (3) interpersonal measures (i.e., measures that relate to others or the self in the context of others) and (4) measures that were coded by the first two authors as fitting the description of being immediate or delayed (i.e., questions related to during- or after the game, respectively). We ran the analyses for the different measures for the two time points separately (i.e., first and last measure).

The different panels in Figure 1 show the results for the different simple effects per subset and overall. A comparison of the the results within each panel shows whether the overall results are robust and representative of all subsets, or whether there are nuances per type of measure. The main differences are notable in panels (1), (2) and (5). The first and second panels indicate that the effect of ostracism within both moderator levels is stronger for the subset of fundamental needs measures, and weaker for interpersonal measures. This indicates that in a similar factorial design, fundamental measures show stronger effects and interpersonal measures weaker effects. Panel 5 indicates that the moderation of interpersonal measures is stronger compared to the other subsets. This suggests that interpersonal measures are more subject to moderation, whereas the effects of ostracism on fundamental needs are larger initially.

**Composition.** We ran a mixed-effects model on the ostracism effect of the first measure (as in Hypothesis 1) for the composition effects, because we expected that the effects would be largest in this setting and the highest number of studies could be included in this analysis. The predictors in the mixed effects model were (1) country (US, other Western country, Asian, other), (2) proportion of males in the study, (3) mean age of the sample, (4) number of players in the game, (5) length of the game (≤ 5min, 5-10 min or > 10 min) and (6) type of needs scale referenced (by assigning unique values for every unique reference). This model (*k =* 47) showed clear residual heterogeneity, QE(36) = 478, p < .0001, estimated τ2 = 0.94, 95% CI [0.57, 1.63], but no overall moderation, QM(10) = 10.78, p = .374. Inspecting the predictors individually also showed no indication for moderation (*p*s > .173; see Table 3), The different types of need scales (e.g., Van Beest & Williams, 2006; Williams, 2009; Zadro et al., 2004) did not significantly moderate effect sizesThese analyses showed large heterogeneity in the effect sizes, necessitating the moderators to have a very large influence on the effect (i.e., half a standard deviation at least) in order to be detectable. We found no indication for such moderation due to study composition.

**Homogeneity?** The analysis of the simple ostracism effect on the first measure showed that differences of underlying effects made up 93% of the variability in study outcomes. We performed an additional exploratory analysis in a more homogenous subset of studies. This subset only included typical Cyberball studies that involved three players in the game, 30 throws, and lasted less than five minutes. In addition, the homogeneous subset of typical Cyberball studies only involved measures of fundamental needs (single or composite). Performing a meta-analysis on this homogeneous subset of studies (*k =* 37) showed an I2 value of 92%, indicating that 92% of the total variability is ascribable to heterogeneity in the effect sizes. We noted that the mean simple ostracism effect in these 37 studies was relatively strong and estimated at d = -2.09, 95% CI [-2.41, -1.77]. In other words, the heterogeneity found in the overall analyses does not appear to be an artefact from the inclusion of different measures and the use of alternative Cyberball setups.

In sum, the secondary analyses indicate several things. First, the effect of ostracism appears relatively large for fundamental needs, and somewhat weaker for the interpersonal measures. The interpersonal measures are also more prone to moderation of the ostracism effect. Second, sample gender distribution appears to predict effect size found, where effect sizes were somewhat larger if there are more males present, although the exact size of the moderation was hard to establish given the residual heterogeneity in this analysis. Third, there is substantial heterogeneity in the effect sizes, even when considering a homogeneous subset of studies in terms of the Cyberball game and measure of the ostracism effects. The latter result suggests that responses to ostracism are more variable than they have been construed until now.

**Discussion**

In this comprehensive meta-analysis of ostracism studies with the Cyberball game, we focused on two confirmatory hypotheses based on theory as well as several exploratory hypotheses. Our results showed that the ostracism effect is quite large on average, but that it varies according to different factors. For the confirmatory hypotheses, two questions were central: (1) “does the effect size of ostracism decrease over time?” and (2) “does the average moderation of ostracism increase over time?” The results showed that there is indication for a decrease in the effect from the first to the last measure, but that this is not moderated by the estimated time between first and last measure. Our analyses also showed that variability of the simple ostracism effect was larger on the first measure (τ2 = 0.90, 95% CI [0.70, 1.24]) when compared to the standard ostracism effect on the last measure (τ2 = 0.38, 95% CI [0.27, 0.54]). We simulated effect sizes under the different estimates to visually show the large difference in variability (see Figure 2). Because reflexes are expected to be fairly homogeneous, this difference in variability suggests that the simple ostracism effect is less of a reflex than has been theorized previously (Williams, 2009). Nevertheless, the change in the effect itself is in accordance Williams’ (2009) theory.

The interaction effect was present on both the first and last measure, but on both occasions the effect was relatively weak. Simple effects indicated that, on average, the ostracism effect operates similarly on both levels of the moderator factor and that both effects are relatively large. At the same time, the moderator factor in the primary studies showed a mean positive effect within the ostracism level and a mean negative effect within the inclusion level. Both theses effects were relatively weak (d = ??) For example, if we apply the findings to the factors of Bernstein, Sacco, and Young (2010), the relative between-subjects effect of ostracism, compared to inclusion, is similar for both being ostracized by the outgroup *and* the ingroup, on the dependent measure need for control. However, those ostracized by the outgroup, outscore those ostracized by the ingroup (and vice versa for the included). Substantively, this means that those in the moderated ostracism group average higher scores on measures such as fundamental needs, when compared to the non-moderated group, an effect which only holds for the first time point. Vice versa, the moderated inclusion group scores lower on measures such as fundamental needs, when compared to non-moderated inclusion, an effect which holds for both the first and last time point. This suggests that a moderator decreases negative feelings in the ostracism conditions, but makes these feelings worse in the inclusion conditions. Substantively, this suggests that what makes the bad feel good, makes the good feel bad.

The exploratory analyses focused on…. These analyses revealed … We also ran exploratory analyses to study whether the make-up of a study predicted study-level effect sizes. The analyses showed that this was the case neither for mean age of the participants, nor the setup of the Cyberball manipulation. Additionally, the gender distribution in the sample did not influence the effect size found, counter to what has been hypothesized before (e.g., Hawes et al., 2012).

Exploratory analyses also showed that the results were robust across subsets of dependent measures and the overall set of dependent measures (see Figure 1), except that interpersonal measures showed relatively weak ostracism effects, while fundamental need measures showed somewhat stronger ostracism effects on the first measure. This suggests that psychological effects of ostracism are large, but that this effect is smaller for interpersonal behaviors. On top of this, interpersonal measures also show more moderation, suggesting that interpersonal behaviors are more easily moderated. In sum, we see that the effects we found are robust, except for three deviations, which only differ in size and not in direction.

In sum, the results of the current meta-analysis extend the theory by Williams (2009) and compliment the findings of previous meta-analyses (Blackhart et al., 2009; Cacioppo et al., 2013; Gerber & Wheeler, 2009), by indicating that the effect of ostracism decreases, that the ostracism effect can be moderated, that the initial response to ostracism is highly variable and we note our findings were robust across a wide set of sensitivity analyses.

Hi Kip, this would be an excellent moment to share your magic….here .. explain that experiments say one thing and that meta does the other. And a call for further experimentation.

**Williams’s Model of Ostracism: Supported or Not?** Regarding the test of Williams’s model, there are a few important observations and limitations. First, Williams proposes fundamental need threat as a result of even a brief episode of ostracism. This was supported by the meta-analysis. The model asserts that negative emotional reactions are also induced by ostracism; this was also supported. The proposition that the initial reaction to ostracism is pain is difficult to assess with the criterion of inclusion in this meta-analysis. The most convincing evidence of ostracism-induced-pain comes from the neuroscience literature, which was not included in this meta-analysis. That moderation is predicted to occur in the reflective stage, when the context and meaning of the ostracism event can be appraised, is also supported by the present meta-analysis. Both individual differences and situational context are shown to moderate downstream reactions. The final stage of Williams’s model—resignation—is outside the aims of the present meta-analysis, as it requires long-term exposure to ostracism. Thus, most of propositions set forth in Williams’s model that can be tested within this meta-analysis are supported.

The one proposition that appears to be unsupported is the argument that, because reactions to ostracism are functionally adaptive, there is a universal distress reaction to the initial onset of ostracism that is resistant to moderation. Williams specifically refers to physiological, online, or immediate retrospective reports to assess reflexive reactions. In many instances, the *first* reaction is not isomorphic with reflexive measures. Anything taken after the game, or assessed by wording indicating present state (rather than the participants’ state during the game), is not assumed to be reflexive, nor predicted to be resistant to moderation. Indeed, empirical studies published by Williams and his colleagues for which reflexive and reflective need satisfaction measures are used show repeatedly that the reflexive measures are unaffected by cross-cutting situational factors or individual differences, whereas reflective measures are moderated. The only exceptions to this pattern are the Wirth, Lynam, & Williams (2010) study for which trait measures of personality disorder do show initial moderation (those with higher levels of trait personality disorder show less pain than those with lower levels of trait personality disorder), and Goodwin, Williams, & Carter-Sowell (2010), for which a small difference in reflexive distress was noted in a sample of over 600 participants such that females showed stronger negative effects than males. That leaves well over 15 publications, some with multiple studies, which show the predicted pattern. Thus, studies that specifically test this proposition have, within the studies, supported the hypothesis of more moderation at the reflective than reflexive stage. However, meta-analytically combining across studies does not provide similar support. Thus, we are left deciding which method is more compelling.

Further, the seemingly large heterogeneity of the data to the first measures is suggested to indicate hidden moderators that have yet been discovered. Whereas this is certainly interesting and may lead to further experimentation and discoveries, we need to be caution when interpreting this heterogeneity in a conceptual vacuum. What paradigms that have been meta-analyzed show significantly less heterogeneity? A control or comparison group is needed to allow for any speculation regarding the presumed high level of heterogeneity.

In sum, there is substantial meta-analytic support for many of the propositions put forth in Williams’s model of ostracism. Future research is needed to clarify whether (or under what conditions) immediate (reflexive) reactions are more resistant to moderation than delayed (reflexive) reactions.

**Limitations.** Within the current meta-analysis there are several limitations. First, our test of differences between the first and last measure was indirect. In its current setting, the meta-analysis makes comparisons between the first and last measures based upon the confidence intervals of these estimates. This is an indirect and informal test of whether the effects differ. A direct test would provide more conclusive evidence on whether or not the effect is equal across the first and last measurements. However, such a direct test requires correlations between the measurements per study, per cell, which are (usually) not reported in papers. This would thus require a direct request for data from each paper, which would possibly yield low response rates (Wicherts, Borsboom, Kats, & Molenaar, 2006), lowering the sample size of the meta-analysis overall.5 This lack of direct testing was thus chosen as a way of retaining sample size within the meta-analysis.

Second, not all measures were included and tested in a repeated-measures meta-analysis. Initially, a pre-test was run including all measures, but this showed that many papers did not include all statistics required for all measures. Requesting all of this information from the authors yielded a limitation that was similar to the first: a trade-off between retaining a sufficiently large set of studies and comprehensiveness. Another reason for only including the first and last measures was that every measure would require two separate meta-analyses to test both the main- and interaction effect (increasing Type I error rates) if a similar analytical model was used. If all measures were included, it would increase the importance of including an effect size correction due to correlations between measures, to facilitate repeated-measures analyses to minimize Type I error rates. In other words, the failure to include all measures was to prevent the problem of multiple testing and nonresponse to data requests, which would lead to a smaller set of useful studies and hence less powerful analyses.

Third, the heterogeneity in the effect sizes poses a problem for the power of finding any moderator effects (Hedges & Pigott, 2004). This could pose the problem that several of the non-effects found are actually there, but not detected (Type II errors). However, the subset of typical Cyberball studies still showed substantial variability in the effect sizes: I2 = 92%. This indicates that the effects are very variable to begin with, and makes it unlikely that the effects are misrepresented.

Additionally, the specific null-effect of time as a predictor could be due to one of three reasons: (1) the heterogeneity in the effect sizes is too large to find moderation by time, (2) imprecise reporting of the measures in the papers led to inaccurate time estimation, or (3) the difference in the effect size is not due to time but differences between the type of measures administered at the different time points. For the imprecise reporting of the measures, authors could be contacted, but this also poses new problems (i.e., nonresponse or authors might not be willing to admit that measures were left out in the paper; LeBel et al., 2013). The difference in measures can be inspected by creating a difference index between the types of measures and regressing that on the effect sizes. Doing this for the standard ostracism effect on the last measure, showed no significant predictive effect of this difference (b = -0.03, p = .531), indicating that the effect is not driven by difference in measures on the first and last time point. In short, there are some limitations of the analyses with time as a moderator, but these limitations are either hard to address (i.e., imprecise reporting or heterogeneity), or the data indicates the opposite (i.e., difference in measures). Inspecting whether the types of measures used across all studies are different, and not the difference within a study, shows that these are similarly distributed across timepoints (maximum discrepancy of 4.9 percentage points). Substantive differences in proportions of measures across time points are minimal, and form an unlikely driving force for our findings. In sum, we conclude that the findings are not an artefact of selecting the first and last measures.

Fourth, the current meta-analysis only examined between-subjects designs. Possibly there is a difference for the ostracism effect in between- and within-subjects designs, something that we have not directly investigated. Also, the within-subjects designs often used fMRI data or other physiological data such as EEG (27 out of 49 at least), which pose an interesting avenue for further research in a meta-analytic domain of neurophysiological measures to add to the work of Cacioppo and colleagues (2013) within the physiological framework. These references can easily be retrieved from the database of examined papers, as is available on the OSF page of this paper.

A final note is that this paper only summarized the results of the measures included in the studies. However obvious this might be, it should be pointed out, because the validity of the conclusions are reliant on the validity of the measures. Most prominently represented in the current meta-analysis are the fundamental need measures, which have no proper psychometric validation up-to-date, notwithstanding their wide use.6 Other kinds of included measures possibly have the same, and one has been openly criticized (e.g., the Hot Sauce aggression paradigm; Ritter & Eslea, 2005). Considering this lack of validation or even questionable validity, we note that results in this paper are conditional on that these measures *are* valid.

Mff We coded the articles in a specific way. In each article we only condired the first and last measures that was subsequently coded as being an interpersonal, intrapersonal or need satisfaction meausure. An alternative method for coding would be to check in each article whether it contains an interpsonal, intrapersonal or needs satisfaction measures, and then code for each of these measures whehther it is first or last. In our coding we only have two means for every article. In the second coding it would be possible to have 6 means for every article. The consequence is that we can state whether T1 differs from T2, and see how different measures change this relations. But, can not directly test how a specific measure is affected by temperoal changes.

**Conclusion.**. Our results show that the ostracism effect is large, with an effect size approaching 1.5 standard deviation units. Moreover, the results show that the impact of ostracism decreases from first to last measure. An interesting avenue for future research concerns the explanation of why ostracism can be moderated equally across the first and last measure within the studies. In addition, the large heterogeneity in the effect sizes (even in a homogeneous subset of studies) highlight that there are more potentially relevant moderators of ostracism in need of further study and raises the question whether the immediate ostracism effect is actually as reflexive as has been previously proposed, These findings extend the need-threat model (Williams, 2009), which has played a major role in ostracism research until now. We invite fellow researchers to think and test ideas that might provide some explanation for these findings.

References

References marked with an asterisk indicate studies included in the meta-analysis.

Bakker, M., van Dijk, A., & Wicherts, J. M. (2012). The rules of the game called psychological science. *Perspectives on Psychological Science, 7*, 543-554. doi: 10.1177/1745691612459060

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*, 497-529.

Baumeister, R. F., Twenge, J. M., & Nuss, C. K. (2002). Effects of social exclusion on cognitive processes: Anticipated aloneness reduces intelligent thought. *Journal of Personality and Social Psychology*, *83*(4), 817–827. doi:10.1037//0022-3514.83.4.817

Blackhart, G. C., Nelson, B. C., Knowles, M. L., & Baumeister, R. F. (2009). Rejection elicits emotional reactions but neither causes immediate distress nor lowers self-esteem: a meta-analytic review of 192 studies on social exclusion. *Personality and social psychology review : an official journal of the Society for Personality and Social Psychology, Inc*, *13*(4), 269–309. doi:10.1177/1088868309346065

Borenstein, M. (2009). Effect sizes for continuous data. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed.). New York, NY: Russell Sage Foundation.

Boyes, M. E., & French, D. J. (2009). Having a Cyberball: Using a ball-throwing game as an experimental social stressor to examine the relationship between neuroticism and coping. *Personality and Individual Differences*, *47*(5), 396–401. doi:10.1016/j.paid.2009.04.005

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629–634. doi:10.1136/bmj.315.7109.629

Gerber, J., & Wheeler, L. (2009). On Being Rejected: A Meta-Analysis of Experimental Research on Rejection. *Perspectives on Psychological Science*, *4*(5), 468–488. doi:10.1111/j.1745-6924.2009.01158.x

Gonsalkorale, K., & Williams, K. D. (2007). The KKK won’t let me play: ostracism even by a despised outgroup hurts. *European Journal of Social Psychology*, *37*(6), 1176–1186. doi:10.1002/ejsp.392

Hawes, D. J., Zadro, L., Fink, E., Richardson, R., O’Moore, K., Griffiths, B., … Williams, K. D. (2012). The effects of peer ostracism on children’s cognitive processes. *European Journal of Developmental Psychology*, *9*(5), 599–613. doi:10.1080/17405629.2011.638815

Hunter, J. E., & Schmidt, F. L. (1990). Dichotomization of Continuous Variables: The Implications for Meta-Anafysis. *Journal of Applied Psychology*, *75*(3), 334–349.

Larry V. Hedges. (1981). Distribution theory for Glass’s estimator of effect size and related estimators, *6*(2), 107–128.

Leary, M. R., Kowalski, R. M., Smith, L., & Phillips, S. (2003). Teasing, rejection, and violence: Case studies of the school shootings. *Aggressive Behavior*, *29*(3), 202–214. doi:10.1002/ab.10061

Lustenberger, D. E., & Jagacinski, C. M. (2010). Exploring the Effects of Ostracism on Performance and Intrinsic Motivation. *Human Performance*, *23*(4), 283–304. doi:10.1080/08959285.2010.501046

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, *7*(1), 19–40. doi:10.1037//1082-989X.7.1.19

Oaten, M., Williams, K. D., Jones, A., & Zadro, L. (2008). The effects of ostracism on self-regulation in the socially anxious. *Journal of Social and Clinical Psychology*, *27*(5), 471–504.

Pharo, H., Gross, J., Richardson, R., & Hayne, H. (2011). Age-related changes in the effect of ostracism. *Social Influence*, *6*(1), 22–38. doi:10.1080/15534510.2010.525852

R Core Team. (2013). R: A Language and Environment for Statistical Computing. Vienna, Austria. Retrieved from http://www.r-project.org/

Ritter, D., & Eslea, M. (2005). Hot Sauce, toy guns, and graffiti: A critical account of current laboratory aggression paradigms. *Aggressive Behavior*, *31*(5), 407–419. doi:10.1002/ab.20066

Sterne, J. A. C., & Egger, M. (2005). Regression Methods to Detect Publication and Other Bias in Meta-Analysis. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication bias in meta-analysis*. Chichester: John Wiley & Sons.

Van Beest, I., Carter-Sowell, A. R., van Dijk, E., & Williams, K. D. (2012). Groups being ostracized by groups: Is the pain shared, is recovery quicker, and are groups more likely to be aggressive? *Group Dynamics: Theory, Research, and Practice*, *16*(4), 241–254. doi:10.1037/a0030104

Van Beest, I., & Williams, K. D. (2006). When inclusion costs and ostracism pays, ostracism still hurts. *Journal of personality and social psychology*, *91*(5), 918–928. doi:10.1037/0022-3514.91.5.918

Viechtbauer, W. (2005). Bias and Efficiency of Meta-Analytic Variance Estimators in the Random-Effects Model. *Journal of Educational and Behavioral Statistics*, *30*(3), 261–293. doi:10.3102/10769986030003261

Viechtbauer, Wolfgang. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/wics.10/full

Webb, T. L., Harris, P., & McAtamney, K. (*n.d.*). Self-Affirmation Moderates the Negative Effect of Ostracism on Perceptions of Control, But Not in Predictable Ways. Unpublished manuscript, University of Sheffield, Sheffield, UK.

Wicherts, J. M., Borsboom, D., Kats, J., & Molenaar, D. (2006). The poor availability of psychological research data for reanalysis. *The American psychologist*, *61*(7), 726–8. doi:10.1037/0003-066X.61.7.726

Williams, K D, Cheung, C. K., & Choi, W. (2000). Cyberostracism: effects of being ignored over the Internet. *Journal of personality and social psychology*, *79*(5), 748–62. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11079239

Williams, K.D. (2009). Ostracism: a temporal need-threat model. *Advances in Experimental Social Psychology*, *41*, 275–314. doi:10.1016/S0065-2601(08)00406-1

Williams, Kipling D., & Jarvis, B. (2006). Cyberball: A program for use in research on interpersonal ostracism and acceptance. *Behavior Research Methods*, *38*(1), 174–180. doi:10.3758/BF03192765

Zadro, L., Williams, K. D., & Richardson, R. (2004). How low can you go? Ostracism by a computer is sufficient to lower self-reported levels of belonging, control, self-esteem, and meaningful existence. *Journal of Experimental Social Psychology*, *40*(4), 560–567. doi:10.1016/j.jesp.2003.11.006

**Footnotes**

1. Note that exclusion encompasses both social rejection and social ostracism.
2. The direct link: <http://openscienceframework.org/project/HT25n/>
3. It has been updated since, but the list that was used can be found on the Open Science Framework page.
4. Oaten, Williams, Jones and Zadro (2008) was applicable, but was excluded due to being an outlier with respect to effect size (ds > 15). This in similar vein as in Gerber and Wheeler (2009; p. 473): “*One study (Oaten, Williams, Jones, & Zadro, 2007) had need effect sizes that were clear outliers (effect sizes were 5–7 standard deviations above the means)* […and…] *were excluded from the analyses.*”
5. Note that out of the 72 data requests, we received timely replies of 52 (i.e., ~72%). However, these requests were only for specific information and not for raw datasets, as was the case in Wicherts et al. (2006).
6. Some might say that *because* they are widely used, they do not require psychometric validation, that is, they would have been selected out of research if they were improper. This however is no argument for a lack of psychometric validation, because there are no arguments to oppose psychometric validation of any scale, that is there are no arguments for using unvalidated scales over validated scales.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 1 - Hypothetical data example of coding correction | | | | |  |  |  |  |  |  |  |
| (a) Negative moderator, negative measure | | | |  |  | (b) Positive moderator, negative measure | | | |  |  |
|  |  | Moderated | Not-moderated/control | Raw | Correct |  |  | Moderated | Not-moderated/control | Raw | Correct |
| Ostracism factor | Ostracism | 13 | 11 | 2 | 2 | Ostracism factor | Ostracism | 9 | 11 | -2 | 2 |
|  | Inclusion | 8 | 8 | 0 | 0 |  | Inclusion | 8 | 8 | 0 | 0 |
|  | Raw | 5 | 3 |  |  |  | Raw | 1 | 3 |  |  |
|  | Correct | -5 | -3 |  |  |  | Correct | -1 | -3 |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
| (c) Negative moderator, positive measure | | | |  |  | (d) Positive moderator, positive measure | | | |  |  |
|  |  | Moderated | Not-moderated/control | Raw | Correct |  |  | Moderated | Not-moderated/control | Raw | Correct |
| Ostracism factor | Ostracism | 3 | 5 | -2 | 2 | Ostracism factor | Ostracism | 7 | 5 | 2 | 2 |
|  | Inclusion | 8 | 8 | 0 | 0 |  | Inclusion | 8 | 8 | 0 | 0 |
|  | Raw | -5 | -3 |  |  |  | Raw | -1 | -3 |  |  |
|  | Correct | -5 | -3 |  |  |  | Correct | -1 | -3 |  |  |
| Note: raw denotes the simple effect in the hypothetical data before correction whereas correct denotes the simple effect after correction. Column wise effects are multiplied by the type of measure only, whereas column wise effects are multiplied by both the type of moderator and type of measure. | | | | | | | | | | | |

| Table 2 – Effect sizes per study for the confirmatory hypotheses | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| First author | Year | N | d T1 | (SE) | d T2 | (SE) | Δd T1 | (SE) | Δd T2 | (SE) |
| Alvares | 2010 | 74 | -1.21 | 0.12 | -0.10 | 0.10 | -0.15 | 0.24 | 1.12 | 0.23 |
| Ambrosini | 2013 | 40 | -1.69 | 0.13 | -0.97 | 0.11 | - | - | - | - |
| Anonymous | - | 49 | -3.97 | 0.24 | -1.32 | 0.10 | - | - | - | - |
| Anonymous | - | 91 | -3.17 | 0.20 | -0.48 | 0.09 | 0.75 | 0.56 | 0.53 | 0.18 |
| Aydin | 2012 | 68 | -0.95 | 0.13 | -0.40 | 0.12 | -1.19 | 0.24 | 0.72 | 0.23 |
| Banki | 2012 | 89 | -1.87 | 0.07 | -0.35 | 0.05 | - | - | - | - |
| Bastian | 2010 | 72 | -2.75 | 0.11 | -1.42 | 0.07 | - | - | - | - |
| Bernstein | 2012 | 24 | -0.41 | 0.16 | - | - | - | - | - | - |
| Bernstein | 2012 | 25.50 | -1.04 | 0.17 | - | - | - | - | - | - |
| Bernstein | 2010 | 73 | -1.63 | 0.16 | -1.63 | 0.16 | -0.86 | 0.37 | -1.11 | 0.40 |
| Bernstein | 2010 | 138 | -2.67 | 0.10 | -1.96 | 0.08 | -0.53 | 0.22 | -0.51 | 0.17 |
| Bernstein | 2012 | 67 | -2.00 | 0.17 | -0.99 | 0.13 | -1.07 | 0.45 | -0.80 | 0.30 |
| Bernstein | 2012 | 27 | -1.39 | 0.17 | - | - | - | - | - | - |
| Boyes | 2009 | 89 | -0.43 | 0.05 | -0.80 | 0.05 | - | - | - | - |
| Boyes | 2009 | 87 | -0.20 | 0.05 | -0.84 | 0.05 | - | - | - | - |
| Brochu | - | 35 | -2.51 | 0.20 | -0.48 | 0.11 | - | - | - | - |
| Brown | 2009 | 52 | -0.64 | 0.08 | - | - | - | - | - | - |
| Carter | 2008 | 143 | -0.28 | 0.06 | 0.20 | 0.06 | 0.34 | 0.11 | 0.17 | 0.11 |
| Carter-Sowell | 2008 | 65 | -2.86 | 0.12 | -1.48 | 0.08 | - | - | - | - |
| Carter-Sowell | 2010 | 74 | -1.60 | 0.14 | -1.49 | 0.13 | -1.23 | 0.33 | -1.15 | 0.34 |
| Carter-Sowell | 2010 | 70.67 | -2.09 | 0.17 | -0.56 | 0.11 | -0.65 | 0.39 | -0.63 | 0.24 |
| Chen | 2012 | 60 | -1.04 | 0.14 | - | - | -1.35 | 0.27 | - | - |
| Chen | 2012 | 83 | -1.32 | 0.11 | - | - | -1.32 | 0.21 | - | - |
| Chernyak | 2010 | 76 | -1.52 | 0.10 | 0.15 | 0.08 | - | - | - | - |
| Chow | 2008 | 75 | -1.20 | 0.06 | -1.31 | 0.06 | - | - | - | - |
| Chrisp | 2012 | 77 | -0.70 | 0.06 | -0.15 | 0.05 | - | - | - | - |
| Coyne | 2011 | 40 | -0.56 | 0.10 | - | - | - | - | - | - |
| De Waal-Andrews | 2012 | 136 | -3.55 | 0.16 | -2.55 | 0.11 | -1.29 | 0.24 | -0.87 | 0.18 |
| De Waal-Andrews | 2012 | 112 | -4.21 | 0.22 | -2.17 | 0.11 | -1.56 | 0.31 | -1.20 | 0.18 |
| DeBono | - | 57 | -1.07 | 0.15 | -0.05 | 0.13 | -1.55 | 0.29 | -0.48 | 0.27 |
| DeBono | - | 81 | -1.07 | 0.11 | -0.10 | 0.09 | -0.33 | 0.21 | 0.24 | 0.19 |
| DeBono | - | 83 | -0.13 | 0.09 | - | - | -0.75 | 0.19 | - | - |
| Dietrich | 2010 | 75 | 1.43 | 0.07 | - | - | - | - | - | - |
| Duclos | 2012 | 59 | -0.63 | 0.07 | - | - | - | - | - | - |
| Eisenberger | 2006 | 48 | -0.15 | 0.08 | -1.24 | 0.10 | - | - | - | - |
| Fayant | - | 60 | -2.04 | 0.20 | -1.12 | 0.15 | 0.22 | 0.38 | -0.44 | 0.28 |
| Floor | 2007 | 88 | -1.92 | 0.13 | -0.73 | 0.09 | -0.21 | 0.28 | -0.59 | 0.19 |
| Gallardo-Pujol | 2012 | 57 | -1.18 | 0.16 | -0.52 | 0.15 | -1.17 | 0.31 | 0.11 | 0.29 |
| Gan | 2012 | 72 | -0.54 | 0.03 | -0.07 | 0.03 | -0.62 | 0.06 | 0.02 | 0.06 |
| Garczynski | 2013 | 83 | -1.51 | 0.19 | 0.39 | 0.15 | -1.29 | 0.33 | -0.01 | 0.29 |
| Geniole | 2011 | 74 | 0.19 | 0.06 | -0.11 | 0.06 | - | - | - | - |
| Gerber | - | 38 | -2.09 | 0.16 | - | - | - | - | - | - |
| Gerber | - | 89 | -3.38 | 0.21 | - | - | - | - | - | - |
| Gonsalkorale | 2007 | 97 | -1.31 | 0.14 | 0.26 | 0.12 | 0.49 | 0.30 | 1.31 | 0.25 |
| Goodwin | 2010 | 300 | -1.81 | 0.04 | -0.94 | 0.03 | 0.20 | 0.08 | -0.43 | 0.07 |
| Goodwin | 2010 | 314 | 0.13 | 0.02 | -0.09 | 0.02 | 0.35 | 0.06 | -0.10 | 0.06 |
| Greitemeyer | 2012 | 56 | -0.48 | 0.07 | -0.23 | 0.07 | - | - | - | - |
| Gruijters | - | 113 | -0.26 | 0.06 | -1.07 | 0.07 | - | - | - | - |
| Hackenbracht | 2013 | 51 | -1.92 | 0.11 | -0.18 | 0.08 | - | - | - | - |
| Hawes | 2012 | 55 | -2.16 | 0.23 | 0.69 | 0.15 | 0.00 | 0.38 | -1.05 | 0.28 |
| Hellmann | - | 76 | -1.21 | 0.12 | 0.19 | 0.10 | -1.40 | 0.22 | 0.74 | 0.21 |
| Hess | 2010 | 162 | -2.34 | 0.04 | -0.87 | 0.03 | - | - | - | - |
| Hess | 2011 | 38 | -0.64 | 0.11 | - | - | - | - | - | - |
| Horn | - | 68 | -0.77 | 0.12 | -0.99 | 0.13 | -0.99 | 0.23 | 1.49 | 0.24 |
| Ijzerman | 2012 | 86 | -1.67 | 0.12 | - | - | -1.07 | 0.22 | - | - |
| Jamieson | 2010 | 33 | -1.56 | 0.15 | -1.06 | 0.13 | - | - | - | - |
| Jamieson | 2010 | 68 | -1.94 | 0.09 | -1.47 | 0.07 | - | - | - | - |
| Johnson | 2010 | 104 | -0.73 | 0.04 | -0.79 | 0.04 | - | - | - | - |
| Kassner | - | 85 | -1.72 | 0.13 | -1.02 | 0.11 | -0.87 | 0.31 | -0.30 | 0.21 |
| Kassner | 2012 | 49 | -2.11 | 0.12 | -1.78 | 0.11 | - | - | - | - |
| Kerr | 2008 | 250 | -1.66 | 0.02 | -0.05 | 0.02 | - | - | - | - |
| Kesting | 2013 | 76 | -0.28 | 0.05 | -0.79 | 0.06 | - | - | - | - |
| Knowles | 2010 | 62 | -0.38 | 0.12 | - | - | -0.99 | 0.25 | - | - |
| Knowles | 2012 | 60 | -0.60 | 0.07 | - | - | - | - | - | - |
| Krijnen | 2008 | 144 | -4.74 | 0.11 | -0.18 | 0.03 | - | - | - | - |
| Krill | 2008 | 119 | -2.11 | 0.05 | -0.57 | 0.03 | - | - | - | - |
| Lakin | 2008 | 36 | -1.53 | 0.14 | -0.51 | 0.11 | - | - | - | - |
| Lau | 2009 | 56 | -2.50 | 0.23 | -1.09 | 0.15 | -0.06 | 0.58 | 1.36 | 0.46 |
| Lustenberger | 2010 | 71 | -0.83 | 0.06 | 0.04 | 0.06 | - | - | - | - |
| Lustenberger | 2010 | 156 | -0.70 | 0.03 | - | - | - | - | - | - |
| MacDonald | 2008 | 63 | -0.15 | 0.06 | - | - | - | - | - | - |
| McDonald | 2012 | 270 | -0.06 | 0.02 | -2.40 | 0.03 | - | - | - | - |
| Nordgren | 2011 | 71 | -0.74 | 0.06 | - | - | - | - | - | - |
| Nordgren | 2011 | 74 | -0.80 | 0.06 | - | - | - | - | - | - |
| Nordgren | 2011 | 46 | -2.24 | 0.14 | - | - | - | - | - | - |
| Nordgren | 2011 | 44.67 | -0.55 | 0.09 | -0.75 | 0.09 | - | - | - | - |
| Nordgren | 2011 | 58.67 | -0.65 | 0.07 | - | - | - | - | - | - |
| Oberleitner | 2012 | 88 | -2.36 | 0.08 | 0.42 | 0.05 | - | - | - | - |
| O’Brien | 2012 | 125 | -0.58 | 0.03 | -0.69 | 0.03 | - | - | - | - |
| Peterson | 2011 | 40 | -0.89 | 0.11 | -0.91 | 0.11 | - | - | - | - |
| Pharo | 2011 | 74 | -1.33 | 0.13 | -0.58 | 0.11 | -1.01 | 0.30 | -0.84 | 0.23 |
| Plaisier | 2012 | 149 | -0.36 | 0.05 | 0.23 | 0.05 | -0.40 | 0.11 | -0.56 | 0.11 |
| Ramirez | 2009 | 121 | -2.26 | 0.05 | -1.02 | 0.04 | - | - | - | - |
| Ren | 2012 | 53 | -2.18 | 0.12 | -0.17 | 0.07 | - | - | - | - |
| Renneberg | 2011 | 60 | -1.46 | 0.16 | -1.30 | 0.15 | 0.47 | 0.29 | 0.51 | 0.29 |
| Riva | 2011 | 100 | -2.10 | 0.13 | -1.09 | 0.09 | - | - | - | - |
| Ruggieri | - | 91 | -0.39 | 0.04 | -0.57 | 0.05 | - | - | - | - |
| Ruggieri | - | 74 | -0.06 | 0.13 | -0.23 | 0.13 | -0.31 | 0.24 | -0.68 | 0.23 |
| Sacco | 2011 | 51 | -2.40 | 0.13 | -1.45 | 0.10 | - | - | - | - |
| Sacco | 2011 | 21 | -2.28 | 0.29 | -1.46 | 0.22 | - | - | - | - |
| Sacco | 2011 | 38 | -1.74 | 0.14 | -1.04 | 0.11 | - | - | - | - |
| Salvy | 2010 | 59 | -1.45 | 0.08 | -1.43 | 0.08 | - | - | - | - |
| Salvy | 2009 | 103 | -1.48 | 0.05 | -1.31 | 0.05 | - | - | - | - |
| Schaafsma | 2012 | 720 | -1.42 | 0.02 | -0.49 | 0.02 | 0.09 | 0.03 | 0.33 | 0.03 |
| Segovia | 2012 | 56 | 0.14 | 0.13 | - | - | -1.89 | 0.32 | - | - |
| Staebler | 2011 | 68 | -0.79 | 0.12 | -0.05 | 0.12 | 0.50 | 0.23 | 0.42 | 0.23 |
| Stillman | 2009 | 121 | -0.74 | 0.15 | -1.13 | 0.16 | 0.57 | 0.22 | -1.19 | 0.24 |
| Stock | 2011 | 155 | -2.00 | 0.04 | -0.13 | 0.03 | - | - | - | - |
| Van Beest | 2011 | 87 | -0.94 | 0.10 | -0.58 | 0.09 | -0.40 | 0.24 | -0.44 | 0.19 |
| Van Beest | 2011 | 183 | -2.64 | 0.13 | -0.50 | 0.07 | -0.76 | 0.22 | -0.11 | 0.13 |
| Van Beest | 2006 | 135 | -1.29 | 0.07 | -0.65 | 0.06 | -0.10 | 0.14 | -0.13 | 0.12 |
| Van Beest | 2006 | 111.33 | -2.11 | 0.11 | 0.09 | 0.07 | -0.09 | 0.22 | -0.19 | 0.14 |
| Van Beest | 2012 | 125 | -2.68 | 0.11 | -1.24 | 0.07 | 0.06 | 0.35 | -0.23 | 0.15 |
| Van Beest | 2012 | 85 | -3.10 | 0.20 | 0.05 | 0.09 | -0.28 | 0.44 | 0.07 | 0.18 |
| Van Dijk | - | 51 | -1.50 | 0.10 | -0.04 | 0.08 | - | - | - | - |
| Webb | - | 170 | -0.91 | 0.05 | -0.38 | 0.05 | 0.03 | 0.10 | 0.04 | 0.09 |
| Weik | 2010 | 65 | 0.16 | 0.12 | -0.22 | 0.12 | -0.43 | 0.24 | 0.66 | 0.24 |
| Wesselmann | 2009 | 82 | -0.71 | 0.10 | -2.03 | 0.14 | -1.30 | 0.24 | -0.20 | 0.28 |
| Wesselmann | 2012 | 91 | -1.46 | 0.06 | - | - | - | - | - | - |
| Williams | 2002 | 390 | -0.39 | 0.01 | -2.35 | 0.02 | - | - | - | - |
| Williams | 2000 | 732 | -0.79 | 0.01 | -1.44 | 0.01 | - | - | - | - |
| Williams | 2000 | 111 | -0.26 | 0.06 | -1.01 | 0.07 | -0.20 | 0.15 | -0.98 | 0.15 |
| Wirth | 2009 | 159.33 | -2.29 | 0.08 | -0.76 | 0.05 | 0.05 | 0.17 | 0.46 | 0.11 |
| Wirth | 2010 | 76 | -0.96 | 0.06 | -1.64 | 0.07 | - | - | - | - |
| Zadro | 2004 | 62 | -1.63 | 0.16 | -0.19 | 0.12 | -0.11 | 0.32 | -1.12 | 0.28 |
| Zadro | 2004 | 77 | -1.75 | 0.14 | -0.33 | 0.10 | -0.29 | 0.28 | -0.70 | 0.21 |
| Zadro | 2006 | 56 | -3.70 | 0.19 | -0.87 | 0.08 | - | - | - | - |
| Zhong | 2008 | 52 | -0.72 | 0.15 | - | - | - | - | - | - |
| Zoller | 2010 | 57 | -0.24 | 0.07 | -0.09 | 0.07 | - | - | - | - |
| Zwolinski | 2012 | 56 | -2.01 | 0.11 | -0.28 | 0.07 | - | - | - | - |
| Note: d T1 refers to ostracism effect on first measure; d T2 refers to ostracism effect on last measure; Δd represent interactions. | | | | | | | | | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Table 3—Meta-regression coefficients for composition effects | | | | | |  |
|  | Estimate | (SE) | Z-value | p-value | 95% CI Lowerbound | 95% CI Upperbound |
| Intercept | -0.33 | 3.24 | -0.10 | 0.919 | -6.68 | 6.02 |
| Country = US | - | - | - | - | - | - |
| Country = Western | -0.45 | 0.37 | -1.22 | 0.222 | -1.18 | 0.27 |
| Country = Asian | -0.65 | 1.13 | -0.58 | 0.564 | -2.87 | 1.56 |
| Proportion male | 1.50 | 1.10 | 1.36 | 0.173 | -0.66 | 3.65 |
| Mean age | -0.02 | 0.05 | -0.48 | 0.629 | -0.13 | 0.08 |
| Nr. of players | -0.45 | 1.07 | -0.42 | 0.673 | -2.54 | 1.64 |
| Ostracism <5 min | - | - | - | - | - | - |
| Ostracism 5-10 min | 0.86 | 0.83 | 1.04 | 0.297 | -0.76 | 2.48 |
| Need scale = Williams (2000) | - | - | - | - | - | - |
| Need scale = Zadro et al. (2004) | -0.31 | 0.42 | -0.74 | 0.456 | -1.13 | 0.51 |
| Need scale = Van Beest & Williams (2006) | -0.28 | 0.52 | -0.54 | 0.586 | -1.29 | 0.73 |
| Need scale = Williams Zadro | 0.04 | 0.63 | 0.06 | 0.950 | -1.20 | 1.28 |
| Need scale = Gonsalkorale & Williams (2007) | 0.87 | 0.83 | 1.04 | 0.299 | -0.77 | 2.50 |
| Note: this can be interpreted as a standard regression formula. Empty rows represent reference categories | | | | | | |

C:\Users\Chris\Downloads\Rplot.tiff

*Figure 1.* A dotplot of the average estimated simple effects with 95% confidence intervals, where T1 represents first measure, and T2 represents last measure. Traditional ostracism effect refers to the between-subjects effect of being ostracized with *no* moderator present, whereas moderated ostracism effect refers to being ostracized *with* a moderator present. Vice versa, moderator effect within ostracism/inclusion level refers to the between-subjects effect of the moderator factor, within the ostracized/inclusion conditions.

C:\Users\Chris\Dropbox\MSc ReMa\Cyberball Meta-analysis\5.Writing\simEffects.tiff

*Figure 2.* Simulated effects under the model estimates for the standard ostracism effect, showing there is higher estimated heterogeneity on the first measure.

Appendix

All formulae reported below originate from the chapter by Michael Borenstein (2009). Hedges’ g was calculated as

where d is the standardized main effect. For the standardized interaction effect d was calculated as

where the first term in the nominator is the ostracism effect and the second term is the ostracism effect in the moderator conditions. This Δd corresponds to the partial eta-squared of the interaction. Sampling variance of g was calculated by multiplying the sampling variance of d by the squared correction factor, that is

where the sampling variance of the interaction was calculated as the sum of the sampling variances of both the simple main effects.

**Supplementary materials**

Reference list of included studies

Alvares, G. A., Hickie, I. B., & Guastella, A. J. (2010). Acute effects of intranasal oxytocin on subjective and behavioral responses to social rejection. *Experimental and Clinical Psychopharmacology*, *18*, 316–321.doi:10.1037/a0019719

Ambrosini, E., Blomberg, O., Mandrigin, A., & Costantini, M. (2013). Social exclusion modulates pre-reflective interpersonal body representation. *Psychological Research*. doi:10.1007/s00426-012-0476-2

Anonymous. (*in preparation*). Partial Ostracism, Partial Pain, Partial Revenge.

Aydin, N., Krueger, J. I., Fischer, J., Hahn, D., Kastenmüller, A., Frey, D., & Fischer, P. (2012). “Man’s best friend:” How the presence of a dog reduces mental distress after social exclusion. *Journal of Experimental Social Psychology*, *48*, 446–449. doi:10.1016/j.jesp.2011.09.011

Banki, S. (2012). *How much or how many? Partial ostracism and its consequences*. University of Toronto. Retrieved from https://tspace.library.utoronto.ca/bitstream/1807/32659/3/Banki\_Sara\_201206\_PhD\_thesis.pdf

Bastian, B., & Haslam, N. (2010). Excluded from humanity: The dehumanizing effects of social ostracism. *Journal of Experimental Social Psychology*, *46*, 107–113. doi:10.1016/j.jesp.2009.06.022

Bernstein, M. J., & Claypool, H. M. (2012a). Social exclusion and pain sensitivity: why exclusion sometimes hurts and sometimes numbs. *Personality & Social Psychology Bulletin*, *38*, 185–196. doi:10.1177/0146167211422449

Bernstein, M. J., & Claypool, H. M. (2012b). Not all social exclusions are created equal: Emotional distress following social exclusion is moderated by exclusion paradigm. *Social Influence*, *7*, 113–130. doi:10.1080/15534510.2012.664326

Bernstein, M. J., Sacco, D. F., Young, S. G., Hugenberg, K., & Cook, E. (2010a). Being “in” with the in-crowd: the effects of social exclusion and inclusion are enhanced by the perceived essentialism of ingroups and outgroups. *Personality & Social Psychology Bulletin*, *36*, 999–1009. doi:10.1177/0146167210376059

Bernstein, M. J., Sacco, D. F., Young, S. G., Hugenberg, K., & Cook, E. (2010b). Unpublished study of “Being ‘in’ with the crowd” paper.

Boyes, M. E., & French, D. J. (2009). Having a Cyberball: Using a ball-throwing game as an experimental social stressor to examine the relationship between neuroticism and coping. *Personality and Individual Differences*, *47*, 396–401. doi:10.1016/j.paid.2009.04.005

Brochu, P. M., Garcia, D. M., Smith, H. J., & Esses, V. M. (*working paper*). Ironic Effects of Ambiguous Weight Stigmatization on the Regulation of Eating Behavior.

Brown, C. M., Young, S. G., Sacco, D. F., Bernstein, M. J., & Claypool, H. M. (2009). Social Inclusion Facilitates Interest in Mating. *Evolutionary Psychology*, *7*(1), 11–27.

Carter, B. E. (2008). *When does ostracism decrease self-regulation?* Montana State University. Retrieved from http://etd.lib.montana.edu/etd/2008/carter/CarterB0508.pdf

Carter-Sowell, A R, Wesslemann, E. C., Wirth, J. H., Law, A. T., Chen, Z., Kosasih, M. W., Van der Lee, R., et al. (2010). Strides for belonging trump strides for superiority: Effects of being ostracized for being superior or inferior to the others. *The Journal of Individual Psychology*, *66*(1), 68–92.

Carter-Sowell, Adrienne R, Chen, Z., & Williams, K. D. (2008). Ostracism increases social susceptibility. *Social Influence*, *3*, 143–153. doi:10.1080/15534510802204868

Chen, Z., DeWall, C. N., Poon, K.-T., & Chen, E.-W. (2012). When destiny hurts: Implicit theories of relationships moderate aggressive responses to ostracism. *Journal of Experimental Social Psychology*, *48*, 1029–1036. doi:10.1016/j.jesp.2012.04.002

Chernyak, N., & Zayas, V. (2010). Being excluded by one means being excluded by all: Perceiving exclusion from inclusive others during one-person social exclusion. *Journal of Experimental Social Psychology*, *46*, 582–585. doi:10.1016/j.jesp.2010.01.004

Chow, R. M., Tiedens, L. Z., & Govan, C. L. (2008). Excluded emotions: The role of anger in antisocial responses to ostracism. *Journal of Experimental Social Psychology*, *44*, 896–903. doi:10.1016/j.jesp.2007.09.004

Chrisp, J. A. L. (2012). *Exploring the role of belonging in intergroup discrimination.* University of Otago. Retrieved from http://otago.ourarchive.ac.nz/bitstream/handle/10523/2662/ChrispJulia

Coyne, S. M., Gundersen, N., Nelson, D. A., & Stockdale, L. (2011). Adolescents’ Prosocial Responses to Ostracism: An Experimental Study. *The Journal of Social Psychology*, *151*, 657–661. doi:10.1080/00224545.2010.522625

De Waal-Andrews, W., & Van Beest, I. (2012). When you don’t quite get what you want: psychological and interpersonal consequences of claiming inclusion. *Personality & Social Psychology Bulletin*, *38*, 1367–1377. doi:10.1177/0146167212450463

DeBono, A. (*in preparation*). Understanding Maladaptive Responses to Social Exclusion: Aggression with an Audience.

DeBono, A., & Muraven, M. (*in preparation*). Personality traits that reduce and exacerbate aggression from social exclusion when depleted.

Dietrich, D., Hawkinson, K., & Palo, A. (2010).Self-monitoring as a moderating factor of aggressive responses to ostracism. *SPSP Conference abstract*.

Duclos, R., Wan, E. W., & Jiang, Y. (2013). Show Me the Honey! Effects of Social Exclusion on Financial Risk-Taking. *Journal of Consumer Research*, *40*, 122–135. doi:10.1086/668900

Eisenberger, N. I., Jarcho, J. M., Lieberman, M. D., & Naliboff, B. D. (2006). An experimental study of shared sensitivity to physical pain and social rejection. *Pain*, *126*, 132–138. doi:10.1016/j.pain.2006.06.024

Fayant, M. P., Lantian, A., Muller, D., & Hartgerink, C.H.J. (*submitted*). Is ostracism by a despised group suffering? A replication of Gonsalkorale and Williams (2007).

Floor, L. (2007). *De effecten van groepslidmaatschap op ostracisme en pesten. Wat is erger: genegeerd of gepest worden? [The effects of group membership on ostracism and bullying. What is worse: being ignored or being bullied?]*. University Leiden.

Gallardo-Pujol, D., Andrés-Pueyo, A., & Maydeu-Olivares, A. (2012). MAOA genotype, social exclusion and aggression: An experimental test of a gene-environment interaction. *Genes, Brain, and Behavior*. doi:10.1111/j.1601-183X.2012.00868.x

Gan, Y., & Liu, J. (2012).The mechanism by which interpersonal coping flexibility influences self-esteem. *The Psychological Record*, *62*, 735–746.

Garczynski, A. M., Brown, C. M., & Harvey, R. D. (2013). Temporal perspective moderates self-reported reactions to social rejection. *SPSP Conference abstract*.

Geniole, S. N., Carré, J. M., & McCormick, C. M. (2011). State, not trait, neuroendocrine function predicts costly reactive aggression in men after social exclusion and inclusion. *Biological Psychology*, *87*, 137–145. doi:10.1016/j.biopsycho.2011.02.020

Gerber, J. P., Williams, K. D., & Wheeler, L. (*in preparation*). Clarifying the relationship between ostracism and relational devaluation.

Gonsalkorale, K., & Williams, K. D. (2007). The KKK won’t let me play: ostracism even by a despised outgroup hurts. *European Journal of Social Psychology*, *37*, 1176–1186. doi:10.1002/ejsp.392

Goodwin, S. A., Williams, K. D., & Carter-Sowell, A. R. (2010). The psychological sting of stigma: The costs of attributing ostracism to racism. *Journal of Experimental Social Psychology*, *46*, 612–618. doi:10.1016/j.jesp.2010.02.002

Greitemeyer, T., Fischer, P., & Kastenmüller, A. (2012).The effects of social exclusion on confirmatory information processing. *European Journal of Social Psychology*, *42*, 462–469. doi:10.1002/ejsp.1851

Gruijters, S. (*unpublished raw data*). Social connection and anthropomorphism.

Hackenbracht, J., & Gasper, K. (*in press*). I’m All Ears: The Need to Belong Motivates Listening to Emotional Disclosure. *Journal of Experimental Psychology*.

Hawes, D. J., Zadro, L., Fink, E., Richardson, R., O’Moore, K., Griffiths, B., Dadds, M. R., et al. (2012). The effects of peer ostracism on children’s cognitive processes. *European Journal of Developmental Psychology*, *9*, 599–613. doi:10.1080/17405629.2011.638815

Hellmann, J. H., & Echterhoff, G. (*in preparation*). Ostracism and Facebook.

Hess, Y. D., & Pickett, C. L. (2010). Social rejection and self- versus other-awareness. *Journal of Experimental Social Psychology*, *46*, 453–456. doi:10.1016/j.jesp.2009.12.004

Hess, Y., & Pickett, C. (2011). I’ll stick with my idea: exclusion increases socially dominant behaviors. *SPSP Conference abstract*.

Horn, N. R. (*n.d.*). *Social exclusion and the bitter coldness of rejection: can drinking a warm drink after being socially excluded counterbalance the negative effects?* Retrieved from https://dspace.washcoll.edu/handle/10090/19133?show=full

Ijzerman, H., Gallucci, M., Pouw, W. T. J. L., Weiβgerber, S. C., Van Doesum, N. J., & Williams, K. D. (2012). Cold-blooded loneliness: social exclusion leads to lower skin temperatures. *Acta Psychologica*, *140*, 283–288. doi:10.1016/j.actpsy.2012.05.002

Jamieson, J. P., Harkins, S. G., & Williams, K. D. (2010). Need threat can motivate performance after ostracism. *Personality & Social Psychology Bulletin*, *36*, 690–702. doi:10.1177/0146167209358882

Johnson, C. (2010). *Personality and ostracism: do hope, optimism, and forgiveness moderate the effects of social exclusion?* Purdue University. Retrieved from https://scholarworks.iupui.edu/bitstream/handle/1805/2778/CBJohnson\_Thesis\_FINAL.pdf?sequence=1

Kassner, M. P., Dongning, R., Law, A. T., & Williams, K. D. (*in preparation*). Effects of Mental Visualization and Degraded Presentation on Detection and Influence of Ostracism.

Kassner, M. P., Wesselmann, E. D., Law, A. T., & Williams, K. D. (2012). Virtually ostracized: studying ostracism in immersive virtual environments. *Cyberpsychology, Behavior and Social Networking*, *15*, 399–403. doi:10.1089/cyber.2012.0113

Kerr, N. L., Seok, D.-H., Poulsen, J. R., Harris, D. W., & Messé, L. A. (2008). Social ostracism and group motivation gain. *European Journal of Social Psychology*, *38*, 736–746. doi:10.1002/ejsp

Kesting, M.-L., Bredenpohl, M., Klenke, J., Westermann, S., & Lincoln, T. M. (2012). The impact of social stress on self-esteem and paranoid ideation. *Journal of Behavior Therapy and Experimental Psychiatry*, *44*, 122–128. doi:10.1016/j.jbtep.2012.07.010

Knowles, M. (2010).Use of social media in the service of belonging needs. *SPSP Conference abstract*.

Knowles, M., & Caroll, C. (2012).Shifting perspectives: the impact of social rejection and acceptance on perspective-taking. *SPSP Conference abstract*.

Krijnen, J. (*n.d.*). Research Report “BP29 Complete.”

Krill, A. L., Platek, S. M., &Wathne, K. (2008). Feelings of control during social exclusion are partly accounted for by empathizing personality. *Personality and Individual Differences*, *45*, 684–688. doi:10.1016/j.paid.2008.07.016

Lakin, J. L., Chartrand, T. L., & Arkin, R. M. (2008). I am too just like you: nonconscious mimicry as an automatic behavioral response to social exclusion. *Psychological science*, *19*, 816–822. doi:10.1111/j.1467-9280.2008.02162.x

Lau, G., Moulds, M. L., & Richardson, R. (2009). Ostracism: how much it hurts depends on how you remember it. *Emotion*, *9*, 430–434. doi:10.1037/a0015350

Lustenberger, D. E., & Jagacinski, C. M. (2010). Exploring the Effects of Ostracism on Performance and Intrinsic Motivation. *Human Performance*, *23*, 283–304. doi:10.1080/08959285.2010.501046

MacDonald, G. (2008). Use of pain threshold reports to satisfy social needs. *Pain Research & Management*, *13*(4), 309–319.

McDonald, M. M., & Donnellan, B. M. (2012). Is ostracism a strong situation? The influence of personality in reactions to rejection. *Journal of Research in Personality*, *46*, 614–618. doi:10.1016/j.jrp.2012.05.008

Nordgren, L. F., Banas, K., &MacDonald, G. (2011). Empathy gaps for social pain: why people underestimate the pain of social suffering. *Journal of Personality and Social Psychology*, *100*, 120–128. doi:10.1037/a0020938

Nordgren, L. F., McDonnell, M.-H.M., & Loewenstein, G. (2011). What constitutes torture?: psychological impediments to an objective evaluation of enhanced interrogation tactics. *Psychological Science*, *22*, 689–694. doi:10.1177/0956797611405679

O’Brien, E., Ellsworth, P. C., & Schwarz, N. (2012). Today’s misery and yesterday's happiness: Differential effects of current life-events on perceptions of past wellbeing. *Journal of Experimental Social Psychology*, *48*, 968–972. doi:10.1016/j.jesp.2012.02.018

Oberleitner, D. E. (2012). *Accessibility for aggression and negative self-views following ostracism*. Wayne State University. Retrieved from http://digitalcommons.wayne.edu/cgi/viewcontent.cgi?article=1553&context=oa\_dissertations

Peterson, C. K., Gravens, L. C., & Harmon-Jones, E. (2011).Asymmetric frontal cortical activity and negative affective responses to ostracism. *Social Cognitive and Affective Neuroscience*, *6*, 277–285. doi:10.1093/scan/nsq027

Pharo, H., Gross, J., Richardson, R., & Hayne, H. (2011). Age-related changes in the effect of ostracism. *Social Influence*, *6*, 22–38. doi:10.1080/15534510.2010.525852

Plaisier, X. S., & Konijn, E. A. (2012). Rejected by Peers-Attracted to Antisocial Media Content: Rejection-Based Anger Impairs Moral Judgment Among Adolescents. *Developmental Psychology*. doi:10.1037/a0029399

Ramirez, M. C. (2009). *The influence of contingent self-esteem and self-esteem variability on reactions to ostracism*. University of Texas at Arlington. Retrieved from http://dspace.uta.edu/bitstream/handle/10106/2034/Ramirez\_uta\_2502M\_10456.pdf?sequence=1

Ren, D., & Williams, K. D. (2012). Self-construals matter when coping with ostracism: but not when experiencing it. *SPSP Conference abstract*.

Renneberg, B., Herm, K., Hahn, A., Staebler, K., Lammers, C.-H., & Roepke, S. (2011). Perception of social participation in borderline personality disorder. *Clinical Psychology & Psychotherapy*. doi:10.1002/cpp.772

Riva, P., Wirth, J. H., & Williams, K. D. (2011). The consequences of pain: The social and physical pain overlap on psychological responses. *European Journal of Social Psychology*, *41*, 681–687. doi:10.1002/ejsp.837

Ruggieri, S., Bendixen, M., Gabriel, U., & Alsaker, F. (*submitted*). Do victimization experiences accentuate reactions to ostracism? An experiment using Cyberball.

Ruggieri, S., Bendixen, M., Gabriel, U., & Alsaker, F. (*in press*). Cyberball: the impact of ostracism on early adolescents’ well-being. *Swiss Journal of Psychology*. doi:10.1024/1421-0185/a000103

Sacco, D. F., Wirth, J. H., Hugenberg, K., Chen, Z., & Williams, K. D. (2011). The world in black and white: Ostracism enhances the categorical perception of social information. *Journal of Experimental Social Psychology*, *47*, 836–842. doi:10.1016/j.jesp.2011.03.001

Salvy, S.-J., Bowker, J. C., Nitecki, L. A., Kluczynski, M. A., Germeroth, L. J., & Roemmich, J. N. (2010). Impact of simulated ostracism on overweight and normal-weight youths’ motivation to eat and food intake. *Appetite*, *56*, 39–45. doi:10.1016/j.appet.2010.11.140

Salvy, S.-J., Bowker, J. C., Nitecki, L. A., Kluczynski, M. A., Germeroth, L. J., & Roemmich, J. N. (2012). Effects of ostracism and social connection-related activities on adolescents’ motivation to eat and energy intake. *Journal of Pediatric Psychology*, *37*, 23–32. doi:10.1093/jpepsy/jsr066

Schaafsma, J., & Williams, K. D. (2012).Exclusion, intergroup hostility, and religious fundamentalism. *Journal of Experimental Social Psychology*, *48*, 829–837. doi:10.1016/j.jesp.2012.02.015

Segovia, K. Y., & Bailenson, J. N. (2012). Virtual imposters: Responses to avatars that do not look like their controllers. *Social Influence*, *7*, 285–303. doi:10.1080/15534510.2012.670906

Staebler, K., Renneberg, B., Stopsack, M., Fiedler, P., Weiler, M., & Roepke, S. (2011).Facial emotional expression in reaction to social exclusion in borderline personality disorder. *Psychological Medicine*, *41*, 1929–1938. doi:10.1017/S0033291711000080

Stillman, T. F., Baumeister, R. F., Lambert, N. M., Crescioni, A. W., Dewall, C. N., & Fincham, F. D. (2009). Alone and Without Purpose: Life Loses Meaning Following Social Exclusion. *Journal of Experimental Social Psychology*, *45*, 686–694. doi:10.1016/j.jesp.2009.03.007

Stock, M. L., Gibbons, F. X., Walsh, L. A., &Gerrard, M. (2011). Racial identification, racial discrimination, and substance use vulnerability among African American young adults. *Personality & Social Psychology Bulletin*, *37*, 1349–1361. doi:10.1177/0146167211410574

Van Beest, I., Williams, K. D., & Van Dijk, E. (2011). Cyberbomb: Effects of being ostracized from a death game. *Group Processes & Intergroup Relations*, *14*, 581–596. doi:10.1177/1368430210389084

Van Beest, I., Carter-Sowell, A. R., Van Dijk, E., & Williams, K. D. (2012). Groups being ostracized by groups: Is the pain shared, is recovery quicker, and are groups more likely to be aggressive? *Group Dynamics: Theory, Research, and Practice*, *16*, 241–254. doi:10.1037/a0030104

Van Beest, I,, & Williams, K. D. (2006). When inclusion costs and ostracism pays, ostracism still hurts. *Journal of Personality and Social Psychology*, *91*, 918–928. doi:10.1037/0022-3514.91.5.918

Van Dijk, W. W., & Williams, K. D. (*unpublished raw data*). Social exclusion and schadenfreude.

Webb, T. L., Harris, P., & McAtamney, K. (*unpublished manuscript*). Self-Affirmation Moderates the Negative Effect of Ostracism on Perceptions of Control, But Not in Predictable Ways.

Weik, U., Maroof, P., Zöller, C., &Deinzer, R. (2010). Pre-experience of social exclusion suppresses cortisol response to psychosocial stress in women but not in men. *Hormones and behavior*, *58*, 891–897. doi:10.1016/j.yhbeh.2010.08.018

Wesselmann, E. D., Bagg, D., & Williams, K. D. (2009). “I Feel Your Pain”: The effects of observing ostracism on the ostracism detection system. *Journal of Experimental Social Psychology*, *45*, 1308–1311. doi:10.1016/j.jesp.2009.08.003

Wesselmann, E. D., Wirth, J. H., Mroczek, D. K., & Williams, K. D. (2012). Dial a feeling: Detecting moderation of affect decline during ostracism. *Personality and Individual Differences*, *53*, 580–586. doi:10.1016/j.paid.2012.04.039

Williams, K D, Cheung, C. K., & Choi, W. (2000). Cyberostracism: effects of being ignored over the Internet. *Journal of Personality and Social Psychology*, *79*(5), 748–762.

Williams, K. D, Govan, C. L., Croker, V., Tynan, D., Cruickshank, M., & Lam, A. (2002). Investigations into differences between social- and cyberostracism. *Group Dynamics: Theory, Research, and Practice*, *6*, 65–77. doi:10.1037//1089-2699.6.1.65

Wirth, J H, & Williams, K. D. (2009). `They Don’t Like Our Kind': Consequences of Being Ostracized While Possessing a Group Membership. *Group Processes & Intergroup Relations*, *12*, 111–127. doi:10.1177/1368430208098780

Wirth, James H, Lynam, D. R., & Williams, K. D. (2010). When social pain is not automatic: Personality disorder traits buffer ostracism’s immediate negative impact. *Journal of Research in Personality*, *44*, 397–401. doi:10.1016/j.jrp.2010.03.001

Zadro, L., Boland, C., & Richardson, R. (2006). How long does it last? The persistence of the effects of ostracism in the socially anxious. *Journal of Experimental Social Psychology*, *42*, 692–697. doi:10.1016/j.jesp.2005.10.007

Zadro, L., Williams, K. D., & Richardson, R. (2004). How low can you go? Ostracism by a computer is sufficient to lower self-reported levels of belonging, control, self-esteem, and meaningful existence. *Journal of Experimental Social Psychology*, *40*, 560–567. doi:10.1016/j.jesp.2003.11.006

Zhong, C.-B., & Leonardelli, G. J. (2008). Cold and lonely: does social exclusion literally feel cold? *Psychological Science*, *19*, 838–842. doi:10.1111/j.1467-9280.2008.02165.x

Zöller, C., Maroof, P., Weik, U., & Deinzer, R. (2010). No effect of social exclusion on salivary cortisol secretion in women in a randomized controlled study. *Psychoneuroendocrinology*, *35*, 1294–1298. doi:10.1016/j.psyneuen.2010.02.019

Zwolinski, J. (2012). Psychological and Neuroendocrine Reactivity to Ostracism. *Aggressive Behavior*, *38*, 108–125. doi:10.1002/ab.21411