Supplemental Material A – Visual Representation of Outliers and Influential Cases

Calculated effects include (1) externally standardized residuals, (2) DFFITS values, (3) Cook's distances, (4) covariance ratios, (5) leave-one-out estimates of the amount of heterogeneity, (6) leave-one-out values of the test statistics for heterogeneity, (7) hat values, and (8) weights. Syntax and additional information can be found at: [http://www.metafor-project.org/doku.php/plots:plot\_of\_influence\_diagnostics](http://www.metafor-project.org/doku.php/plots%3Aplot_of_influence_diagnostics) .

Supplemental Material A Continued – Visual Representation of Outliers and Influential Cases



Supplemental Material A Continued – Visual Representation of Outliers and Influential Cases



Supplemental Material B – Justification for Removing Van de Castle (1971)

After performing the initial outlier analysis, a second integration of the analyses (Supplemental Material A) inculcated that Van de Castle (1971) was a notable outlier that should be addressed. While I agree with Storm et al.’s (2001) assessment that dated studies should not be excluded from meta-analyses due to their age, this study does raise some concerns. The weighted effect size reported by large-n studies was .02 when excluding Van de Castle, whereas this study found an effect size of .29. Van de Castle is an outlier whether compared to the overall database or only large-n studies.

Also, the methods applied by Van de Castle contradict most beliefs about obtaining large effects via dream ESP studies. The author provided a magazine picture to a youth camp counselor, and the counselor “periodically concentrated on the target during the night” while housed several hundred yards from about 70 sleeping campers (Van de Castle, 1971, p. 312). The next morning, the target picture along with four others were presented to the campers and they ranked the pictures based on correspondence. A rank of 1 or 2 represented a hit, whereas a rank of 4 or 5 represented a miss. This process was repeated for four total nights. A total of 95 hits and 55 misses were obtained. Because about 70 campers participated each night for four nights, this would suggest that the total number of trails is between 240 and 320. Thereby, this would suggest that 95 trials were a hit (1 or 2), 90 to 170 trials were in the middle (3), and 55 trials were misses (4 or 5); however, Van de Castle did not mention this anomalous number of middle responses, which is an even larger outlier than the hits. Likewise, it seems unusual that a single person housed several hundred yards away could have such a strong influence on the ESP of others, given that similar studies have not produced such strong effects when more closely collocated. Even two authors of Storm et al., Sherwood and Roe (2003), previously expressed concerns regarding the methodology of Van de Castle (1971), perhaps because they too doubted such a strong effect from this research design (although I do not want to speak on their behalf).

 Lastly, only an abstract can be found for Van de Castle, because it was a conference presentation with a research brief published in the Journal of Parapsychology. Van de Castle may have presented additional information that absolves these concerns, but this cannot be known given current records. Given these concerns, I chose not to include Van de Castle in the current reanalysis of Storm et al.’s meta-analytic results.

**References**

Sherwood, S., & Roe, C. A. (2003). A review of dream ESP studies conducted since the Maimonides dream ESP programme. *Journal of Consciousness Studies*, *10*(6-7), 85-109.

Storm, L., & Ertel, S. (2001). Does psi exist? Comments on Milton and Wiseman's (1999) meta-analysis of Ganzfield research. *Psychological Bulletin*, 127(3), 424-433.

Supplemental Material C – Reanalysis of Results with Van de Castle (1971)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  | Implied Missing |
|  | k | *I2* | Fail Safe k | Egger’s Test *β0* | Egger’s Test t | Left of Mean | Right of Mean |
| Overall | 45a | .567 | 133 | .391 | 1.687\* | 7 | 0 |
| Telepathy | 22a | 0.00 | 15 | .155 | .616 | 2 | 0 |
| Clairvoyance | 12 | 26.713 | 7 | -.535 | .532 | 0 | 3 |
| Precognition | 9 | 27.487 | 0 | .573 | .834 | 1 | 0 |

† p < .10

\* p < .05

\*\* p < .01

\*\*\* p < .001

a Indicates difference between primary analysis and supplemental analysis.

Supplemental Material C Continued - Reanalysis of Results with Van de Castle (1971)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | k | n | $$\overbar{r}$$ | 95% C.I. | z-value | p-value | Q (df) |
| Overall | 45a | 1858 | .09 | .04, .14 | 3.658 | <.001\*\*\* | 44.251 (44) |
| Small n | 24 | 202 | .20 | .03, .36 | 2.299 | .022\* | 13.393 (23) |
| Medium n | 13 | 345 | .13 | -.01, .27 | 1.820 | .069† | 18.683 (12) |
| Large n | 8a | 1311 | .07 | .00, .13 | 2.057 | .040\* | 8.935 (7) |
| Telepathy | 22a | 599 | .13 | .04, .21 | 2.939 | .003\*\* | 12.416 (21) |
| Small n | 17 | 140 | .14 | -.07, .33 | 1.309 | .191 | 6.921 (16) |
| Medium n | 3 | 106 | .11 | -.09, .30 | 1.075 | .282 | .686 (2) |
| Large n | 2a | 353 | .14 | -.09, .36 | 1.213 | .225 | 4.763 (1) |
| Clairvoyance | 12 | 236 | .19 | .03, .35 | 2.252 | .024\* | 15.009 (11) |
| Small n | 4 | 34 | .31 | -.10, .63 | 1.495 | .135 | 2.000 (3) |
| Medium n | 8 | 202 | .17 | -.03, .36 | 1.664 | .096† | 12.673 (7) |
| Large n | 0 | 0 | - | - | - | - | - |
| Precognition | 9 | 823 | .04 | -.06, .13 | .796 | .426 | 11.032 (8) |
| Small n | 3 | 28 | .39 | -.18, .76 | 1.365 | .172 | 3.337 (2) |
| Medium n | 2 | 37 | -.05 | -.64, .57 | -.150 | .881 | 3.954 (1) |
| Large n | 4 | 758 | .03 | -.04, .10 | .863 | .388 | 1.600 (3) |

Note: k = number of studies ; n = sample size ; $\overbar{r}$ = inverse-variance weighted effect size ; 95% CI = 95 percent confidence interval ; Q = Cochran’s Q ; df = degrees of freedom; Small n = studies with 15 or fewer participants ; Medium n = studies with 15 to 99 participants ; Large n = studies with 100 or more participants.

† p < .10

\* p < .05

\*\* p < .01

\*\*\* p < .001

a Indicates difference between primary analysis and supplemental analysis.

Supplemental Material C Continued - Reanalysis of Results with Van de Castle (1971)

|  |
| --- |
| Analyses Separated by Author |
|  | k | N | $$\overbar{r}$$ | 95% C.I. | z-value | p-value | Q (df) |
| Braud | 2 | 66 | .12 | -.13, 36 | .950 | .342 | .657 (1) |
| Child | 2 | 15 | .63 | .09, .88 | 2.219 | .027\* | .194 (1) |
| Dalton | 3 | 76 | .38 | .16, .57 | 3.292 | .001\*\* | .610 (2) |
| Foulkes | 2 | 16 | .00 | -.55, .55 | .003 | .997 | .153 (1) |
| Harley | 2 | 40 | .00 | -.52, .52 | .000 | 1.000 | 2.987 (1) |
| Hearne | 6 | 42 | .08 | -.31, .45 | .406 | .685 | .510 (5) |
| Kanthamani | 3 | 34 | .28 | -.11, .59 | 1.429 | .153 | .777 (2) |
| Krippner | 11 | 294 | .07 | -.05, .19 | 1.095 | .273 | 7.484 (10) |
| Luke | 2 | 411 | -.00 | -.10, .09 | -.079 | .937 | .220 (1) |
| Markwick | 2 | 200 | .08 | -.07, .21 | 1.055 | .291 | .276 (1) |
| Roe | 5 | 138 | .06 | -.12, .23 | .650 | .516 | 3.202 (4) |
| Other | 5 | 526 | .12 | -.06, .28 | 1.312 | .190 | 10.946 (4) |
| Analyses Separated by Publication Outlet |
|  | k | N | $$\overbar{r}$$ | 95% C.I. | z-value | p-value | Q (df) |
| ACPPA | 4 | 50 | .30 | -.00, .56 | 1.937 | .053† | .836 (3) |
| Book | 2 | 16 | .00 | -.55, .55 | .000 | 1.000 | .155 (1) |
| EJPP | 2 | 120 | .07 | -.12, .26 | .742 | .458 | 1.026 (1) |
| EMS | 2 | 20 | .00 | -.55, .55 | .002 | .999 | 1.361 (1) |
| IJPP | 2 | 19 | .15 | -.38, .60 | .533 | .594 | .118 (1) |
| JASPR | 3 | 48 | .52 | .25, .71 | 3.571 | <.001\*\*\* | .467 (2) |
| JPP | 3a | 409 | .21 | .02, .39 | 2.136 | .033\* | 5.134 (1) |
| JSPR | 12 | 200 | .03 | -.13, .18 | .325 | .746 | 5.687 (11) |
| PMS | 2 | 16 | .00 | -.55, .55 | .003 | .997 | .153 (1) |
| RPP | 7 | 298 | .09 | -.07, .25 | 1.079 | .281 | 8.671 (6) |
| Other | 6 | 662 | .02 | -.06, .10 | .566 | .571 | 3.127 (5) |

Note: k = number of studies ; n = sample size ; $\overbar{r}$ = inverse-variance weighted effect size ; 95% CI = 95 percent confidence interval ; Q = Cochran’s Q ; df = degrees of freedom ; ACPPA = Annual Convention of the Parapsychological Association ; Book = Dream Telepathy: Experiments in nocturnal ESP ; EJPP = European Journal of Parapsychology ; EMS = Experimental Medicine & Surgery ; IJPP = International Journal of Parapsychology ; JASPR = Journal of the American Society for Psychical Research ; JPP = Journal of Parapsychology ; JSPR = Journal of the Society for Psychical Research ; PMS = Perceptual and Motor Skills ; RPP = Research in Parapsychology ; Other = Publication outlet with only one represented study.

† p < .10

\* p < .05

\*\* p < .01

\*\*\* p < .001

a Indicates difference between primary analysis and supplemental analysis.

Supplemental Material C – Reanalysis of Results with Van de Castle (1971)

* I tested whether the meta-analytic effects differed by ESP mode via calculating a between-group Q. The ESP mode had a marginally significant effect (QB = 5.482, df = 2, p = .065), suggesting that the meta-analytic effect sizes may indeed differ by the ESP mode.
* I reanalyzed the effect of MDL. The 11 MDL studies produced an inverse-variance weighted effect size of .07 (z-value = 1.095; p = .273; 95% CI [-.05, .19]), which is less than one-fourth of the originally reported effect (.33). The 34 non-MDL studies produced an inverse-variance weighted effect size of .10 (z-value = 3.255; p = .001; 95% CI [.04, .16]), which is smaller than the originally reported effect (.14). The confidence interval of the non-MDL studies was entirely contained in the MDL studies, indicating that there was not a significant difference in the two groups. This was further supported by the meta-regression that failed to reach statistical significance (Int = .095, B = -.022, SE = .071, p = .754; 95% CI [-.16, .12]).
* I reanalyzed the effect of REM. The 18 REM studies produced an inverse-variance weighted effect size of .06 (z-value = 1.107; p = .268; 95% CI [-.05, .17]), which is one-fourth of the originally reported effect (.24). The 27 non-REM studies produced an inverse-variance weighted effect size of .11 (z-value = 2.986; p = .003; 95% CI [.04, .18]), which is smaller than the originally reported effect (.16). Because the confidence intervals greatly overlapped, there was not a statistically significant difference between the groups. This was supported by the meta-regression (Int = .096, B = -.031, SE = .065, p = .637; 95% CI [-.16, .10]).
* I performed a meta-regression to identify the impact of publication year, which was marginally significant (Int = 4.940; B = -.002, SE = .001, p = .073; 95% CI [-.01, .00]).
* I determined whether effect sizes differed by the author. Studies were grouped in 12 author categories as defined by Storm et al., and the results showed that this grouping approach was not statistically significant (QB = 16.235, df = 11, p = .133).
* I conducted analyses to probe the effect of sample size on study results. Three groups were created that logically appeared in the database. The first included studies with a sample size over 99 (k = 8), the second included studies with a sample size of 15 to 99 (k = 13), and the third included studies with a sample size below 15 (k = 24). The large-n studies produced a small effect that was statistically significance (r = .066; z-value = 2.057; p = .040; 95% CI [.00, .13]); the moderate-n studies produced a small effect that was marginally significant (r = .131; z-value = 1.820; p = .069; 95% CI [-.01, .27]); and the small-n studies produced a moderate effect that was statistically significant (r = .199; z-value = 2.299; p = .022; 95% CI [.03, .36]). The Begg and Mazumdar rank correlation test, which analyzes the relationship between standard errors and effect sizes, was marginally significant (Kendall’s τ = .159, z = 1.536, p = .062). Because this test is often underpowered (Borenstein, 2005), this result suggests that a significant relationship likely exists between sample size and effect size.
* I conducted analyses to determine whether effect sizes differed by publication outlet. Eleven categories were identified, and the results showed that grouping the studies by outlet produced a marginally significant effect (QB = 17.516, df = 10, p = .064). Also, the 35 studies published in parapsychology journals produced an inverse-variance weighted effect size of .13 (z-value = 4.148; p < .001; 95% CI [.07, .19]), whereas the 10 studies published in non-parapsychology journals produced an inverse-variance weighted effect size of .02 (z-value = .543; p = .587; 95% CI [-.06, .10]). The small overlap in confidence intervals suggests that this difference is significant, a meta-regression further supported the statistical significance of this effect (Int = .021, B = .110, SE = .049, p = .026; 95% CI [.01, .21]).