HOW TO ENTER AND SYNTHESIZE THE DATA

by Simon Moss

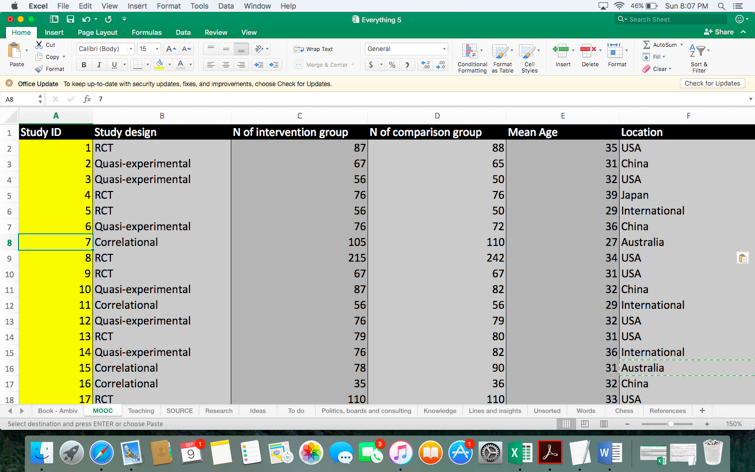
Introduction

Suppose you want to conduct a systematic review to explore whether mindfulness meditation—a variant of meditation—diminishes the incidence of colds. After you have collected the relevant studies, you need to extract, enter, analyze, and synthesize the data in these studies. This document offers some insights into how you should extract, enter, analyze, and synthesize these data effectively. However, this document does not demonstrate how you could conduct a meta-analysis, in which you analyze these data statistically.

## Construct a data extraction form

To help you extract the data, you first need to develop a data extraction form—a template in which you can record these data. As the following example illustrates, many researchers utilize Microsoft Excel or other spreadsheets to record these data. In this form, each row represents one study. Each column represents one characteristic of each study.

Figure 1: Excerpt of a data extraction form



The characteristics, sometimes called fields, varies across studies. For example, if researchers plan to conduct a meta-analysis as well, they will tend to include more fields that revolve around statistical information. Nevertheless, this table presents some typical characteristics as well as the data from two studies. This table could guide your choices in which characteristics or fields to include.

Table 1: Data extraction form

|  |  |  |
| --- | --- | --- |
| Study ID | S001 | S002 |
| Reference | Smith, A. (2018). Mindfulness and infections. Journal of Infections, 1, 1-7. | Jones, B. (2017). Meditation and immunity. Journal of Health, 15, 98-108. |
| Funding body | ARC Discovery Grant | None specified |
| People who extracted the data | Brown, H. | Brown, H. |
| People who assessed this extraction | Anderson, F. | Anderson, F. |
| **Participant characteristics** |  |  |
| N in intervention group | 87 | 132 |
| Withdrawal rate from intervention group | Before treatment: 10%  Before measure: 5%  After measure: 6% | 15% |
| % female in intervention group | 52% | 59% |
| N in comparison group | 86 | 131 |
| % female in comparison group | 48% | 41% |
| Withdrawal rate from comparison group | Before treatment: 6%  Before measure: 2%  After measure: 0% | 8% |
| Age | Mean: 32  Range: 18-60 | Median: 31  Standard deviation: 4.5 |
| Education | 52% Bachelor degree or higher | 28% Bachelor  6% Honors or Graduate Diploma  8% Masters or Doctorate |
| Location | English-speaking nations | China |
| Source of recruitment | Amazon Mechanical Turk | One government organization |
| Procedure to recruit | Advertisement on Amazon Mechanical Turk | Email invitation |
| Remuneration | .50 $US | None |
| **Intervention** |  |  |
| Magnitude of intervention, such as duration | 1 hour a week for 12 weeks | 3 hours a week for 10 weeks |
| Magnitude of comparison, such as duration | 1 hour a week for 12 weeks | 3 hours a week for 10 weeks |
| Activities that only the intervention group completed | Instructions to observe thoughts, feelings, and sensation without judging | Meditation, coupled with instructions to observe thoughts, feelings, and sensation without judging |
| Activities that only the comparison group completed | Instructions to diminish unpleasant thoughts, feelings, and sensation | Exercise |
| Activities that both groups completed | Meditation guiding individuals to orient their attention to pleasant images | None |
| **Outcome measures** |  |  |
| Self-report measures | Frequency of colds | Duration of colds |
| Medical records |  | Frequency of colds |
| HR records |  | Number of absent days ascribed to colds |
| **Design features** |  |  |
| Procedure to randomly allocate participants | Computer random number generator; allocated to intervention group if the number is even | None |
| Procedure to conceal allocation from researchers who interact with participants | Researchers distributed audio to participants, oblivious to the instructions on this audio recording | None |
| Procedure to conceal allocation from researchers who measure participants | Researchers restricted conversation to instructions about measures |  |
| **Data analysis** |  |  |
| Main technique | Binary logistic regression | Independent t-test |
| Control variables | Age, education | None |
| Transformations deployed | Log transformation of frequency | None |
| **Key results** |  |  |
| Measures that were significantly higher in the intervention group | None | None |
| Measures that were significantly higher in the comparison group | Frequency of colds | Frequency of colds  Number of absent days ascribed to colds |
| Measures that did not vary significantly between the groups | None | Duration of colds |
| Effect sizes | Frequency of colds: d = .5 | Duration of colds: d = .23  Frequency of colds: d = .31  Number of absent days ascribed to colds: d = .19 |
| **Other information** |  |  |
| Limitations the authors acknowledged | No baseline measures | None |

## Pilot the data extraction form

Once you have designed a preliminary data extraction form, you should test the utility of this form. In particular, at least two researchers—usually yourself and one supervisor—should conduct this task independently. Both researchers, for example, could utilize this form to extract data from 10 studies. The researchers would then

* Record instances in which they could not identify enough details to complete a field
* Record instances in which they felt no field was available to enter information that seemed important
* Discuss whether the form could be improved to address these concerns. For example, you might write more instructions on how to enter the data for specific fields.

## Enter the data into the data extraction form

After you have modified and improved the data extraction form, you need to enter the data of all the studies. Occasionally, because of human error or ambiguous reports, some of the data may be entered incorrectly. Consequently, researchers will tend to utilize one or more approaches to assess the accuracy of these data. You could

|  |
| --- |
| Approaches to assess the reliability of data |
| Contact the authors of these studies to clarify ambiguous information. |
| Enter the data for all studies twice. You can then calculate a statistic called Cohen’s Kappa coefficient to assess whether your entries changed over time. |
| Invite an independent person—such as a supervisor—to enter data for a subset of studies. You could again compute Cohen’s Kappa coefficient to measure the extent to which the two researchers enter the same categorical data. |
| Disagreements between researchers could be discussed and, preferably, resolved. |
| If disagreements are not resolved, you could present the most favorable and least favorable interpretation of the results, similar to a sensitivity analysis. |

## Design tables to display the data

To report the results, some researchers construct a table that resembles the data extraction form—similar to Table 1. However, if you extract many fields and collect many studies, this form might be too long to include in a paper. Instead, to solve this problem, you may consider a variety of approaches.

First, rather than include a table that presents information about each study, you could construct a summary table instead. An extract of a summary table appears below.

|  |  |
| --- | --- |
| Study characteristic | Summary |
| Remuneration of participants | * 20% remunerated participants * In these studies, mean payment was $1.50 (sd = .23) * Effect sizes was .34 if participants were renumerated and .32 if participants were not remunerated |
| Statistical techniques utilized | * 30% utilized independent t-tests * 25% utilized logistic regression * 15% utilized MANOVA * 30% utilized ANCOVAs * Proportion of significant results was roughly 30% regardless of the statistical test utilized |

Second, you could include two tables: one table that presents some information about each study and one summary table. The table that presents information about each study would include a subset of fields only. Fields that are not as likely to affect the results—such as whether the participants were remunerated and the statistical technique that was utilized—could be included in the summary table only.

Third, to truncate the table, you can collapse some of the fields. For example, if the writer is concise, the five rows that describe the intervention in Table 1 could be collapsed into one row.

Fourth, you could construct a graph, such as a forest plot, to display numerical fields—such as the number of participants in each condition and the effect sizes. This plot can integrate several numerical fields into one graph. The length of lines, numbers, and shapes may correspond to distinct fields. Many software programs, such as the Comprehensive Meta-analysis 3.0, can be used to construct funnel plots



## Narrative synthesis

Finally, you need to construct a narrative to summarize or describe these tables. For example, for each intervention or measure, you could describe the significant and non-significant results. Typical sentences might include

* Mindfulness meditation, relative to other variants of meditation, significantly reduced the frequency of colds in 40% of analyses, the duration of colds in 30% of analyses, and the number of absent days in 25% of the analyses. The other analyses generated non-significant results.
* Compared to no meditation at all, mindfulness meditation significantly decreased the frequency of colds in 80% of analyses and generated no significant effect in 20% of analyses.

In addition, for all these tables and narratives, you should comment on which characteristics of the participants, interventions, comparison groups, designs, or settings affect the results. These characteristics are sometimes referred to as moderators. For instance, you might write

* When participants were living in European nations, mindfulness meditation, relative to other variants of meditation, significantly reduced the frequency of colds in 40% of analyses. However, in Asian nations, mindfulness meditation significantly reduced the frequency of colds in only 10% of analyses.