Exploratory and Confirmatory Analyses

The terms *confirmatory* and *exploratory* are used differently by different researchers. Some researchers apply the term confirmatory only to confirmation of a previous empirical study. For these researchers, the initial research testing a theoretical hypothesis is described as exploratory. However, other researchers apply the term confirmatory to the initial research testing (confirming) a theoretical hypothesis.

The KPU Registry intends to be flexible and handle either perspective. However, well-designed confirmatory research should meet certain conditions. The following discussion is offered for those seeking clarification about categorizing an analysis as confirmatory or exploratory. Because a study may include both exploratory and confirmatory components, we discuss categorizing each analysis within a study rather than categorizing an entire study. The creativity in scientific research often emerges during exploratory research. Confirmatory research provides the compelling evidence that makes scientific research valid and self correcting. Both exploration and confirmation are essential for scientific progress.

The distinction between exploratory and confirmatory analyses is well established in regulated medical research (FDA, 1998; NIH, 2016). Phase 1 clinical trials explore the safety and side effects when a new treatment is administered to people and Phase 2 trials explore whether the new treatment is effective in improving a health condition. Phase 3 clinical trials confirm the efficacy and safety effects in larger, carefully planned studies with pre-specified methodology. Phase 3 research is required in regulated medical research, and similar confirmatory research is needed to resolve scientific questions in other areas. However, the distinction between exploratory and confirmatory research has often not been clear in academic research. One symptom of this is when studies are designed without using power analysis to set sample size. Power analysis is an essential step for designing Phase 3 trials, and for any well-designed confirmatory research.

Well-Designed Confirmatory Analyses

A well-designed confirmatory analysis should be capable of providing evidence that an experimental hypothesis is false as well as true. For many analyses in parapsychology and psychology, the interpretation of a nonsignificant outcome is ambiguous due to uncertainty in the methods or theory, or due to low power. These analyses cannot provide clear evidence that a hypothesis is false. They are not well-designed confirmatory analyses. Unfortunately, the statistical rationale for conducting falsifiable research with classical statistics and related methods for optimal Bayesian analyses for confirmatory research are not well known among psychological researchers. These methods are discussed in Kennedy and Watt (2018).

A well-designed confirmatory analysis has the following characteristics.

- The analysis uses data from measurement methods and experimental interventions that are established or have precedent. The interpretation of the results is not confounded by methodological uncertainties or ambiguous theory.
- Sample size is set to usefully test the hypothesis, and typically is based on power analysis or *operating characteristics* with a power of at least .90 and optimally .95 or higher for detecting the minimum effect size of interest.

- All analysis decisions that could affect the results are made before data collection starts. These include: (a) the specific statistical test, (b) whether the test is one-sided or two-sided, (c) the specific numerical *inference criteria* (e.g., magnitude of *p*-value or Bayes factor) for acceptable evidence, (d) any transformations or adjustments to the data, (e) any criteria for excluding or deleting data, and (f) any corrections for multiple analyses.
- The inference criteria for acceptable evidence should include criteria for evidence the hypothesis is false as well as criteria for evidence the hypothesis if true.

Examples:

- Confirming a previous finding using the methods from the previous experiment(s) and power analysis based on the previous experiment(s) or (preferably) a minimum effect of practical or theoretical interest.
- Providing evidence for (confirming) a theoretical hypothesis using established methods. Samples size is set based on a power analysis testing for a reasonably small a priori effect size (i.e., the analysis has a specified power for a certain a priori effect size that would be of interest or would be convincing).

Exploratory Analyses

Exploratory analyses are usually the starting point for a line of research, but typically involve methodological or theoretical uncertainties that can confound the interpretation of the results. The effects being investigated may be inflated by selection from different analysis possibilities, or the interpretation of nonsignificant outcomes may be confounded by methodological or theoretical uncertainty. Exploratory analyses include (a) analyses when specific statistical hypotheses are not precisely pre-specified and the statistical tests evaluate or emerge from different possibilities, (b) analyses of data from measurement methods or experimental interventions that are being developed as part of the study, or (c) small pilot or initial studies when nonsignificant outcomes could be the result of low power rather than evidence that the experimental hypothesis is false.

Examples:

- A vague hypothesis that will be developed into a specific statistical test as the data are analyzed.
- An analysis based on measurement methods or interventions that are not established and are being developed as part of the study.
- An analysis with data transformations, adjustments, or exclusion criteria that are not pre-specified and will be developed as the data are analyzed.
- An analysis of the relationships among several different factors without clear precedents or theoretical predictions about the specific relationships among the factors.
- The experimenter considers the hypothesis as exploratory rather than well developed.

Fully Specified Exploratory Analyses

Certain exploratory analyses provide greater evidence than others. An exploratory analysis with the statistical tests pre-specified in detail and the sample size set to usefully test the hypothesis provides greater evidence than an exploratory analysis based on a vague hypothesis investigated with a small sample size. Fully specified exploratory analyses with pre-specified statistical tests and adequate power are often useful when investigating the mechanism of operation for an established effect or when developing more effective measurement methods or experimental interventions.

The distinction between a confirmatory analysis and a fully specified exploratory analysis usually depends on the degree of confidence in interpreting a nonsignificant outcome. The analysis is appropriately classified as exploratory if evidence that the experimental hypothesis is false may be confounded by uncertainties or ambiguities for the methods or theory. An example would be the use of measurement methods with uncertain validity, or when data adjustments that were not pre-specified are reasonably likely.

Interacting Confirmatory and Exploratory Hypotheses

Many studies include both confirmatory and exploratory hypotheses. In some cases, the exploratory hypotheses do not affect the analysis of the confirmatory hypothesis. However, in other cases, the exploratory and confirmatory hypotheses are analyzed together. For example, a two-way ANOVA may have a confirmatory hypothesis for one factor and an exploratory hypothesis for the other factor. The uncertainties for the exploratory hypothesis may impact the analysis of the confirmatory hypothesis in these situations.

For purposes of study registration, the confirmatory hypothesis can be described as either confirmatory or exploratory in these cases. In general, if it is likely that the planned analysis for the confirmatory hypothesis may need to be modified or may have uncertainties of interpretation due to the exploratory factor, then the analyses of both factors would be appropriately described as exploratory.

Reporting

For a confirmatory analysis, any deviation from the pre-specified analysis should be carefully explained and justified. If a deviation is necessary, results from the originally planned analysis should be reported if possible, as well as the results with the deviation. Deviations include any data transformation, adjustment, or exclusion criterion that was not pre-specified, or any modification of the pre-specified statistical test. If deviations are likely, an analysis is more exploratory than confirmatory.

For an exploratory analysis, deviations from the pre-specified analysis are usually a recognized possibility. Deviations should be explained, but the degree of justification can be much lower than for a confirmatory analysis. For an exploratory analysis with deviations, reporting the results for the exact pre-specified analysis will often be optional. A fully specified exploratory analysis without deviations typically has higher evidential value than an analysis with deviations or than an exploratory analysis that was not fully specified.

Other Confirmatory Practices

In addition to the statistical practices described above and public pre-registration of the study, other experimental practices are recommended for well-designed confirmatory studies—and particularly in controversial areas of research (Kennedy, 2014, 2016, 2017). These practices address potential methodological problems that would not be detected by study registration. The practices include:

- experimental procedures that make it difficult for one person to make undetected intentional (fraudulent) or unintentional data alterations—fraud by one person acting alone should not be easy and tempting with little risk of detection;
- formal documented validation of software used to collect, manage, or analyze data;
- handling incomplete data, dropouts, and other protocol violations carefully and conservatively to avoid subtle, difficult-to-recognize biases;
- open data practices that make research data available to others.

References

FDA (1998). *Guidance for Industry: E9 Statistical Principles for Clinical Trials*. Available at http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm073137.pdf.

Kennedy, J.E. (2014). Experimenter Misconduct in Parapsychology: Analysis Manipulation and Fraud. Available at https://jeksite.org/psi/misconduct.pdf.

Kennedy, J.E. (2016). Is the Methodological Revolution in Psychology Over or Just Beginning? *Journal of Parapsychology*, 80, 56-68. Available at https://jeksite.org/psi/jp16.pdf.

Kennedy, J.E. (2017). Experimenter Fraud: What are Appropriate Methodological Standards? *Journal of Parapsychology*, 81, 63-72. Available at https://jeksite.org/psi/jp17.pdf.

Kennedy, J.E., & Watt, C.A. (2018). How to Plan Falsifiable Confirmatory Research. Available at https://jeksite.org/psi/falsifiable_research.pdf.

NIH (2016). Glossary of Common Terms: Phases of Clinical Trials. Available at https://www.nih.gov/health-information/nih-clinical-research-trials-you/glossary-common-terms.

Return to KPU Registry Home (https://koestlerunit.wordpress.com/study-registry/)

Koestler Parapsychology Unit Study Registry

Version of January 27, 2018 — which has updated references and associated information. The initial version of this guidance document was first posted in November, 2013.