

## Vocabulary<sup>1</sup>

Comparison group	A group of individuals that serves as the basis for comparison when assessing the effects of an intervention on a treatment group. A comparison group typically receives some treatment other than they would normally receive and is therefore distinguished from a control group, which often receives no treatment or "usual" treatment. To make the comparison valid, the composition and characteristics of the comparison group should resemble that of the treatment group as closely as possible. Some studies use a control group in addition to a comparison group.
Confounding variables	In an experiment, any characteristic that differs between the experimental group and the comparison group and is not the independent variable under study. These characteristics or variables "confound" the ability to explain the experimental results because they provide an alternative explanation for any observed differences in outcome. In assessing a classroom curriculum, for example, a confounding variable would exist if some students were taught by a highly experienced instructor while other students were taught by a less experienced instructor. The difference in the instructors' experience level makes it harder to determine if the differences in student outcomes (e.g., grades) were caused by the effects of the curriculum or by the variation in instructors. The likelihood that confounding variables might have affected the outcomes of a study is one of the six NREPP criteria used to rate Quality of Research.
Control group	A group of individuals that serves as the basis of comparison when assessing the effects of an intervention on a treatment group. Depending upon the study design, a control group may receive no treatment, a "usual" or "standard" treatment, or a placebo. The composition and characteristics of the control group should resemble that of the treatment group as closely as possible to make the comparison valid.
Evidence-based	Approaches to prevention or treatment that are based in theory and have undergone scientific evaluation. "Evidence-based" stands in contrast to approaches that are based on tradition, convention, belief, or anecdotal evidence.
Externalizing behaviors	Social behaviors and other external cues that reflect an individual's internal emotional or psychological conflicts. Examples include spontaneous weeping, "acting out," and uncharacteristic aggression. Reduction of externalizing behaviors is a frequently used measure of the success of treatment or intervention for mental or emotional disorders.
Fidelity	Fidelity of implementation occurs when implementers of a research-based program or intervention (e.g., teachers, clinicians, counselors) closely follow or adhere to the protocols and techniques that are defined as part of the intervention. For example, for a school-based prevention curriculum, fidelity could involve using the program for the proper grade levels and age groups, following the developer's recommendations for the number of sessions per week, sequencing multiple program components correctly, and conducting assessments and evaluations using the recommended or provided tools.

<sup>1</sup> NREPP: National Registry of Evidence-based Programs and Practices...  
<http://www.nrepp.samhsa.gov/help-glossary.htm>

Internalizing behaviors	Behaviors that reflect an individual's transfer of external social or situational stresses to emotional, psychological, or physical symptoms. One well-known internalizing behavior is a child's development of stomach cramps when the parents argue; another is insomnia during a high-stress situation at work. Reduction of internalizing behaviors is a frequently used measure of the success of treatment or intervention for mental or emotional disorders.
Outcome	A change in behavior, physiology, attitudes, or knowledge that can be quantified using standardized scales or assessment tools. In the context of NREPP, outcomes refer to measurable changes in the health of an individual or group of people that are attributable to the intervention.
Quality of Research	One of the two main categories of NREPP ratings. Quality of Research (QOR) is how NREPP quantifies the strength of evidence supporting the results or outcomes of the intervention. Each outcome is rated separately. This is because interventions may target multiple outcomes, and the evidence supporting the different outcomes may vary. These QOR ratings are followed by brief "Strengths and Weaknesses" statements where reviewers comment on the studies and materials they reviewed and explain what factors may have contributed to high or low ratings. For more information on the scientific reviewers who rate QOR and how ratings are derived, see the NREPP page on Review Process.
Reliability	Consistency in measurement, or the ability to get the same result from an instrument or experiment over multiple trials; often expressed in terms of test/retest reliability (the ability to get the same results in a repeated test) and internal consistency (evaluating an instrument to see if multiple items are measuring the same concepts in the same way). Reliability is one of the six NREPP criteria used to rate Quality of Research.
Research designs	Research designs can be divided into three broad categories: experimental, quasi-experimental, and preexperimental. Experimental designs are generally considered the most rigorous because they use control groups, which improve one's ability to draw strong conclusions about the effectiveness of the intervention. This category includes randomized controlled trials (RCTs). Quasi-experimental designs provide strong but more limited scientific evidence compared with experimental designs. These designs are often the best option when it is impossible or very difficult to obtain a control group (as with some community and workplace interventions, for example). Preexperimental designs include simple observational and case studies; they provide the most limited scientific evidence of the three categories of research designs.
Validity	The degree to which a given experiment or study actually measures what it claims to measure. There are many types of validity. Three of the most commonly described types of validity are internal validity (the extent to which the experimental techniques used are good enough to support a cause-effect relationship), external validity (how well outside variables affecting the results are controlled), and construct validity (whether the measurement tools or techniques capture or approximate the theoretical ideal of what is being measured).