Theory testing and evaluation

What's a *good* theory?

Validation, verification, testing, evaluation

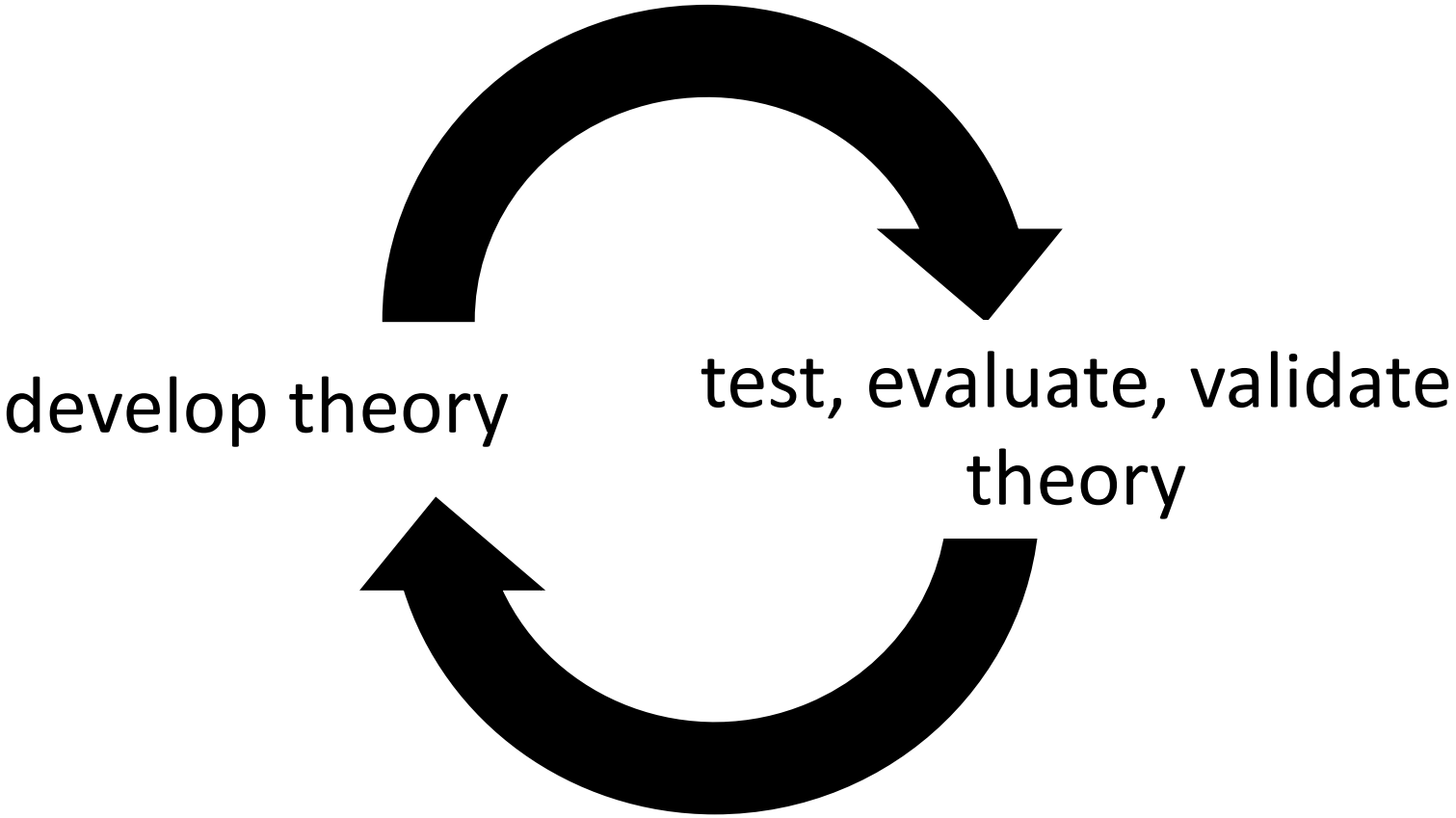
What's a *good* theory?

- Criteria:
  - Importance. A theory should not be limited to a few situations; rather, it should have relevance to real-world situations. Importance may be difficult to evaluate, as acceptance by professionals or recognition and persistence in the literature may be the only real indication of importance. Also, if a theory meets the other following criteria, then it is probably important as well.
  - Preciseness and Clarity. A theory should be understandable, internally consistent, and free of ambiguities.
  - Parsimony and Simplicity. Parsimony has long been considered an important criterion for theory. This means that the theory has a minimum of complexity and few assumptions.
  - …

What's a *good* theory?

- Comprehensiveness. A theory should be complete, covering the area of interest and including all known data in the field.
- Operationality. A theory should be capable of being reduced to procedures for testing its propositions or predictions.
- Empirical Validity or Verifiability. A theory must be supported by experience or experiments that confirm it.
- Fruitfulness. The capacity of a theory to lead to predictions that can be tested, leading to the development of new knowledge, has often been referred to as its fruitfulness.
- Practicality. Whether the theory is useful to practitioners in organizing their thinking and practice by providing an organizing framework for practice.

The PhD project redefined



Theory evaluation criteria

- Einstein
  - external validation
  - inner perfection
- Bacharach
  - falsifiability
  - utility

Such criteria as "naturalness" or "logical simplicity" or "economy" or "unity and parsimony" are not easy to defend or even specify, for their exact formulation ... meets with great difficulties; that formulation requires from us not a mere enumeration of logically independent premises, but a kind of reciprocal weighing of incommensurable qualities, hence a judgment into which aesthetic considerations and other preferences can enter prominently.

Holton, G. (2001). Constructing a Theory: Einstein's Model. *The American Scholar*.

Theory evaluation framework

Bacharach, B. (2011). Organizational Theories: Some Criteria for Evaluation. *Academy of Management Review, 14*, 496-515.

	Falsifiability	Utility
Variables	Measurement Issues	Variable Scope
Constructs	Construct Validity	Construct Scope
Relationships	Logical Adequacy Empirical Adequacy	Explanatory Potential Predictive Adequacy

Construct validity

Cronbach, X., Meel, P. (1962). Construct Validity in Psychological Texts. *Psychological bulletin, 59*, 257-272.

1. A construct is defined implicitly by a network of associations or propositions in which it occurs. Constructs employed at different stages of research vary in definiteness.

2. Construct validation is possible only when some of the statements in the network lead to predicted relations among observables.

3. The network defining the construct, and the derivation leading to the predicted observation, must be reasonably explicit so that validating evidence may be properly interpreted.

4. Many types of evidence are relevant to construct validity, including content validity, interitem correlations, intertest correlations, test - "criterion" correlations, studies of stability over time, and stability under experimental intervention.

Construct validity, continued

5. When a predicted relation fails to occur, the fault may lie in the proposed interpretation of the test or in the network. Altering the network so that it can cope with the new observations is, in effect, redefining the construct.

6. Construct validity cannot generally be expressed in the form of a single simple coefficient. The data often permit one to establish upper and lower bounds for the proportion of test variance which can be attributed to the construct. The integration of diverse data into a proper interpretation cannot be an entirely quantitative process.

7. Constructs may vary in nature from those very close to "pure description" (involving little more than extrapolation of relations among observation- variables) to highly theoretical constructs involving hypothesized entities and processes, or making identifications with constructs of other sciences.

8. The investigation of a test's construct validity is not essentially different from the general scientific procedures for developing and confirming theories.

Theory validation forms

- External validity - correspondence with facts/external reality – most social science
- Mathematical proof – some aspects of computer science
- Logical consistency + fit with existing theories – philosophy (and its application to the disciplines we work with), some forms of social theory
- Utility testing (does it work in practice) – engineering, design science

- *(note) Popper's 'work to falsify your theory' not adopted in practice*

Relation of theory development and testing to research method