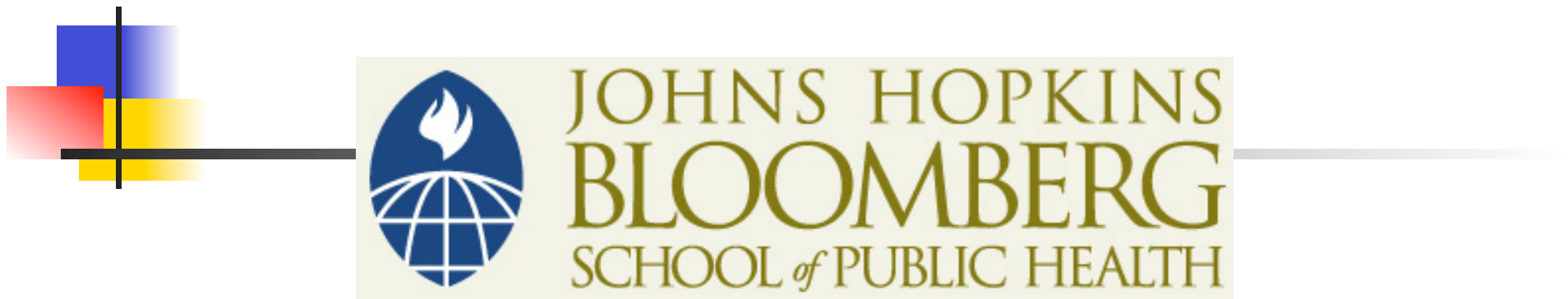


This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike License](https://creativecommons.org/licenses/by-nc-sa/4.0/). Your use of this material constitutes acceptance of that license and the conditions of use of materials on this site.



Copyright 2010, The Johns Hopkins University and Larry Wissow. All rights reserved. Use of these materials permitted only in accordance with license rights granted. Materials provided "AS IS"; no representations or warranties provided. User assumes all responsibility for use, and all liability related thereto, and must independently review all materials for accuracy and efficacy. May contain materials owned by others. User is responsible for obtaining permissions for use from third parties as needed.

ISBT2 - 2010

Overview of Program Evaluation



Larry Wissow

With thanks to Jane Bertrand

Why do evaluation?

- To determine program effectiveness
- To identify ways of improving on the existing program design
- To satisfy donor/funder requirements
- For “political” reasons, PR
- To advance science/theory
- As a step to scaling up interventions

Why do evaluation?

- As a key step in building a model for later testing (Campbell BMJ 2007)
 - Have a goal (improve quality of care)
 - Develop a model
 - What will be the indicators of greater quality
 - What will be the forces that facilitate or hinder quality improvements
 - Develop an evaluation plan
 - Do a pilot and evaluate
 - Refine the model
 - Design the larger-scale intervention

Why talk about it in this class?

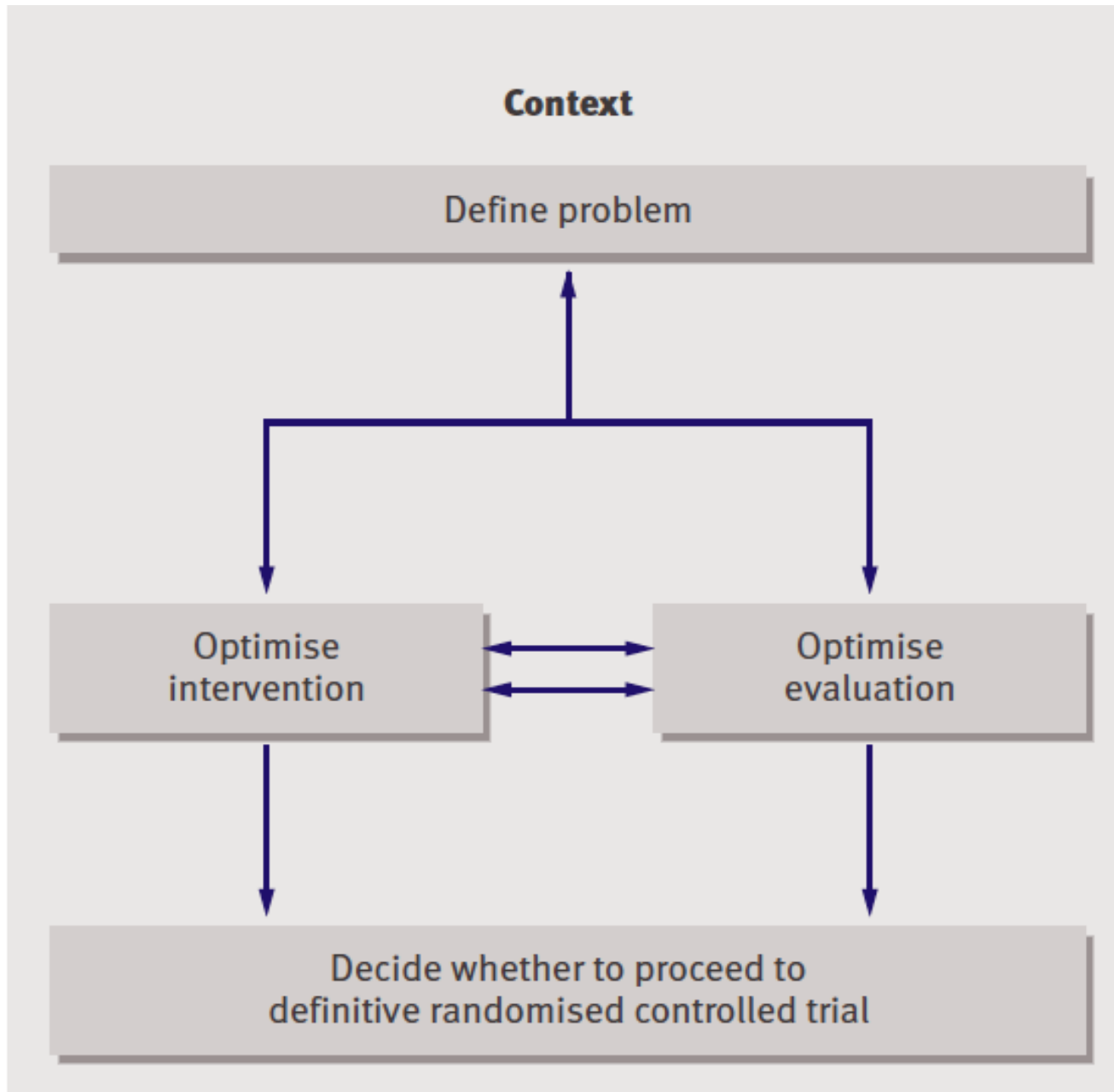
- To emphasize the links
 - Theory
 - Program structure
 - Process measurements
 - Outcome measurements
 - Refining theory
- To think about how social and behavioral sciences inform evaluation

Types of evaluation designs

- Experimental (randomized control trials)
 - “Gold standard” for assessing impact
 - Controls threat to (internal) validity
 - But weak external validity
- Quasi-experimental
 - Controls some threats to (internal) validity
 - Better external validity
- Non experimental
 - Do not control threats to validity
- Observational with statistical controls:
 - Controls some threats to validity

Limitations of the experimental design

- Difficult to use with “full-coverage” programs (no control group)
- Generalizability (external validity) is low
- Politically unpopular
- (In some cases) unethical
- Not enough knowledge yet to justify effort
- Humans are not corn plants



Conceptual frameworks

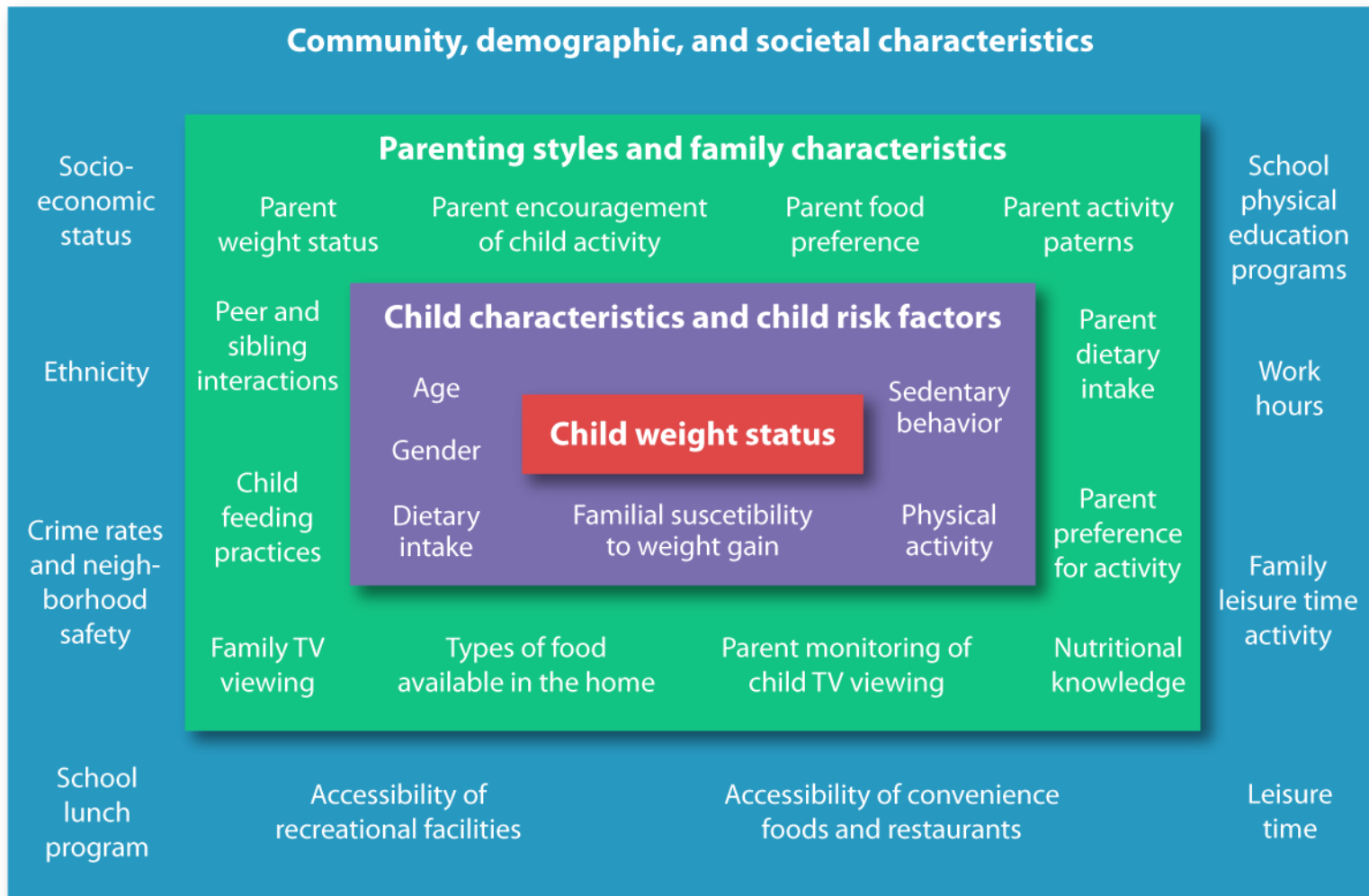
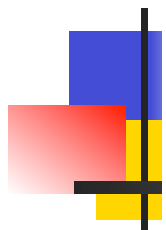
Also known as:

- Logic model
- Program model
- Outcome line
- Cause map
- Action theory
- “Pathways to...”

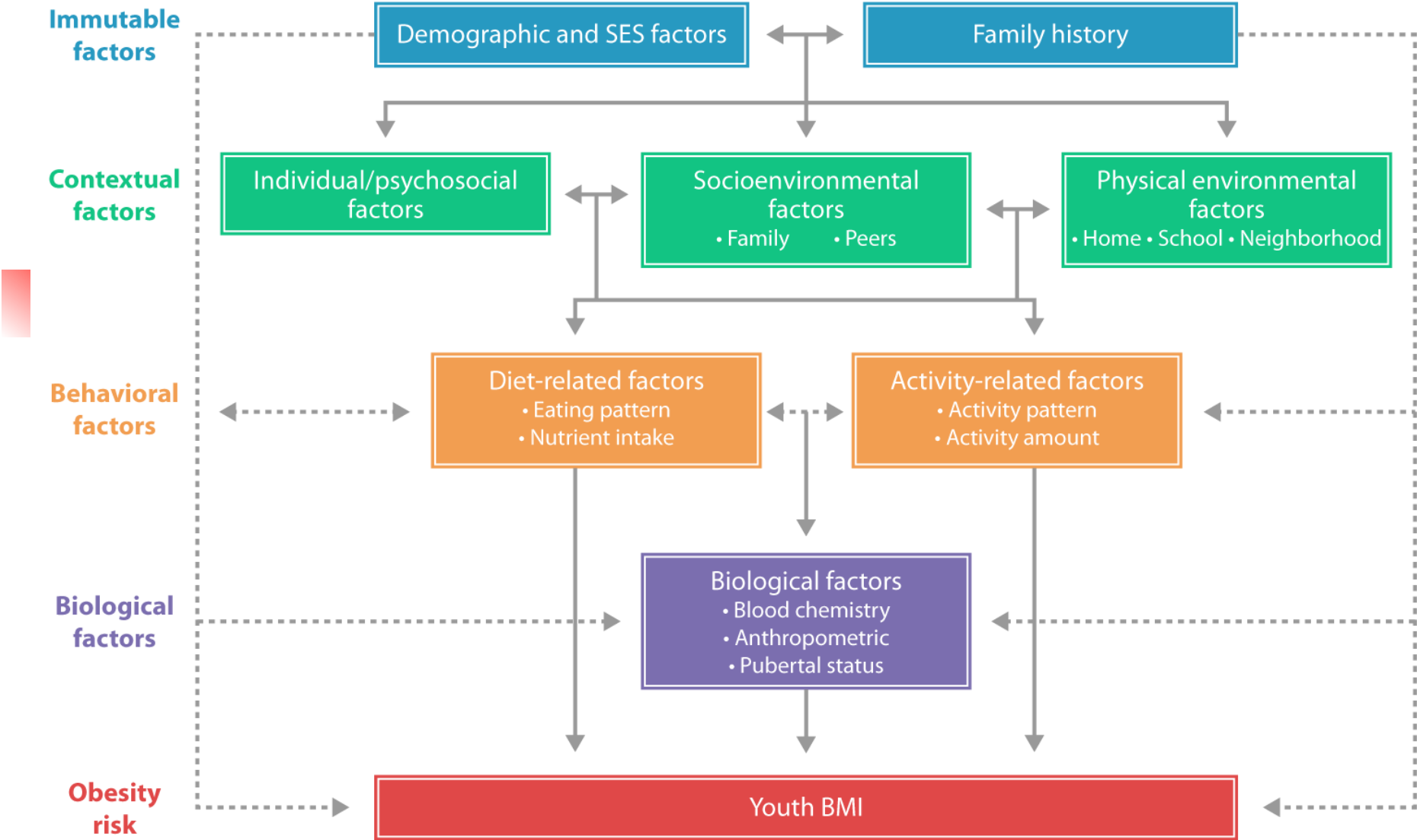
Importance of a conceptual framework (or theory)

- Articulates the pathways by which an intervention is expected to cause the desired outcomes
- Provides evaluator with specific elements to assess
- Identifies outcomes and potential measures that the program designers may not have imagined
- Identifies key explanatory variables that may explain outcomes, even if not in design

Conceptual Model: Etiology of Childhood Obesity (hard to measure)



Etiology of Childhood Obesity (ECHO) Model (easier to measure)



Ecologically-guided points for teen tobacco evaluation (vanLenthe, 2009)

Table 2: Settings for and types of health promotion in TEENAGE, including some examples of potential interventions in cells of the table.

Setting	Health education	Environmental interventions			
	Individual	Physical	Socio-cultural	Economical	Political
Individual	- mass media - brochures, leaflets - tailored interventions				
Micro					
Home		- availability cigarettes, alcohol, exercise equipment, fruit	- provision family support		- targeting parenting style
School	- school-based health education	- availability healthy and unhealthy canteen food, exercise equipment	- targeting peer pressure	- changing costs of healthy and unhealthy food - providing free breakfast in schools	- changing school policies with regard to health promotion - increasing numbers of hours of physical activity
Neighbourhood/ friends	- bill boards	- availability shops selling cigarettes, alcohol, fruit, vegetables, exercise places, parks <i>Laws and policies</i>	- peer pressures	- prices of exercise facilities	
Macro					
National		- opening hours bars		- taxation	- bans on advertisement

Formative evaluation

- Guides the design of a program or of the evaluation
- Different types:
 - Needs assessment
 - Baseline or pre-testing

Process evaluation

- Fidelity (to design):
 - Implemented according to protocol
 - Critically important and often overlooked
- Measures of program activity
 - Population reached and characteristics
 - Units of service or intervention
- Might include some “outcomes”
 - Participant or staff satisfaction
- Can include costs

Summative evaluation

- Results of the intervention - what has changed and by how much
- Ideally goes beyond change in process measures
- Some outcomes remain proxy measures
 - Hospitalizations, ED use
 - Knowledge and attitudes

Possible outcomes:

- Initial outcomes (attitudes, knowledge)
 - Useful if topic is new, difficult (e.g., meaning of continuity of care)
- Intermediate outcomes (behaviors)
 - Focus of much evaluation
- Long-term outcomes (health status)
 - Important but they change slowly; other factors contribute

How to collect the data?

- Indicators built into the process
 - Bills, clinical records, registries
- Additional assessments conducted at key time points
 - Pre-post surveys of behavior or health status
- Direct observation of processes
- Interviews with staff and participants

Significance of Design for Public Health

“Far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise” – John Tukey

(West et al. AJPH)

Significance of Design for Public Health (West et al.)

- Need to pay attention to assumptions of each design
- RCT is not best design in all cases
- Right questions
- Strongest design
- Probing assumptions of design/analysis choices

The process in the lab - 1

- Generate a model or “set of analytic concepts” from the literature
- Think about what those concepts might look like when operationalized in the setting in which you are working

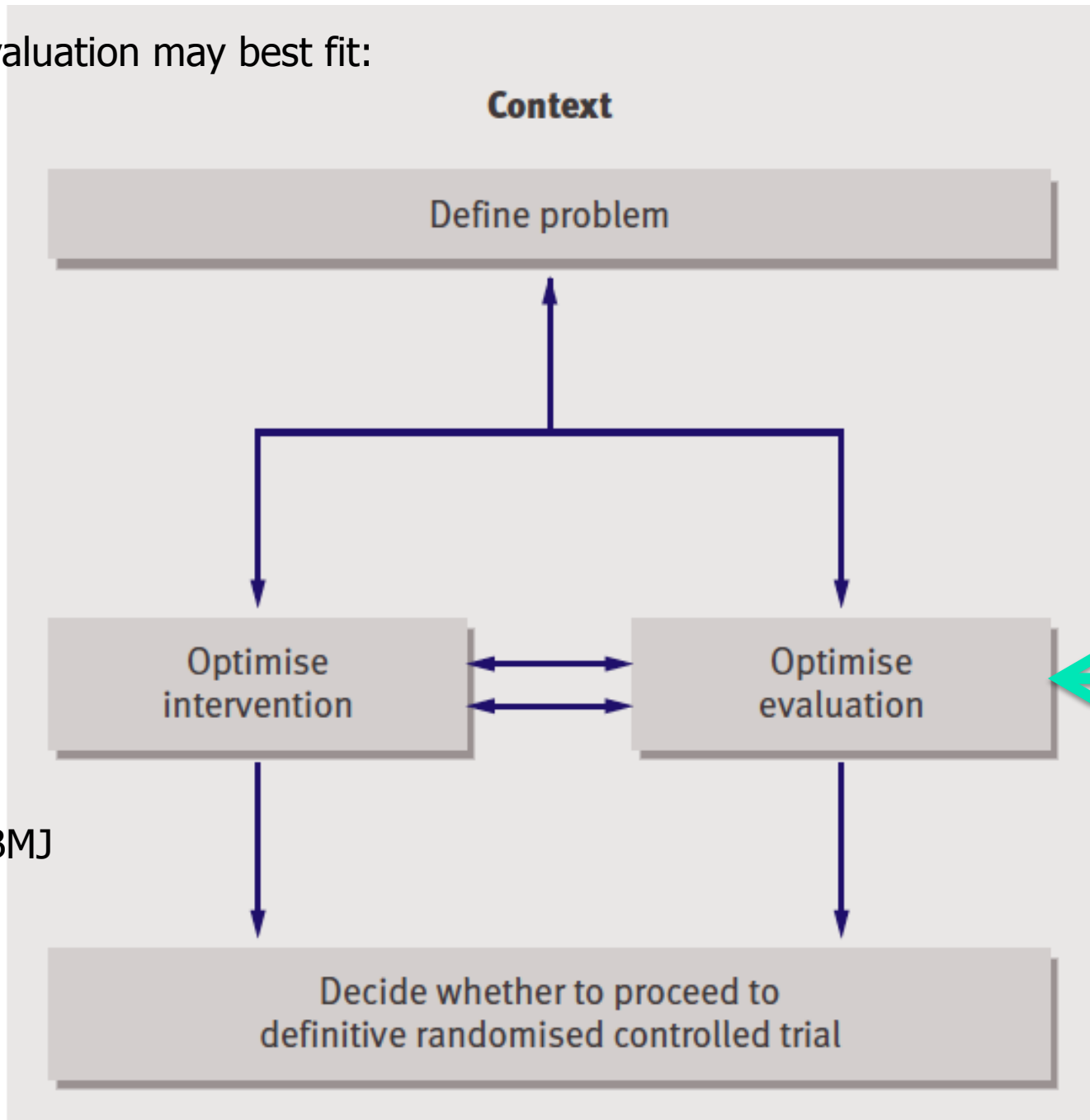
The process in the lab - 2

- Think about how to ask some open-ended questions that will get people to tell you about their behavior in a way that should generate data about the applicability of those constructs, but without leading or forcing them into giving shallow “yes” or “no” answers.

The process in the lab -3

- Analyze the answers to the questions looking for evidence that supports your model, and for evidence that might cause you to modify it (or that maybe even better supports and alternative model)

Where evaluation may best fit:



Campbell BMJ
2007