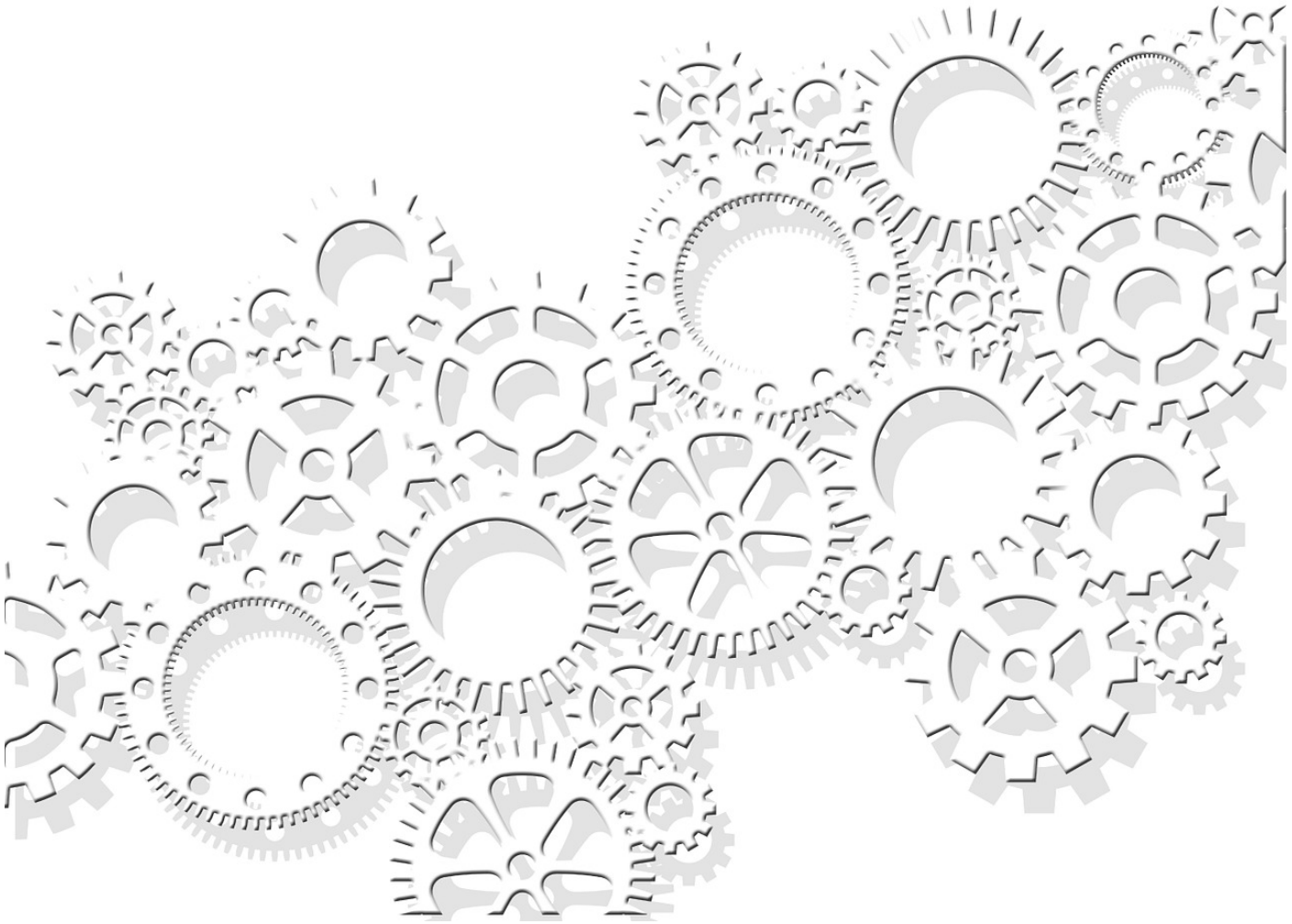
SESSION 4 APPRAISING EVIDENCE

Evidence Informed Policy Making Training

RECAP OF PREVIOUS DAY OR SESSION



SESSION 4 OBJECTIVES

At the end of this session participants will:

- Identify characteristics of basic research designs and methods
- Describe the types of evidence generated from different designs
- Describe how characteristics of critical thinking apply to assessing quality of evidence
- Know characteristics and questions to use for appraising the strength of a research paper/article – and a body of evidence.
- Demonstrate assessing levels and measures of strength of evidence for their policy issue.

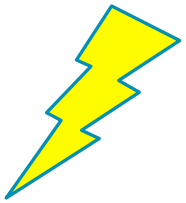
RESEARCH DESIGNS PRIMER

BUT FIRST…

DEFINITIONS: AT LIGHTENING SPEED

Research is…

- Process to discover new knowledge
- A systematic investigation
- Designed to produce generalizable knowledge



Systematic is…

- Done or acting according to a fixed plan or system; methodical

Generalizable is…

- Applied to other populations
- Published and disseminated

THE SCIENTIFIC METHOD

EXAMPLE: SCIENTIFIC METHOD AND SMOKING

Observation

- A lot of the people dying of lung cancer were smokers

Hypothesis

- People who smoke are more likely to get lung cancer than people who don't smoke

Experiment

- Follow group of smokers to see how many get lung cancer. Follow group of non-smokers to see how many get lung cancer. Compare lung cancer rates between smokers and non-smokers.
- Did the results support the hypothesis?

WHY DO RESEAR

https://pixabay.com/static/uploads/photo/2015/04/04/19/22/question-mark-706906_960_720.jpg

- To find the truth (or get closer); expand knowledge

- …and to get at the truth, the research has to be designed in a certain way
- The research design is part of the protocol
- The protocol is the set of rules/activities to be followed

RESEARCH DESI

What are they and why important?

Source: http://www.bartaste.com/wp-content/uploads/2012/05/Square-hole-round-peg-web-l.png

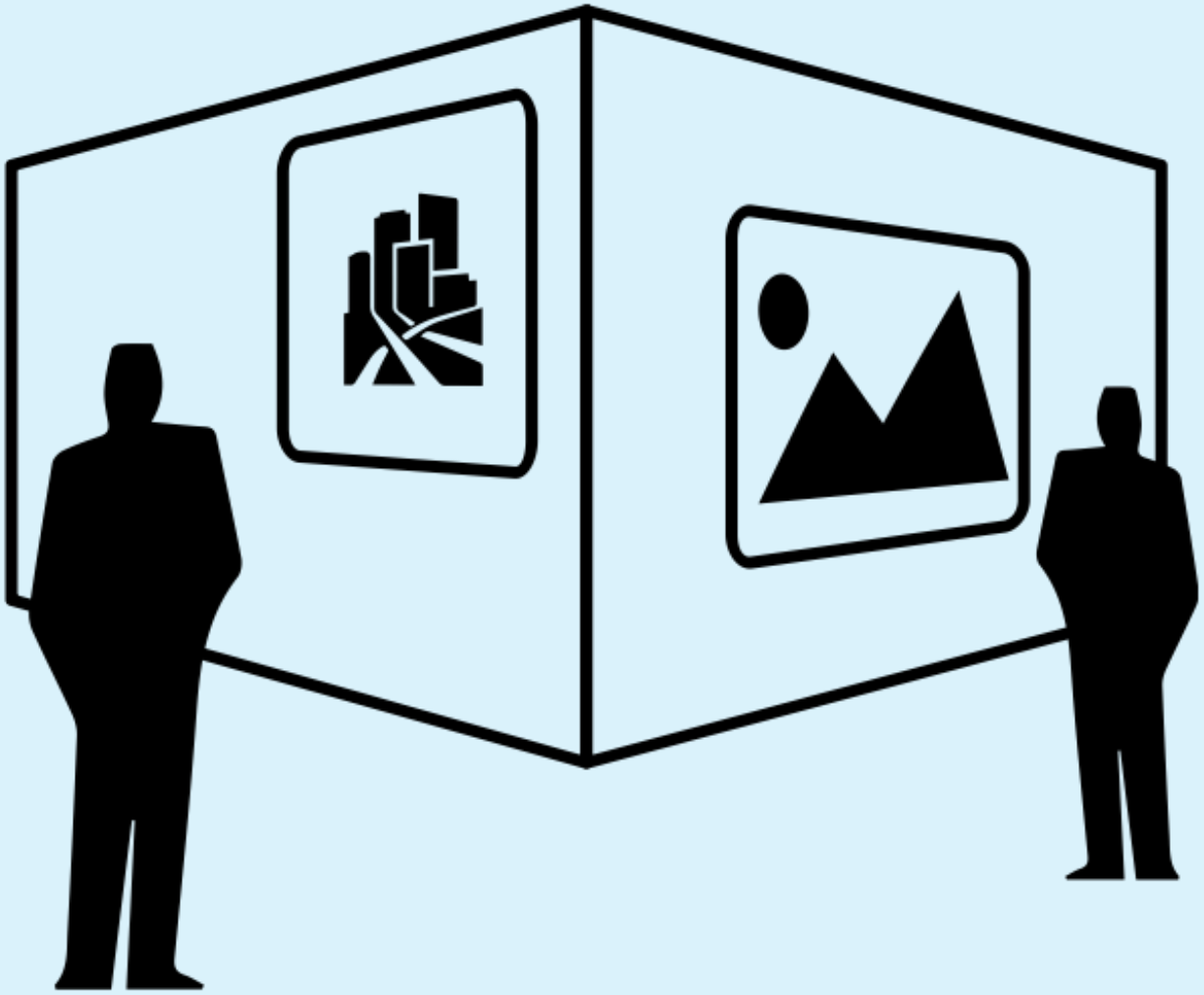
MAJOR RESEARCH DESIGNS

1. Action Research Design
2. Case Study Design
3. Causal Design
4. Cohort Design
5. Cross-Sectional Design
6. Descriptive Design
7. Experimental Design
8. Exploratory Design
9. Historical Design
10. Longitudinal Design
11. Meta-Analysis Design
12. Observational Design

Detail in pre-reading and Participant Guide

GROUP ACTIVITY

GALLERY WALK



Created by Vijay Ragavan
from Noun Project

TYPES OF EVIDENCE

Primary research studies empirically observe a phenomenon first hand. Typically:

- Experimental - Quasi-experimental - Observational

Secondary review studies re-examine primary studies. Typically:

- Systematic reviews - Non-systematic reviews

Theoretical or conceptual studies focus almost exclusively on the construction of new theories versus generating or synthesizing evidence

QUALITATIVE RESEARCH

Qualitative research:

- Gathers understanding of human behavior & reasons for such behavior
- Investigates the 'why & how' of decision-making, not just 'what, when & where'
- Highly useful in policy & evaluation studies

Qualitative data:

- Text-based
- Derived from in-depth interviews, observations, analysis of written documents, FGDs, or open-ended questionnaires

Adapted from Australian Bureau of Statistics (2010). *Understanding statistical concepts and terminology.*

QUANTITATIVE RESEARCH

Quantitative research:

- Systematic scientific investigation of quantitative properties, phenomena & their relationships
- Objective is to develop & employ statistical models, theories and/or hypotheses pertaining to phenomena & relationships

Quantitative data:

- Numerical data that can be manipulated using mathematical procedures to produce statistics

The process of measurement is central to quantitative research because it provides the fundamental connection between empirical observation & statistical expression of quantitative relationships

Adapted from Australian Bureau of Statistics (2010). *Understanding statistical concepts and terminology.*

GROUP DISCUSSION

CRITICAL THINKING

What is it?

Characteristics of critical thinkers?

How does it relate to my work? To appraising evidence?

ASSESSING STRENGTH OF EVIDENCE

1. Single study
2. Bodies of evidence

Evidence-informed policy is not just about getting research used, but getting 'good' research used

Scenario: You have an article/report from a new study in front of you. What is your thought process for deciding whether to read it and take it seriously? What questions do you ask yourself to make a determination?

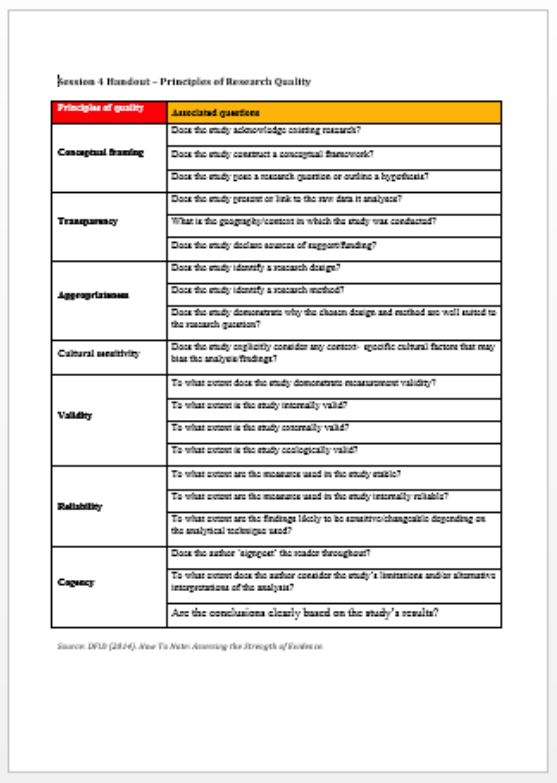
10 QUESTIONS FOR CRITICALLY APPRAISING RESEARCH ARTICLE

1. Is the study question relevant?
2. Does the study add anything new?
3. What type of research question is being asked?
4. Was the study design appropriate for the research question?
5. Did the study methods address the most important potential sources of bias?
6. Was the study performed according to the original protocol?
7. Does the study test a stated hypothesis?
8. Were the statistical analysis performed correctly?
9. Do the data justify the conclusions?
10. Are there any conflicts of interest?

Source: Young and Solomon (2009). *How to critically appraise an article. https://www.researchgate.net/profile/Michael_Solomon2/publication/23801220_How_to_critically_appraise_an_article/links/5567adb508aeccd777378c24.pdf*

GROUP DISCUSSION

CHECKLIST OF PRINCIPLES OF RESEARCH QUALITY – SINGLE STUDY



Session 4 Handout – Principles of Research Quality

Principles of quality	Associated questions
Conceptual framing	Does the study acknowledge existing research?
	Does the study construct a conceptual framework?
	Does the study pose a research question or outline a hypothesis?
Transparency	Does the study present or link to the raw data it analyses?
	What is the geography/context in which the study was conducted?
	Does the study declare sources of support/funding?
Appropriateness	Does the study identify a research design?
	Does the study identify a research method?
	Does the study demonstrate why the chosen design and method are well suited to the research question?
Cultural sensitivity	Does the study explicitly consider any context- specific cultural factors that may bias the analysis/findings?
Validity	To what extent does the study demonstrate measurement validity?
	To what extent is the study internally valid?
	To what extent is the study externally valid?
	To what extent is the study ecologically valid?
Reliability	To what extent are the measures used in the study stable?
	To what extent are the measures used in the study internally reliable?
	To what extent are the findings likely to be sensitive/changeable depending on the analytical technique used?
Cogency	Does the author 'signpost' the reader throughout?
	To what extent does the author consider the study's limitations and/or alternative interpretations of the analysis?
	Are the conclusions clearly based on the study's results?

Source: DFID (2014). How To Note: Assessing the Strength of Evidence

RESEARCH ARTICLE Open Access

Facility-level intervention to improve attendance and adherence among patients on anti-retroviral treatment in Kenya—a quasi-experimental study using time series analysis

See Participant's Guide

TWENTY TIPS FOR INTERPRETING SCIENTIFIC CLAIMS

COMMENT

Twenty tips for interpreting scientific claims

This list will help non-scientists to interrogate advisers and to grasp the limitations of evidence, say **William J. Sutherland, David Spiegelhalter** and **Mark A. Burgman.**

Calls for the closer integration of science in political decision-making have been commonplace for decades. However, there are serious problems in the application of science to policy — from energy to health and environment to education.

One suggestion to improve matters is to encourage more scientists to get involved in politics. Although laudable, it is unrealistic to expect substantially increased political involvement from scientists. Another proposal is to expand the role of chief scientific advisers[1], increasing their number, availability and participation in political processes. Neither approach deals with the core problem of scientific ignorance among many who vote in parliaments.

Perhaps we could teach science to politicians? It is an attractive idea, but which busy politician has sufficient time? In practice, policy-makers almost never read scientific papers or books. The research relevant to the topic of the day — for example, mitochondrial replacement, bovine tuberculosis or nuclear-waste disposal — is interpreted for them by advisers or external advocates. And there is rarely, if ever, a beautifully designed double-blind, randomized, replicated, controlled experiment with a large sample size and unambiguous conclusion that tackles the exact policy issue.

In this context, we suggest that the immediate priority is to improve policy-makers' understanding of the imperfect nature of science. The essential skills are to be able to intelligently interrogate experts and advisers, and to understand the quality, limitations and biases of evidence. We term these interpretive scientific skills. These skills are more accessible than those required to understand the fundamental science itself, and can form part of the broad skill set of most politicians.

To this end, we suggest 20 concepts that should be part of the education of civil servants, politicians, policy advisers and journalists — and anyone else who may have to interact with science or scientists. Politicians with a healthy scepticism of scientific advocates might simply prefer to arm themselves with this critical set of knowledge.

We are not so naive as to believe that improved policy decisions will automatically follow. We are fully aware that scientific judgement itself is value-laden, and that bias and context are integral to how data are collected and interpreted. What we offer is a simple list of ideas that could help decision-makers to parse how evidence can contribute to a decision, and potentially to avoid undue influence by those with vested interests. The harder part — the social acceptability of different policies — remains in the hands of politicians and the broader political process.

Of course, others will have slightly different lists. Our point is that a wider

See Participant's Guide

ASSESSING STRENGTH OF EVIDENCE

Weigh the rigor of the evidence you found.

Ask:

- What makes the study important?
- Do the findings make sense?
- Who conducted the research and wrote the report?
- Who published the report?
- Did the researcher select an appropriate group for study?

Source: The Guttmacher Institute (2006). Interpreting Research Studies.

ASSESSING STRENGTH OF EVIDENCE (CONT.)

- If comparison groups are used, how similar are they?
- What has changed since information was collected?
- Are the methods appropriate to the research purpose?
- Does the study establish causation?
- Is the time frame long enough to identify an impact?
- Could data be biased due to poor research design?
- Are the results statistically significant?

Source: The Guttmacher Institute (2006). *Interpreting Research Studies.*

ASSESSING CONTENT QUALITY -- IN ADDITION TO STRENGTH OF EVIDENCE

Consider:

- Completeness – missing anything?
- Uniqueness – original?
- Timeliness – up-to-date?
- Coverage – depth?

Levels of Evidence Pyramid

This evidence pyramid provides a concept of higher to lower levels of evidence.

Source: UIC Evidence Based Practice Tutorial, ebp.lib.uic.edu

MEASURES OF STRENGTH

Internal validity

- The intervention is actually causing the desired outcome. Are the changes observed due to the intervention or due to other possible factors?
- How confident we are that the observed changes are due to the intervention
- Ability to rule-out competing explanations for observed changes

External validity

- The program is replicable, producing similar results in different settings

Program fidelity

- How well a program is implemented according to established standards. Research on implementation of evidence-based programs shows that fidelity to core program elements is critical to success.

P-VALUES

A p-value tells you if the relationship is strong enough to pay attention to.

P-values represent how likely the result would occur by chance.

Used to determine whether observed differences between experiment and control groups are due to systematic effects of treatments or simply to chance factors.

Look for p-values lower than .05, or 5%, when reading journal papers.

Adapted from: Lovestats (2011). Really simple statistics: p values #MRX. Greene, L. (2008). A Brief Explanation of Statistical Significance and P Values. The Guttmacher Institute (2006). Interpreting Research Studies.

APPRAISING QUALITY OF NON-SCIENTIFIC INFORMATION

This type is still important – even though it is not gathered through a scientific process with conceptual and analytical framework, research design, methods, etc.

Examples: newspaper articles, blogs, reports of commissions, government policy documents, or guidelines.

How do you go about appraising quality for this type of information?

See Session 4 Handout – Appraising Quality of Non-Scientific Information

GROUP ACTIVITY

EVALUATING STRENGTH OF BODY OF EVIDENCE

1. Very Strong
2. Strong
3. Medium
4. Limited
5. No evidence

Evaluating the overall strength of a body of evidence

Categories of evidence	Quality + size + consistency + context	Typical features of the body of evidence	What it means for a proposed intervention
Very Strong	High quality body of evidence, large in size, consistent, and contextually relevant.	Research questions aimed at isolating cause and effect (i.e. what is happening) are answered using **high quality experimental and quasi- experimental research designs**, sufficient in number to have resulted in production of a systematic review or meta-analysis. Research questions aimed at exploring meaning (i.e. why and how something is happening) are considered through an **array of structured qualitative observational research methods** directly addressing contextual issues.	We are very confident that the intervention does or does not have the effect anticipated. The body of evidence is very diverse and highly credible, with the findings convincing and stable.
Strong	High quality body of evidence, large or medium in size, highly or moderately consistent, and contextually relevant.	Research questions aimed at isolating cause and effect (i.e. what is happening) are answered using **high quality quasi-experimental research designs and/or quantitative observational studies**. They are sufficient in number to have resulted in the production of a systematic review or meta-analysis. Research questions aimed at exploring meaning (i.e. why and how something is happening) are considered through **an array of structured qualitative observational research methods** directly addressing contextual issues.	We are confident that the intervention does or does not have the effect anticipated. The body of evidence is diverse and credible, with the findings convincing and stable.
Medium	Moderate quality studies, medium size evidence body, moderate level of consistency. Studies may or may not be contextually relevant.	Research questions aimed at isolating cause and effect (i.e. what is happening) are answered using **moderate to high- quality quantitative observational designs**. Research questions aimed at exploring meaning (i.e. why and how something is happening) are considered through a **restricted range of qualitative observational research methods** addressing contextual issues.	We believe that the intervention may or may not have the effect anticipated. The body of evidence displays some significant shortcomings. There are reasons to think that contextual differences may unpredictably and substantially affect intervention outcomes.
Limited	Moderate- to-low quality studies, medium size	Research questions aimed at isolating cause and effect (i.e. what is happening) are answered using **moderate to low- quality quantitative observational studies**. Research	We believe that the intervention may or may not have the effect anticipated. The body of

See handout: Evaluating the overall strength of a body of evidence in Participant's Guide. Source: DFID (2014). How To Note: Assessing the Strength of Evidence.

APPRAISING BODIES OF EVIDENCE

1. Summarize **technical quality** of body of evidence

   - Builds directly upon prior assessment of the quality of single research studies conducted individually or as part of a secondary study (e.g., a systematic review)

2. Assess the **overall strength** of a *body* of evidence

   - Directly linked to the quality, size, consistency and context of the collection

- If time or expertise are not available to assess all individual studies in a body of evidence…:

  -Seek to use evidence synthesis products which assess the quality of individual studies

  -Make a judgement about a body of evidence based on the criteria for strength of a body of evidence (e.g., quality, size, consistency, strength)

EIPM  SYSTEMATIC REVIEWS

Systematic reviews may be preferred in EIPM, as opposed to using single studies.

Systematic reviews sum up the best available research on a question by synthesizing results of several studies

Ideally, combine with newer or perhaps 'out-of-the box' single studies which may not have been included in a systematic review

PRACTICAL APPLICATION EXERCISE 3

Part 1

1. Assess the strength of at least one of the research documents you found for answering your policy question
2. Provide a brief, but critical summary of its strength and/or weaknesses, and indicate your decision on whether you will use the research document in your work or not [40 min]

Part 2

1. Individual feedback from facilitators [40 min]

Use Session 4 Worksheet – Appraising Evidence

SESSION 4 OBJECTIVES

At the end of this session participants will:

- Identify characteristics of basic research designs and methods
- Describe the types of evidence generated from different designs
- Know characteristics and questions to use for appraising the strength of a research paper/article – and a body of evidence.
- Demonstrate assessing levels and measures of strength of evidence for their policy issue.
- Describe how characteristics of critical thinking apply to assessing quality of evidence

SESSION REFLECTION AND EVALUATION

Source: https://pixabay.com/en/stones-stacked-balance-842731/

EXTRA SLIDES

REFLECTION

What did you learn that you can use in your work place?

What would you share in a debrief at your work place?

Are there sub-topics from the session you want to explore more?

What ideas did this session generate for you?

Are there tasks or “to-do’s” you want to follow up on later?

Are there topics or areas you want to clarify with the facilitator or group?