HUMANITARIAN ACADEMY · HUMANITAIRES JOBS

Humanitarian Mini MBA: leading a humanitarian organisation

Course handout · Week 4 : Strategic MEAL: accountability, evidence and decisions

A leadership programme of exactly 3 months (12 weeks, 144 hours): leadership, strategy, project cycle, MEAL, finance and donor budgets, compliance and fraud, resource mobilisation, supply chain, HR and safeguarding, risk and security, negotiation and advocacy, change and governance. Corrected numerical exercises, three demanding monthly exams with case studies, and a verifiable diploma from the Humanitarian Academy.

https://humanitairejobs.org/en/academy/mini-mba-humanitaire/semaine-4/

Free document from the Humanitarian Academy. It may be printed and shared freely for training purposes, without modification.

Week 4 : Strategic MEAL: accountability, evidence and decisions

Objectives

- Design a proportionate MEAL system that serves management decisions, accountability to people and donor requirements.
- Choose rigorous indicators and methods (sampling, attribution, contribution) and interpret results correctly.
- Put accountability to affected people at the centre (CHS) and turn data into organisational learning.

MEAL as a decision system

MEAL (monitoring, evaluation, accountability, learning) is not a reporting function. For a senior manager, it answers three questions: **are we doing what we promised?** (monitoring), **is it really changing people's situation?** (evaluation), **do people have a say and are they safe?** (accountability). Learning links all three to decisions.

A proportionate MEAL system generally devotes 3 to 10% of a project's budget to monitoring and evaluation, depending on its size and complexity. It is formalised in a **MEAL plan**: an indicator tracking table (definition, method, frequency, person responsible, disaggregation), a survey calendar, feedback mechanisms, an evaluation plan and data protection rules.

> Relevance test: for every piece of data collected, you must be able to name the decision it informs and the person who will make it. Data that informs no decision costs the teams and the people interviewed time, and increases data protection risks.

Indicators, sampling and data quality

A good indicator is **valid** (it measures what it claims to measure), **reliable** (two enumerators get the same result), **sensitive** (it changes when the situation changes), **disaggregated** (sex, age, disability using the Washington Group questions) and **realistic to collect**. For food security, standardised indicators (food consumption score, reduced coping strategies index) are preferred so that results can be compared with other actors.

For a simple random sample survey, the sample size is calculated as $n = z^2 \times p \times (1 - p) / e^2$. With a 95% confidence level ($z = 1.96$), an unknown prevalence ($p = 0.5$) and a 5% margin of error ($e = 0.05$): $n = 3.8416 \times 0.25 / 0.0025 \approx 384$. With cluster sampling, multiply by the **design effect** (often 1.5 to 2); then add a margin for non-response (for example 10%). With a design effect of 2 and 10% non-response: $384 \times 2 / 0.9 \approx 854$ households.

Data quality is controlled at every step: training and testing enumerators, a questionnaire piloted in the local language, automatic checks in the tool (KoboToolbox, ODK), re-checking a

sample of interviews (back-checks), daily analysis of outliers and of abnormally short interview durations.

Evaluating: criteria, attribution and contribution

Evaluations most often use the six revised **OECD DAC** criteria (2019): relevance, coherence, effectiveness, efficiency, impact and sustainability. For humanitarian action, ALNAP proposes complementary criteria such as coverage, connectedness and coordination.

The hardest question is **causality**: is the observed change due to the project? Three families of approaches:

- **Experimental approaches** (randomised trials): they establish attribution with the greatest rigour, but are expensive, raise ethical issues in emergencies (you do not deny aid to a control group in need) and do not always explain why something works.
- **Quasi-experimental approaches** (difference-in-differences, propensity score matching, regression discontinuity): they use natural comparisons, for example households just above and just below an eligibility threshold.
- **Theory-based approaches** (contribution analysis, process tracing, outcome harvesting): they check whether the planned causal chain actually happened and rule out alternative explanations.

Worked difference-in-differences example. The average food consumption score rises from 28 to 41 among beneficiary households, and from 30 to 36 in a comparison group. The estimated effect is $(41 - 28) - (36 - 30) = 13 - 6 = 7$ points, not 13: part of the improvement is explained by the harvest season.

Accountability to people and the Core Humanitarian Standard

The **Core Humanitarian Standard** (CHS), revised in 2024, sets nine commitments to affected people and communities: assistance that is appropriate, effective and timely, that strengthens local capacities and avoids negative effects, based on communication, participation and feedback, with complaints handled, coordination, learning, competent and well-managed staff, and resources managed responsibly.

An effective **complaints and feedback mechanism** offers several channels (hotline, boxes, help desks, committees, digital channels), adapted to people who cannot read, older people and people with disabilities; it separates sensitive complaints (sexual exploitation and abuse, fraud, corruption), which are handled confidentially through a dedicated channel; and it **closes the loop** by responding to people within an announced timeframe.

Useful management indicators for a senior leader: rate of responses within the deadline, share of complaints that led to a programme change, share of feedback coming from women and people with disabilities, and time taken to handle sensitive complaints.

> Warning sign: a mechanism that receives almost no complaints is not proof of a perfect programme. Most often it is proof that people do not know about it, cannot access it or do not trust it.

From data to organisational learning

Learning happens when evidence changes decisions. The most effective tools are simple: **quarterly progress reviews** centred on three questions (what works, what does not work, what are we changing?), **after-action reviews** at the end of each emergency phase, a **management response** to the recommendations of every evaluation (accepted, partially accepted, rejected, with an action plan and a person responsible), and follow-up of those plans.

Adaptive management formalises this cycle: explicit assumptions, leading indicators, scheduled decision points, and room for manoeuvre negotiated with the donor to change activities without a heavy amendment. Several donors explicitly encourage it in unstable contexts.

Finally, **data protection** is a leadership obligation: minimising the data collected, informed consent, encryption, restricted access, retention periods, and an impact assessment before sharing beneficiary lists with third parties (authorities, payment providers). The ICRC Handbook on Data Protection in Humanitarian Action (2nd edition, 2020) is the sector's reference.

Exercise: design and critique a MEAL plan

Situation. A 9-month cash transfer project for 3,000 households in 12 villages. The donor asks you to demonstrate the effect on food security. Available MEAL budget: USD 45,000. A second area, not covered for lack of funding, has similar characteristics.

Task. 1) Three standardised outcome indicators with their method. 2) Calculate the sample size of a cluster survey (95% confidence, p = 0.5, 5% margin, design effect 1.5, 10% non-response). 3) Choose a method to estimate the effect and justify it, including from an ethical standpoint. 4) Describe the complaints mechanism, including the channel for sensitive complaints. 5) Write two quarterly review questions and the decision each one can trigger.

Answer key. Sample: 384 × 1.5 = 576; 576 / 0.9 = 640 households. The uncovered area allows a difference-in-differences design: before-and-after comparison in both areas, without denying aid to anyone since the comparison area is not funded (you should nevertheless advocate for its coverage and share the data with the cluster). Possible indicators: food consumption score, reduced coping strategies index, share of expenditure on food.

Key takeaways

- Every piece of data collected must inform an identified decision.
- $n = z^2 \times p \times (1 - p) / e^2$; then × design effect; then ÷ (1 − non-response rate).
- OECD DAC 2019 criteria: relevance, coherence, effectiveness, efficiency, impact, sustainability.
- Difference-in-differences: subtract the change observed in the comparison group.
- CHS 2024: nine commitments; a complaints mechanism with no complaints is a warning sign.

Quiz

1. With z = 1.96, p = 0.5, e = 0.05, a design effect of 2 and 20% non-response, the target sample size is about:
   A. 384
   B. 768
   C. 960
   D. 461

2. Beneficiaries' average income rises from 50 to 80; a comparison group's from 52 to 70. The effect estimated by difference-in-differences is:
   A. 30
   B. 18
   C. 12
   D. 28

3. The "coherence" criterion added by the OECD DAC in 2019 concerns:
   A. The consistency of the budget with the logframe
   B. The compatibility of the intervention with other interventions and policies in the same context
   C. The quality of the report's writing
   D. The consistency of the indicators with one another

4. A complaint of sexual exploitation arrives through a village suggestion box. It must be:
   A. Read out at a community meeting for transparency
   B. Handled by the programme team like other complaints
   C. Passed on confidentially to the dedicated channel (PSEA focal point), protecting the person, following a survivor-centred approach
   D. Ignored if it is anonymous

5. Why is a randomised trial often unsuitable during an emergency?
   A. Because it is always imprecise
   B. Because it raises an ethical problem if aid is denied to people in need to form a control group, and it is slow and costly
   C. Because donors forbid it
   D. Because it only measures activities

Quiz answers : 1. C · 2. C · 3. B · 4. C · 5. B