4.1 Evaluating research

Name: _____ Class: _____ Date: _____ **Total: 24 marks**

KEY WORDS credibility • thematic analysis

Evaluating research

Objective

Distinguish the paradigms and their quality criteria, apply qualitative sampling and analysis with trustworthiness checks, and answer Paper-3 style scenario questions with stimulus-anchored evaluation (HL research extension).

1 Key knowledge and worked examples

HL evaluation toolkit. Quantitative tests hypotheses (reliability, validity); qualitative explores meanings (credibility, transferability, dependability, confirmability). Qualitative sampling: purposive, snowball, convenience – saturation is the stopping rule, not size. Thematic analysis: transcribe, code, theme, check against the whole set and disconfirming cases. Content analysis counts, with inter-rater reliability: 16/20 agreements = 80%. Bias family: observer effect, interviewer bias, expectancy – answered by reflexivity and triangulation. Paper 3: Q1 read the study; Q2 discuss method; Q3 evaluate and modify – every point anchors in the stimulus.

2 Practice

2.1 State one research question suited to a quantitative design and one suited to qualitative interviews. [2]

2.2 Distinguish purposive sampling from snowball sampling, naming one research group for which snowball is the realistic choice. [2]

2.3 Two raters independently code the same 20 statements and agree on 16. Calculate the agreement rate and state one way to improve it. [3]

2.4 Name the four trustworthiness criteria of qualitative research and pair each with the practice that serves it. [4]

3 Exam-style questions

3.1 *(Paper 3, Q1 style.)* In the scenario: six semi-structured interviews with volunteers from one year group, analysed thematically into three themes. Identify the method, the sampling method and the type of data collected. [3]

3.2 *(Paper 3, Q2 style.)* Discuss one methodological strength and one methodological limitation of the scenario's design, each with its consequence. [4]

3.3 *(Paper 3, Q3 style.)* Evaluate the scenario's credibility, transferability and ethics together, and recommend ONE modification, naming its trade-off. [6]

Solutions

2.1 Quantitative: does a semantic-processing condition produce higher recall than an acoustic condition? (a difference to be tested – accept any measurable effect) (1). Qualitative: how do students experience and explain their reluctance to ask questions in class? (a meaning to be explored) (1).

2.2 Purposive sampling deliberately selects information-rich cases chosen for the question (1); snowball sampling recruits each participant through the previous one's network. Snowball is the realistic choice for hard-to-reach or stigmatised groups (e.g. homeless young people, undocumented workers), where no sampling frame exists (1).

2.3 Agreement = $16/20 \times 100 = 80\%$ (2). Improve by sharpening the coding manual with worked examples and boundary rules, retraining the raters on a practice set, and resolving disagreements by discussion with the rule change documented (any one) (1).

2.4 Credibility – triangulation and member checking (1); transferability – thick description of the context (1); dependability – an audit trail of analytic decisions (1); confirmability – reflexivity: the researcher's role and assumptions declared (1).

3.1 Method: semi-structured interviews (qualitative) (1). Sampling: opportunity sampling of volunteers from one year group (1). Data: qualitative – participants' own accounts, transcribed and thematically coded (1).

3.2 Strength: semi-structured interviews can pursue unanticipated answers, so the mechanism behind question-avoidance can emerge – well-suited to a "why" question (2). Limitation: opportunity-sampled volunteers from ONE year group limit transferability – schools differ in culture and teacher norms, so the three themes may not travel, and the claim must be hedged accordingly (2).

3.3 Credibility: served by direct extracts and a describable coding process; weakened by single-researcher coding and no member checking (1). Transferability: one school, one year group, volunteers – thick description of the setting is the honest hedge, not a representativeness claim (1). Ethics: consent and anonymised transcripts are present, but classroom anxiety is a sensitive area – the right to skip questions and distress management should be explicit in the procedure (2). Modification: add a second coder with an agreement check (toward dependability) and member checking of the themes (toward credibility) (1). Trade-off: both cost time, and member checking lets participants recognise and edit their own themes – confirmation moves a little toward rehearsal (1).

Local review guidance: the division before the percentage; each point names a stimulus feature and its consequence; (R1 3.3) the evaluation integrates all three angles before the single modification and its trade-off. Authored classroom guidance.