Online Resource 1: Behavioural marker definitions and a protocol for replication

Breaches in the Flow: A Mismatch Model of Micro-Decisions in Everyday Activity

Atrina Oraee and Saadi Lahlou

Department of Psychological and Behavioural Science, London School of Economics and Political Science, London, United Kingdom

Corresponding author: Atrina Oraee; a.oraee@lse.ac.uk

Journal: *Integrative Psychological and Behavioral Science*

R1. Purpose and evidence status

This resource makes the behavioural descriptions more explicit and specifies records needed to evaluate their reproducibility. It distinguishes **source descriptions** recovered from the doctoral thesis, **retained-data conventions** established for the article's final corpus, and **prospective rules** proposed here for future applications. The prospective rules have not been validated, and the original observations have **not** been recoded under them. All article statistics continue to use the source-coded episodes.

Three measurements must remain separate: an observer's footage-based description of action; the participant's replay classification using the supplied cards; and the participant's report of whether the episode was experienced as deciding. Contextual interpretations and later activity-model codes constitute a further analytic layer. More precise descriptions can improve transparency without turning any of these measures into a cognitive gold standard.

R2. What the original procedure established

Source description. Card terminology developed from participants' language during exploratory work. Two pilot participants each coded approximately 30 minutes of their own recordings, helping refine the procedure (Oraee, 2024, pp. 53, 62). In the main study, 51 of 60 participants completed

Reinforced Replay Interviews (RRIWs) after ordinary Replay Interviews (RIWs). The final self-coding supplied the retained RRIW classifications (p. 67).

The documented RRIW sequence was: explain the cards with examples and check comprehension; ask participants to pause playback when they experienced a break in activity fluidity; let them select the best-fitting card; and, unless already volunteered, ask **“Did you experience this instance as making a decision?”** (p. 65). Participants could control playback. The seven choices were Pause, Hesitation, Scan, Retry, Give Up, I don't know, and None of the available cards. The study intentionally did not impose an objective definition of deciding (p. 216).

The historical unit was a participant-contextualised episode. Although the thesis describes rows as experienced breaks (p. 67), it also explicitly includes open transitions into new activity after previous goals ended, where there was “no breach in the flow per se” (p. 130). Thus, source-coded episodes are broader than strictly observed interruptions of an ongoing motor action. They must not be retrospectively restricted to that narrower definition.

Development was iterative: pilot participants had difficulty separating Pause and Hesitation (p. 358), and Retry and Give Up without reported deciding prompted changes to the initial model (p. 357). These patterns informed model construction; they are not independent validation of a previously fixed account.

FH and NL independently identified and coded randomly selected participants' footage before conducting RIW/RRIW to replicate the protocol. Initial footage coding and random selection are documented in the thesis; the complete procedural sequence is confirmed by the first author. The retained paired sheets represent completed-protocol classifications, not a preserved independent detection benchmark. Appendix A of the article reports the available comparisons and their limitations.

R3. Observable cues and category boundaries

The descriptions below identify visible features that can support an observer's candidate classification. They do not determine the participant's card choice or decision response. Requirements for explicit

feature records and uncertainty are **prospective rules**; they are not claims about an already standardised historical instrument.

Pause

Source description: the current motor action stops or markedly slows, without an immediate next action; the thesis mentions freezing and distinguishes relative head stability from active scanning (p. 87).

Prospective application: identify the preceding operation and the visible change in its progression. Record interruption onset, resumption or transition, and visibility. Stillness while an operation continues, a held tool, absent hands, a camera cut, or occlusion is insufficient. Record task-imposed waiting and resting as contextual states rather than automatically treating them as deciding or excluding them from all analysis. If only slowing is visible, describe the before/after movement and mark an uncertain boundary when necessary. The thesis supplies no quantitative speed criterion.

The published thesis illustration gives conflicting five- and eight-second descriptions of the same pause (pp. 87–88); neither establishes a cutoff. The final corpus contains 30 one-second pause entries. A historical minimum of two seconds is therefore unsupported, despite a descriptive minimum reported on p. 204. Duration cannot determine decision status.

Hesitation

Source description: rapid head/hand changes and interrupted or repeated approaches occur before action continuation (pp. 88–89). The dishwasher illustration shows approaches and withdrawals toward cutlery; a later example describes holding a salt shaker, putting it down and picking it up again (p. 206).

Prospective application: record the checked, withdrawn, reversed or alternating movement and its target. An observer must be able to describe the sequence without saying that the actor looks confused. Distinguish ordinary repeated task movements from an initiated action that is suspended or redirected. A halt alone supports a Pause candidate; scene exploration alone supports a Scan candidate. Both may coexist with checked approaches, so no mandatory category hierarchy is

imposed. If an apparent withdrawal could instead be an unsuccessful completed attempt, retain the competing Hesitation/Retry descriptions until context is clarified.

Uncertainty, remembering and conflicting options are participant reports or interpretations, not visible prerequisites. The salt-shaker replay account supplies an explanation of the movement; the movement itself does not establish that explanation.

Scan

Source description: visual exploration appears as looking around or reorienting the head/view across the environment (pp. 89–90). The later methodological clarification explicitly excludes **continuous, successful visual monitoring during ongoing action** from the index Scan category, distinguishing it from exploration at a pause/breach (p. 208). The source also retains scanning at open activity transitions (p. 130).

Prospective application: describe the locations or objects brought into view, any visible inspection action, and whether the context is an interruption, open transition, or continuous execution. Exclude camera disturbance, locomotion alone, and routine tracking of a manipulated object as sufficient evidence of index Scan. Record continuous monitoring separately. Inspection using hands or an object may provide contextual evidence, but do not convert every object movement into confirmed visual exploration.

A head camera does not measure eye movements. In the washing-machine account, the participant says scanning looked like a pause (p. 157). Preserve that participant-selected Scan while marking the observer feature as unverified if no discriminating visual cue is available. Neither a stationary view nor the report alone verifies visible gaze exploration. No angle, fixation count, sweep speed or duration threshold is established.

Retry

Source description: an actor makes another attempt at substantially the same immediate operation after an earlier attempt failed to achieve the local result (pp. 90–91). Small adjustments in force or technique can occur (p. 130).

The operation need not be a blocked motor manipulation. A thesis account describes repeatedly looking for a cleaning product in the same places after an unsuccessful search, before deciding to buy it (p. 146). Thus, repeated search can receive Retry, while its visible exploratory movements can also support a Scan description. Keep the action form, repeated operation and participant category distinguishable; do not restrict Retry to physical blockage.

Prospective application: name the operation, target and visible evidence of the preceding failure or unsatisfactory result. Mark the initial attempt separately from renewed attempts. Routine wiping strokes, chopping cycles or other planned repetitions are not automatically retries. A substantially different means is a strategy change; it may remain in the same broader episode but should not be counted as another same-operation attempt. Where intended success is unclear, retain an uncertain candidate and obtain contextual information without inferring decision status.

The historical washing-machine illustration contains one initial failed attempt and two retries (three attempts total). The retained workbook's attempt-count convention includes the initial attempt. Future records should explicitly store total attempts and additional retries, avoiding blank cells with implicit numerical meanings. The historical episode convention groups repeated attempts addressing one difficulty; it does not count every movement as a new episode.

Give Up

Source description: a particular task or subgoal pursuit ceases before completion, with disengagement or movement elsewhere (pp. 91–92). The vacuuming example abandons an inaccessible area while the broader cleaning activity continues.

Prospective application: specify the local goal, evidence that it remains unresolved, and the observed cessation. Separate local discontinuation from successful completion, waiting, externally imposed interruption and the end of the recording. Report later resumption only when observed or explicitly reported. If the observation window ends before the outcome is known, mark it indeterminate; absence of return within a short clip does not establish permanent abandonment.

The same visible departure can accompany changing means, postponement or abandonment at different goal levels. Preserve that ambiguity until contextual evidence permits a narrower account. Avoidance, exhaustion and deciding to stop are explanations or reports, not defining visual features.

R4. Prospective episode and feature boundaries

The following rules are proposed for future recording and reliability assessment. Retain an **episode record** for the unresolved local task or open transition, and separate **feature records** for the observable forms within it.

1. **Describe the task chain first.** Identify the preceding operation, its object/target and any supported local goal. Record whether the candidate concerns interruption of an ongoing action, an open transition after completion, or uncertain context. Do not infer that the next action was psychologically unsettled merely because it was not yet visible.

2. **Mark observable onset.** For an interruption, use the first visible departure from the preceding operation. For an open transition, record the end of the preceding task and the first candidate orientation feature separately. If either falls outside the recording or is obscured, retain a bounded uncertainty rather than inventing a timestamp.

3. **Track the same local issue.** Keep repeated attempts and associated pause/scan features together while they address the same unresolved operation. Record changes of means and newly introduced goals. Start a new linked episode when a different task/obstacle becomes the focus; retain a split/merge query when that distinction is unclear.

4. **Record the observed endpoint.** Use resumption, resolution, departure to another task, or the observation limit, and specify which occurred. Following activity can retain the same overarching goal. A clip boundary is an observation limit, not a behavioural outcome.

5. **Keep feature intervals distinct.** Mark successive versus overlapping cues. The final corpus contains 20 episodes with two substantive markers; the categories cannot be made mutually exclusive. A Pause followed by Scan can belong to one episode without implying simultaneous physical immobility and head movement.

These action-relative rules require piloting. The historical dataset does not preserve feature-level onset/offset intervals sufficient to apply them retrospectively. Any future timing resolution, minimum-duration threshold or rule for joining nearby segments must be preregistered, justified using separate training footage, and identified as a revised instrument.

R5. Prospective workflow before and during replay

Preparation. Fix the sample of continuous footage, recording coverage, task scope and manual version before evaluation. Train on separate examples, including boundary and negative cases. Keep study identifiers pseudonymous. Decide whether the project evaluates strict action interruptions, broader orientation episodes, or both; report the strata separately.

Independent observation. Each coder reviews the same continuous footage without another coder's event list, participant replay accounts or decision labels. Log all candidate episodes, visible features, alternative classifications and observation limits. Preserve the unedited independent lists. “Not observable” is different from an observed absence; a candidate without a substantive marker is not automatically uninterrupted activity.

Replay. Follow the documented card explanation, participant-controlled review and decision question in R2. Invite the participant's own account and selection without supplying a preferred answer. Record what was actually said, the exact prompts used, spontaneous answers, requested clarifications and any prompted change. This instruction to preserve those details prospectively does not imply that every historical exchange was unprompted. The thesis includes a participant rejecting a suggested Scan label (p. 67).

Separate outcomes. Retain the observer's original feature description even if the participant selects a different card. Record explicit decision, explicit nondecision, uncertainty, contradictory answers and no usable answer separately before any resolution. A card labelled I don't know concerns marker classification and must not be equated with uncertainty about deciding. Do not infer decision status from silence, the movement form or task outcome.

Resolution and interpretation. Preserve amendments with their evidence and reason. None means an explicit lack of fit to the cards; I don't know means expressed classification uncertainty; missing means no retained classification. Keep them distinct. Later activity-model interpretations, such as Identify Possibilities, must remain separate from visible features; the thesis explicitly describes their interpretive basis (p. 93). The current corpus has final binary decision judgements; additional prospective response states do not alter those historical records.

R6. Prospective agreement and event matching

Evaluate **event detection**, **classification conditional on matched events**, and **extraction of participant reports** separately. Agreement on supplied or matched rows cannot show how often coders miss episodes in continuous footage. Protocol replication through interviews additionally concerns elicitation and interpretation; it is not reducible to static label agreement.

Before examining evaluation results, preregister the matching procedure: identical clip, compatible local action/target referent, and a stated temporal criterion. Select any tolerance on training footage in light of frame rate and segmentation uncertainty. No universal validated tolerance follows from this study. Use a reproducible one-to-one matching method, document ties and split/merge disagreements, and retain unmatched candidates. Do not force matches solely because two timestamps are close, or change the tolerance to maximise agreement.

Report independently detected counts, matched and unmatched candidates, boundary differences, split/merge frequencies and observation exclusions. For matched episodes, report marker-specific confusion and agreement with multi-label coding permitted, plus uncertainty and category prevalence. Evaluate decision-report extraction on the available recorded answers rather than treating a consensus label as cognitive ground truth. Confidence intervals should account for repeated episodes within participants where applicable.

Only after preserving these results should coders discuss disagreements or an adjudicator review them. Store independent and adjudicated records separately, with the evidence for each change. A

consensus event set is useful for analysis but does not erase detection disagreements. No new reliability estimate, sensitivity or specificity is claimed by this manual.

R7. Minimal prospective recording forms

Use linked tables rather than overwriting one label as interpretation develops. Fields below constitute a proposed schema, not a description of all surviving historical records.

Record	Minimum fields
Clip	Anonymous participant/clip ID; duration; activity; coverage/visibility limits; manual version; coder; observation or replay stage
Episode	Coder-specific episode ID; clip ID; onset/end or uncertain bounds; preceding operation; target; local goal and evidence; interruption/open-transition/uncertain context; endpoint type; linked earlier/later episodes
Observer feature	Episode ID; feature ID; candidate marker(s); visible movement description; onset/end; successive/overlapping relation; visibility; competing classification; confidence/uncertainty reason
Attempts/outcome	Episode ID; operation/target; initial and subsequent attempt boundaries; total attempts; additional retries; result/evidence; means change; local completion/discontinuation; observed or reported resumption

Record	Minimum fields
Replay response	Episode ID; interview/time reference; actual prompt; participant description; selected card(s); decision response; spontaneous/prompted status; uncertainty or contradiction; subsequent clarification
Match/adjudication	Original coder IDs; match rule/version; matched/unmatched/split/merge status; boundary difference; adjudicated value; evidence; reason; adjudicator; date

Retain timestamps at the declared precision and distinguish zero, missing, not applicable and unobservable. Preserve the original response text or a locatable recording reference. A numeric field must not silently encode a reported estimate as an independently observed count. Dataset exports should carry code definitions and manual version so that later users can reconstruct every analytic transformation.

R8. Source provenance and limits

Content	Source	Evidential use
Card development and exact RRIW sequence	Oraee (2024), pp. 53, 62–68	Historical procedure and wording
Marker descriptions and illustrated boundaries	pp. 87–92	Source-grounded behavioural cues
Retry includes renewed search in the same places	p. 146	Local operation, not only motor blockage
Interpretation of following-step codes	pp. 93–94	Separation from direct observation
Open transitions without strict interruption	p. 130	Scope of historical episodes

Content	Source	Evidential use
Reported Scan resembling Pause	p. 157	Limit of visible classification
Hesitation movement sequence	p. 206	Visible action versus replay meaning
Continuous monitoring excluded from index Scan	p. 208	Documented category boundary
No imposed objective definition of deciding	p. 216	Participant-defined outcome
Pilot category ambiguity and model revision	pp. 357–358	Iterative development, not independent validation
Twenty multi-marker episodes; thirty one-second pauses	Final retained corpus audit; Appendix A of the article	Data properties, not audiovisual recoding
Boundary logging, forms, matching and adjudication	This resource, R4–R7	New prospective rules requiring evaluation

All thesis pages are printed page numbers. Thesis illustrations and replay accounts are source-reported examples; this resource does not claim fresh verification of the underlying recordings. The more explicit protocol enables future testing of reproducibility. It does not establish that behavioural cues reveal mental mechanisms, that replay reports are error-free, or that absence of reported deciding means absence of cognitive processing.

Source: Oraee, A. (2024). *Micro decision-making in everyday life: An investigation with subjective evidence-based ethnography* [Doctoral thesis, London School of Economics and Political Science].