FOCUS_NOW!!

A Go/No-Go Attention Task, Realized in Rust — A Psychometric Design Note

Jean-François Parent · jfparent.dev@pm.me **February 2025**

Live demo (React port): https://jfparent.ca/focus_now/index.html

---

Abstract

FOCUS_NOW!! is a small Rust game in which rapidly cycling, colour-degraded images are shown under time pressure; the player presses **Space** when a memorised *target* image appears and withholds the press otherwise — including when the target reappears as a red-toned "joker". Stripped of its presentation, the game is a member of a well-studied family of psychometric tasks: the **Continuous Performance Test (CPT)** and, more specifically, the **Go/No-Go** paradigm used to measure sustained attention and response inhibition.

This paper takes the game seriously as a *measurement design*. For each core mechanic we name the psychometric construct it engages and show the exact Rust code — Bevy's Entity-Component-System (ECS) state machine, the per-trial roll table, the grace-timer response window, and the score-linked dithering — that realises it. We then read the game through the lens of **Signal Detection Theory (SDT)**, treat its adaptive difficulty as a crude ability **staircase**, and close with an honest account of what stands between this game and a *validated* instrument.

What this paper is not. FOCUS_NOW!! has not been normed, its reliability has not been estimated, and its scores have not been validated against any established measure of attention. It is a game whose mechanics *rhyme* with psychometric paradigms, not a diagnostic tool. Where we use measurement language (d', criterion, reliability), we are describing what the design *could* support, not what it currently delivers. The gap is the subject of Section 7.

---

1. The measurement paradigm

1.1 The CPT / Go-No-Go lineage

The Continuous Performance Test, introduced by Rosvold and colleagues in 1956, is one of the most durable instruments in cognitive psychology. Its logic is minimal: present a long stream of stimuli, one at a time, at a fixed pace; designate a *target*; ask the participant to respond to the target and to nothing else. Two error types then become observable and dissociable:

- **Omission errors** — failing to respond to a target. Classically read as lapses of **sustained attention** (vigilance).
- **Commission errors** — responding to a non-target. Classically read as failures of **response inhibition** (impulsivity).

The **Go/No-Go** paradigm sharpens the second axis: some stimuli are explicit "No-Go" signals that look enough like a "Go" to invite a response, so that withholding demands active inhibition rather than mere absence of a target.

FOCUS_NOW!! is a compact hybrid of the two. Its *target* trials are Go trials; its *joker* trials are No-Go trials that share the target's semantic content but carry a distinct colour signature; its *random* trials are the neutral background stream.

1.2 The four constructs in play

Construct	Where the game engages it
Sustained attention / vigilance	Maintaining readiness across a stream of mostly-irrelevant images at a fixed pace.
Response inhibition	Withholding a prepared press on joker (No-Go) trials.
Processing speed	The response window equals one stimulus-onset interval; slower processing costs hits.
Perceptual discrimination	Colour-reduced (dithered) images make target identification itself effortful, and increasingly so as the player succeeds.

The remainder of the paper follows these four constructs into the source code.

---

2. The trial as a unit of measurement

The atomic unit of any psychometric task is the **trial**: a controlled stimulus, a bounded response window, and a scored outcome. In FOCUS_NOW!! the trial is not a metaphor — it is a literal iteration of a Bevy ECS schedule.

2.1 The task as a state machine

The game is a five-state machine (`lib.rs`). Only one state, `Playing` , collects measurements; the others are instruction, selection, and reporting scaffolding.

```
enum GameState {
    Loading,       // asset initialisation
    Instructions,  // target is memorised here
    Playing,       // measurement happens here
    Menu,          // difficulty (SOA) is selected here
    Score,         // outcome is reported here
}
```

Psychometrically this matters: the `Instructions` state is where the **target template** is established in the participant's memory, and the `Menu` state is where the **stimulus-onset asynchrony (SOA)** — the task's core difficulty parameter — is fixed. Measurement is cleanly separated from instruction, which is exactly the separation a testing protocol requires.

2.2 The per-trial pipeline

Within `Playing`, five systems run in a fixed order every frame (`game.rs`):

```
(
    update_score_display,   // report running score
    update_grace_timer,     // close the response window when it expires
    handle_player_input,    // score the response
    manage_image_transitions, // advance to the next trial
    apply_pending_image,    // present the stimulus
).chain()
```

The ordering is a measurement decision. Input is scored (`handle_player_input`) *before* the next trial is generated (`manage_image_transitions`), so a response is always attributed to the stimulus that was on screen when the key went down — never to the one about to replace it. Trial attribution, the thing that most often corrupts home-grown reaction-time tasks, is handled by system order.

2.3 The roll table: stimulus scheduling

Each trial's *type* is drawn from a ten-sided roll (`manage_image_transitions`):

```
const ROLL_RANGE: i32 = 10;
const TARGET_THRESHOLD: i32 = 8; // roll >= 8 → target
const JOKER_ROLL: i32 = 0;       // roll == 0 → joker

let roll = rng.gen_range(0..ROLL_RANGE);
if roll >= TARGET_THRESHOLD && !was_target {
    // Go trial (target)
} else if roll == JOKER_ROLL {
    // No-Go trial (joker)
} else {
    // neutral distractor
}
```

This yields an approximate base rate of **~20% Go**, **~10% No-Go**, **~70% neutral**, with one constraint: `!was_target` forbids two targets in a row. The constraint is psychometrically meaningful — back-to-back targets would let a participant "ride" a single readiness state and would inflate hit rates without engaging sustained attention. Prohibiting them keeps each Go trial an independent demand on re-orienting.

A note on honesty: `rand::thread_rng()` is a fresh, unseeded generator per trial. The stimulus stream is therefore **not reproducible** across sessions or players. For a game this is fine; for an instrument it is a limitation (Section 7).

2.4 The grace timer: the response window

When a target is shown, the code spawns a one-shot **grace timer** whose duration is exactly one SOA:

```
game_vars.current_image = CurrentImage::Target;
game_vars.possible_score += 1;
commands.spawn(UpdateGraceTimer {
    timer: Timer::new(Duration::from_millis(timer_duration(&difficulty.value)),
                      TimerMode::Once),
});
```

While this timer runs, a Space press is scored as a hit. When it finishes (`update_grace_timer`), the game reverts the current stimulus to `Random` and re-arms input:

```
if grace_timer.timer.finished() {
    game_vars.current_image = CurrentImage::Random;
    game_vars.space_already_registered = false;
    commands.entity(entity).despawn();
}
```

This *is* the response window of a CPT: a bounded interval, tied to stimulus onset, within which a response "counts". Its width equals the SOA, so faster difficulty levels shrink the window — coupling **processing speed** demands directly to the **vigilance** demands, which we return to in Section 6.

---

3. A Signal Detection Theory view

3.1 The 2×2 that the game already contains

Signal Detection Theory decomposes performance on a detection task into two independent quantities: **sensitivity** (how well signal is told apart from noise) and **bias** (how willing the observer is to say "signal"). Its bookkeeping is a 2×2 contingency table. FOCUS_NOW!! generates every cell of it:

	Player presses (says "target")	Player withholds
Target on screen (signal)	**Hit** — scored **+1**	**Miss** (omission) — scored 0
Joker / Random (noise)	**False alarm** — scored **−1**	**Correct rejection** — scored 0

The mapping to code is direct (`handle_player_input`):

```
match game_vars.current_image {
    CurrentImage::Target => { score += 1; }                  // Hit
    CurrentImage::Random
    | CurrentImage::Joker
    | CurrentImage::EndGame => { score -= 1; }               // False alarm
}
```

3.2 The scoring rule is a payoff matrix

The `+1 / −1` rule is not neutral. In SDT terms it is a **payoff matrix**, and a payoff matrix sets the participant's optimal **response criterion**. Because hits and false alarms are weighted symmetrically (+1 and −1) while misses and correct rejections are weighted at zero, an idealised score-maximising

player should press whenever their internal evidence that "this is the target" exceeds the point where the *expected* gain from a possible hit outweighs the *expected* loss from a possible false alarm. Change the weights — say −2 for a joker press — and you would deliberately shift players toward a more conservative criterion, trading hits for fewer commission errors. The scoring constants are thus a **bias knob**, and the paper flags them as the single most powerful psychometric lever in the game.

3.3 d' and criterion: derivable in principle, not yet logged

Given counts of hits and false alarms over a session, SDT yields two summary scores:

- **d' (sensitivity)** = z(hit rate) − z(false-alarm rate) — how far the signal distribution sits from the noise distribution, in standard-deviation units. Higher d' = sharper attention/discrimination, *independent of bias*.
- **criterion (c)** = $-\frac{1}{2} \cdot [z(\text{hit rate}) + z(\text{false-alarm rate})]$ — how conservative or liberal the player is, *independent of sensitivity*.

FOCUS_NOW!! currently reports only the **net payoff score** (`game_vars.score`), which confounds these two. A player who presses at everything and a player who discriminates perfectly can, in principle, land on similar net scores by different routes. Separating them requires logging the four SDT cell counts per session — a change discussed in Section 7. The important point for now: the design *already produces the events*; it simply does not *record* them.

---

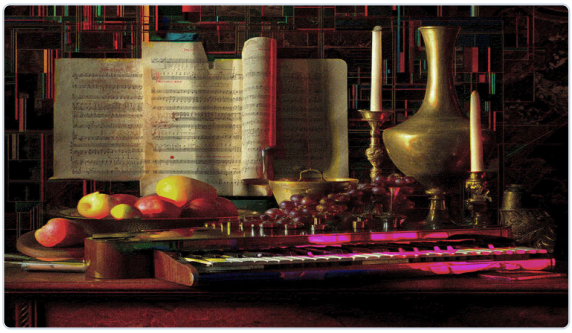
4. The Joker and response inhibition

4.1 A feature-based No-Go signal

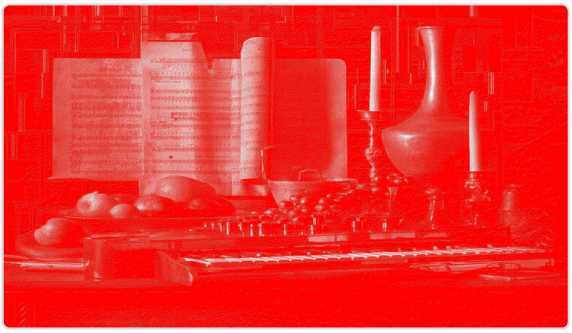
The joker is the paper's most interesting single mechanic. It is **the target image itself**, re-rendered in a red tone — a two-colour dither of red with white highlights (`loading.rs`):

```
fn random_joker_palette() -> [dither::RGB<u8>; 2] {
    [dither::RGB(255, 0, 0), dither::RGB(255, 255, 255)]
}
```

```
let is_joker = game_vars.current_image == CurrentImage::Joker;
let effective_nb_colors = if is_joker { 2 } else { nb_colors };
let dither_colors = if is_joker { Some(random_joker_palette()) } else { /* … */ };
```



(a) **Go** — target, normal palette
Correct response: *press*



(b) **No-Go** — same target, red-toned joker
Correct response: *withhold*

Figure 1. The joker is the target image itself, re-dithered to a red tone (a two-colour dither of red with white highlights). Identical semantic content; the only distinguishing signal is a low-level colour feature — which is exactly what makes withholding a prepotent response effortful.

Semantically, the joker *is* the target. What distinguishes it is a **low-level perceptual feature** — its colour signature — not its identity. This makes the No-Go decision a **feature-based inhibition** demand: the participant has built a readiness to respond to the target's *content*, and must override that prepotent response on the basis of a colour cue that arrives together with the very content that triggers the urge to press.

4.2 Why this is a good inhibition probe

Response inhibition is hardest, and most diagnostic, when the No-Go stimulus is **maximally similar** to the Go stimulus — when the "stop" signal shares most of the features that drive the "go" response. A joker that looked nothing like the target would be trivially rejected and would measure little. By reusing the target content and inverting only the colour, the game maximises response conflict, which is precisely the condition under which commission errors reveal individual differences in inhibitory control.

4.3 One response per stimulus

A single latch prevents a held or double key-press from being scored twice (`handle_player_input` / `manage_image_transitions`):

```
if actions.spaced && !game_vars.space_already_registered {
    /* … score exactly once … */
    game_vars.space_already_registered = true;
}
```

Psychometrically this enforces the **one-decision-per-trial** assumption that the SDT 2×2 relies on. Without it, a nervous double-tap on a joker would be counted as two commission errors and would distort the bias estimate.

5. Adaptive difficulty as an ability staircase

5.1 Two coupled difficulty axes

FOCUS_NOW!! varies difficulty along two axes at once.

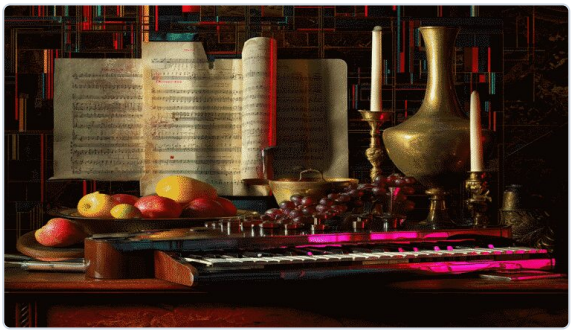
Axis 1 — SOA (fixed per session). The difficulty menu sets the stimulus interval and, with it, the response window (`game.rs`):

```
fn timer_duration(difficulty: &GameDifficulty) -> u64 {
    match *difficulty {
        GameDifficulty::Easy   => 2000,  // ms per image
        GameDifficulty::Medium => 1000,
        GameDifficulty::Hard   =>  800,
    }
}
```

Axis 2 — Perceptual degradation (adaptive within a session). A JSON config (`assets/dither_config.json`) maps score bands to a colour count *and* a dithering algorithm, degrading the image as the running score climbs:

```
{
  "pack": "nature_morte",
  "jpegQuality": 35,
  "defaultColors": 30,
  "defaultAlgorithm": "floyd",
  "jokerAlgorithm": "floyd",
  "rules": [
    { "scores": [1, 2], "colors": 40, "algorithm": "atkinson" },
    { "scores": [3, 4], "colors": 20, "algorithm": "stucki" },
    { "scores": [5, 6], "colors": 5,  "algorithm": "burkes" },
    { "minScore": 9,    "colors": 2,  "algorithm": "sierra3" }
  ]
}
```

Broadly, higher scores map to fewer colours and a harder discrimination: at rest (score ≤ 0, and again at 7–8) the image uses the 30-colour default, and once the score passes 9 it collapses to just 2 colours. The bands are not strictly monotone — score 1–2 actually gets *more* colours (40) than the default — but the intent is the same: performance drives difficulty.



(a) **Many colours** — target easy to recognise

(b) **Two colours** — target much harder to recognise

Figure 2. The two extremes of the current config: (a) 40 colours — the clearest band, reached at score 1–2 — versus (b) 2 colours, the setting once the score passes 9. Fewer colours means a harder perceptual discrimination, so as the running score climbs the target itself becomes harder to recognise: performance drives difficulty.

5.2 This is a crude adaptive staircase

The dithering rule is, in spirit, an **adaptive staircase** — the family of procedures (Method of Constant Stimuli's adaptive cousins; up-down staircases) that psychophysics uses to home in on a participant's *threshold*: the stimulus level at which they perform at some target accuracy. The logic is the same — *make it harder when they succeed, easier when they fail* — so that a skilled and an unskilled player converge toward a comparable subjective difficulty, and the *level they settle at* becomes the ability estimate.

FOCUS_NOW!!'s version is deliberately coarse:

- The step schedule is a fixed table keyed on **net score**, not on a running accuracy estimate, so it inherits the bias/sensitivity confound of Section 3.
- It is **not even monotone in score** with wide, uneven bands: score 1–2 gets 40 colours — *more* than the 30-colour default — and scores 7–8 fall back to that default before score ≥ 9 drops to 2, so "harder with skill" holds only loosely.
- It adjusts **stimulus clarity** — colour count and dithering algorithm — but not **SOA**, so only one of the two difficulty axes actually adapts.

5.3 The ideal it gestures at: IRT / CAT

The principled version of "adapt difficulty to ability" is **Item Response Theory (IRT)** and its runtime cousin **Computerised Adaptive Testing (CAT)**. There, each item has a calibrated difficulty on the same latent scale as the person's ability, and the next item is chosen to be maximally *informative* about that ability given responses so far. FOCUS_NOW!! has no item calibration and no explicit ability scale; its "items" (images × colour levels) are uncalibrated. The staircase is a gameplay feel-good loop that *approximates* the shape of adaptive measurement without its psychometric machinery — a fair thing for a game to do, and a clear target for a future instrument.

5.4 The termination rule

The session ends after a fixed number of **target** presentations, not a fixed number of trials (`game.rs`):

```
const NB_GOAL_IMAGE: i8 = 10;
// …
if game_vars.possible_score >= NB_GOAL_IMAGE { /* end after 10 targets */ }
```

Fixing the count of *scored signal events* rather than total trials keeps the number of measurement-bearing opportunities constant across players, even though the random stream means total trial counts (and session lengths) vary. That is a reasonable stopping rule for a fixed-precision test — though ten targets is far too few for a stable estimate (Section 7).

---

6. Reading speed, vigilance, and their coupling

Because the response window equals the SOA (Section 2.4), FOCUS_NOW!! does **not** separate *how fast* you respond from *whether* you respond in time. A hit on Hard (800 ms window) certifies a faster perception-decision-action loop than a hit on Easy (2000 ms), but the game records both as an identical "+1". Processing speed is therefore *implied by* difficulty selection rather than *measured* as a latency.

A conventional CPT records **reaction time** on every hit and reports its mean and, crucially, its **variability** — intra-individual RT variance is among the most reliable markers of attentional lapses. FOCUS_NOW!! has the timing information in hand (the grace timer knows exactly when the window opened) but discards the key-down timestamp. Capturing it is the highest-value, lowest-cost instrumentation change available, and it is the first item in the agenda below.

---

7. Threats to validity and a measurement research agenda

An honest psychometric appraisal must state plainly: **the game measures a score, not yet a trait.** Turning it into an instrument requires work along three fronts.

7.1 Instrumentation gaps (engineering, low cost)

Gap	Consequence	Fix
Only net score is logged	d' and criterion are unrecoverable; sensitivity and bias are confounded	Log per-trial `{type, response, correct}` → recover the SDT 2×2
Key-down time is discarded	No reaction-time mean or variability (the best lapse marker)	Timestamp the press relative to grace-timer start
Unseeded RNG per trial	Stimulus stream is non-reproducible; sessions are not comparable	Seed the generator; record the seed with the session
10 targets per session	Far too few observations for a stable estimate	Lengthen sessions or aggregate across sessions

7.2 Reliability (does it measure *consistently*?)

None of the standard reliability evidence exists yet:

- **Test–retest reliability** — does the same player score similarly across days? Requires the seeded, logged sessions above, then a correlation across repeat sittings. Attention tasks are notoriously susceptible to practice effects, so retest must control for learning.
- **Internal consistency** — do the trials cohere as measures of one thing? Estimable (e.g. split-half on the trial log) once trial-level data is captured.

Without these, no claim about *individual* scores is defensible — reliability is the ceiling on validity.

7.3 Validity (does it measure *what we claim*?)

- **Convergent validity** — do FOCUS_NOW!! scores correlate with an established CPT (e.g. Conners CPT, the TOVA) in the same participants? This is the load-bearing study; until it exists, "attention training game" is a design intention, not an evidenced property.
- **Construct validity** — do the *dissociations* the design predicts actually appear? For instance, joker commission errors should load on inhibition, target omissions on vigilance; a factor analysis over many players would test whether the two axes separate as the paradigm assumes.
- **Norming** — score interpretation ("is 6 good?") is meaningless without a reference distribution stratified by at least age, and by SOA level.

7.4 A confound the design should own

Because the adaptive dithering makes the target *harder to see* as the player succeeds, the game deliberately entangles **attention** with **perceptual discrimination**. A player might lose points late in a session because their *attention* lapsed or because the 2-colour target became genuinely *hard to recognise* — the net score cannot tell these apart. For a game this coupling is a virtue (it keeps difficulty engaging); for an instrument it is a confound that would need to be isolated, e.g. by holding colour depth fixed within a measurement block.

---

8. Conclusion

FOCUS_NOW!! is a faithful, if compact, expression of the Go/No-Go / CPT paradigm, and its Rust implementation is unusually legible as a *measurement design*: the trial is an ECS iteration, the response window is a one-shot timer, the stimulus schedule is a ten-sided roll, the payoff matrix is a two-line `match`, and the adaptive difficulty is a score-keyed lookup. Every psychometric construct the game touches has a precise line of code behind it.

What the game does not yet have is *evidence* — the logged trial data, the reliability estimates, and the validation studies that separate a score from a trait. The encouraging finding of this review is that the distance is mostly **instrumentation, not redesign**: the events that a real CPT scores are already being generated frame by frame; they are simply not being written down. Capture them, seed the randomness, lengthen the session, and FOCUS_NOW!! becomes a defensible platform for the studies that would earn it the word "psychometric" without qualification.

Appendix — Glossary

- **CPT (Continuous Performance Test)** — a sustained-attention task: respond to a target in a long, paced stimulus stream.
- **Go/No-Go** — a paradigm in which some stimuli require a response (Go) and perceptually similar others require withholding it (No-Go), isolating inhibition.
- **SOA (stimulus-onset asynchrony)** — the time between successive stimulus onsets; here, the per-image interval.
- **Signal Detection Theory (SDT)** — a framework separating sensitivity (d') from response bias (criterion) in detection tasks.
- **d' (d-prime)** — bias-free sensitivity: z(hit rate) – z(false-alarm rate).
- **criterion (c)** — response bias: the evidence threshold at which the observer decides “signal”.
- **Commission / omission error** — pressing on a non-target / failing to press on a target.
- **Staircase** — an adaptive procedure that raises difficulty after successes and lowers it after failures to converge on a performance threshold.
- **IRT / CAT** — Item Response Theory and Computerised Adaptive Testing: the calibrated, information-maximising ideal that the game’s staircase approximates.