Online Stochastic Trial Minimisation

Andrew Bistras*
University of California, Los Angeles
Astrid Cybernetics

Draft, October 17, 2026

Abstract

Human subjects research often gives each participant a set of conditions from a larger pool, such as vignettes, choice sets, intervention components or treatment sequences. Investigators want every condition used equally often at every point of enrollment. Complete randomization balances exposure only in expectation, and least-filled quotas are deterministic, so they fail when participants enroll at the same time or drop out unevenly. We generalize minimisation, the balancing method of clinical trials, to sets of conditions under within-participant constraints, and show that it is a feedback controller with one natural parameter: the bits of randomness each allocation may give up. The algorithm rolls one weighted die whose faces are the conditions, reweighs the faces after every participant, and keeps a roll only when it forms an admissible set. The temperature is solved so that each allocation spends exactly its budget. The rule is the maximum-entropy allocation for its expected balance, and its balance criterion is a second-order Jensen–Shannon divergence between the allocation history and the new allocation. A fixed-temperature version solves an entropy-regularized control problem exactly, and imbalance stays bounded. In simulations of a 2^6 vignette design with five vignettes per participant and no minimal pairs, a budget of 1 bit (of 28.3 available) kept the exposure SD near 0.9 over 1,000 participants, against 8.4 under complete randomization. It stayed stable when 64 participants shared a stale ledger, where deterministic minimisation did three times worse than randomization, and with a ledger of completions it corrected the unequal condition representation caused by differential attrition.

Keywords: minimisation; adaptive randomization; counterbalancing; vignette studies; entropy; online enrollment; attrition.

1 Introduction

Human subjects research often gives each participant several conditions drawn from a larger set, and investigators want every condition used about equally often. The pattern runs across fields: case vignettes rated by clinicians or the public (Peabody et al., 2000; Auspurg and Hinz, 2015), discrete choice experiments in health economics (Johnson et al., 2013), crossover trials (Senn, 2002), and factorial experiments that screen intervention components (Collins, 2018). Our running example is a vignette study in which six binary attributes define 64 vignettes and each participant rates five.

*Correspondence: andrew.bistras@gmail.com. Code, simulations and this manuscript: `https://github.com/andrewbistras/min_sim`.

Complete randomization balances exposure only in expectation. With 64 vignettes and five per participant, the counts after 320 participants scatter with a standard deviation of about 4.6 around a target of 25. The most- and least-shown vignettes then differ by about 21 showings, so precision varies across stimuli, and a study that stops early or loses data to exclusions can leave some vignettes nearly unobserved.

Enrollment is sequential, bursty and leaky, and quota tools handle it badly. Participants arrive one at a time, often at several sites or through several recruitment channels at once, and some leave before they finish. Common tools allocate the least-filled conditions, such as Qualtrics' "evenly present elements," Gorilla's counterbalancing nodes and psiTurk's choice among the least-filled conditions (Gureckis et al., 2016); apart from tie-breaking these rules are deterministic, so simultaneous enrollments receive the same allocation. Counts of assignments rather than completions also drift when dropout differs by condition, as it often does (Zhou and Fishbach, 2016; Bell et al., 2013).

Constraints within a participant make the problem harder. Researchers often forbid "minimal pairs" (two vignettes that differ in one factor only) because such pairs expose the manipulation and invite demand effects (Charness et al., 2012; Horiuchi et al., 2022). The evidence that demand effects are large is mixed (Mummolo and Peterson, 2019; Clifford et al., 2021), so the rule is cheap insurance rather than a necessity. Either way, the least-filled stimuli may be mutually incompatible, and quota rules have no principled way to trade balance against the constraint.

Clinical trials met a version of this problem fifty years ago. Minimisation assigns each new participant to the arm that most reduces imbalance (Taves, 1974; Pocock and Simon, 1975; Altman and Bland, 2005). It appears in roughly one in ten trials published in leading medical journals (Kahan and Morris, 2012; Ciolino et al., 2019), and regulators accept it when each allocation keeps a random element (ICH, 1998; EMA, 2015). A long literature asks how much randomness to keep, from Efron's biased coin (Efron, 1971) and Klotz's maximum-entropy allocation (Klotz, 1978) to modern comparisons of balance and predictability (Atkinson, 2014).

Those methods allocate one treatment per participant. None of them handles constrained sets of conditions, simultaneous enrollment or the unequal condition counts that differential attrition leaves. Minimisation over the 16 cells of a 2^4 factorial trial made 34% of allocations deterministic (Kuhn et al., 2019), social scientists have minimized over respondents' covariates (Moore and Moore, 2013), and adaptive vignette allocation has been limited to a biased coin on a single factor (Kim and Yang, 2024).

This paper generalizes minimisation to allocating sets of conditions as participants enroll, treating it as a feedback controller with one reportable parameter. That parameter is the number of bits of randomness each participant's allocation may give up. The algorithm keeps one weighted die whose faces are the conditions, reweighs the faces after every participant, and rolls it until the faces form an admissible set. The contributions are:

1. A framing of allocation as feedback control, which explains why the controller must stay random when participants enroll at the same time.

2. An exact, simple algorithm whose single parameter has a direct reading: a budget of κ bits leaves each allocation as random as a uniform draw from a fraction $2^{-\kappa}$ of all admissible sets, and caps at κ bits what anyone who knows the history can learn about the next allocation.

3. Theory linking the method to maximum entropy, the Jensen–Shannon divergence and entropy-regularized optimal control, with a stability guarantee.

4. Simulations of a 64-vignette design covering balance, simultaneous enrollment, differential attrition and order effects; a JavaScript implementation; and guidance for protocols and reports.

2 The allocation problem

Each participant receives a deck of k stimuli from a pool of m, and we want the running exposure counts to stay equal while every deck stays random.

Setting. Participants arrive one at a time, or in bursts. Participant t receives a deck S_t: a set of k distinct stimuli, shown in some order. A compatibility rule between pairs of stimuli defines the admissible decks, written $\mathcal{A}$.

The ledger. Let $c_i(t)$ be the number of the first t decks that contain stimulus i. With equal targets, each stimulus should have been shown kt/m times after t participants. The ledger is the only state the allocator needs.

Running example. Six binary factors define $m = 64$ vignettes, each coded as a 6-bit string. Each participant rates $k = 5$ of them, and no two vignettes in a deck may differ in exactly one factor (Hamming distance at least 2). In graph terms, admissible decks are independent sets of size five in the six-dimensional hypercube.

How much room the constraint leaves. Of the 7,624,512 five-vignette sets, 2,682,432 (35.2%) are admissible. With 120 presentation orders each, that gives 321,891,840 ordered decks, or 28.26 bits: the most randomness any single allocation can carry. The order supplies 6.91 of those bits; the choice of set supplies the other 21.36.

Complete randomization is balanced only in expectation. Every vignette has the same six forbidden partners, so a uniform draw from $\mathcal{A}$ includes each vignette with probability exactly 5/64. The problem is the fluctuation around that expectation, which grows with the square root of the number of participants.

Bits spent. We measure how far an allocation departs from complete randomization by the bits it gives up,

$$\kappa_t \;=\; \log_2 |\mathcal{A}| - H(\pi_t) \;=\; \mathrm{KL}\big(\pi_t \,\big\|\, U\big), \tag{1}$$

where π_t is the probability distribution over admissible sets used for participant t and U is the uniform distribution on $\mathcal{A}$. A budget of κ bits leaves the allocation as random as a uniform choice among a fraction $2^{-\kappa}$ of the admissible decks: 1 bit means half of them, 2 bits a quarter.

Table 1: What a good allocator delivers.

Goal	Measure	Ideal
Balanced exposure at every point	SD of $c_i(t)$ around kt/m; range of the counts	as small as integer counts allow
A random allocation for each participant	Entropy H_t of the allocation law, in bits	close to the 28.26-bit maximum
Robustness online	Balance when b participants share one stale ledger; balance among completers when dropout differs	unaffected
No order effects by design	Presentation order	uniform over the 120 orders

Exposure imbalance. We summarize balance by the standard deviation of the counts around their target,

$$\mathrm{sd}(t) \;=\; \Big(\frac{1}{m}\sum_{i=1}^{m}\big(c_i(t)-kt/m\big)^2\Big)^{1/2}. \tag{2}$$

Under complete randomization $\mathrm{sd}(t)$ grows like $\sqrt{t}$. A good minimisation rule keeps it bounded for all t.

3 Minimisation as feedback control

Allocation is a control problem: the ledger accumulates every decision, and the allocator steers it back toward equal exposure.

The ledger integrates the allocations. Write $e_i(t) = c_i(t) - kt/m$ for stimulus i's surplus over its target. Each deck adds one showing to each of its stimuli and raises every target by k/m, so with $x(S)$ the 0/1 inclusion vector of deck S,

$$e(t+1) \;=\; e(t) + x(S_t) - \tfrac{k}{m}\mathbf{1}. \tag{3}$$

Under complete randomization $e(t)$ is a random walk and drifts away like $\sqrt{t}$. Any rule that looks at $e(t)$ before choosing S_t is a feedback controller for this integrator.

Open loop or closed loop. A schedule of decks fixed in advance, such as a Latin square or an offline D-efficient design, is open-loop control. It is exact only if every participant completes, in order, and the study stops where planned. Online, participants drop out, arrive in bursts, and studies are stopped early or extended, so the allocator has to read the ledger and correct what actually happened.

Deterministic minimisation is the high-gain limit. Because every deck has the same number of stimuli, the squared surplus after deck S differs between decks only through the counts of its members:

$$\big\|e + x(S) - \tfrac{k}{m}\mathbf{1}\big\|^2 \;=\; \|e\|^2 + 2\sum_{i\in S} c_i + \text{const.} \tag{4}$$

Minimizing it means choosing the admissible deck with the smallest total count. This is Taves's (1974) minimisation applied to decks; in machine learning the same rule is called herding (Welling, 2009).

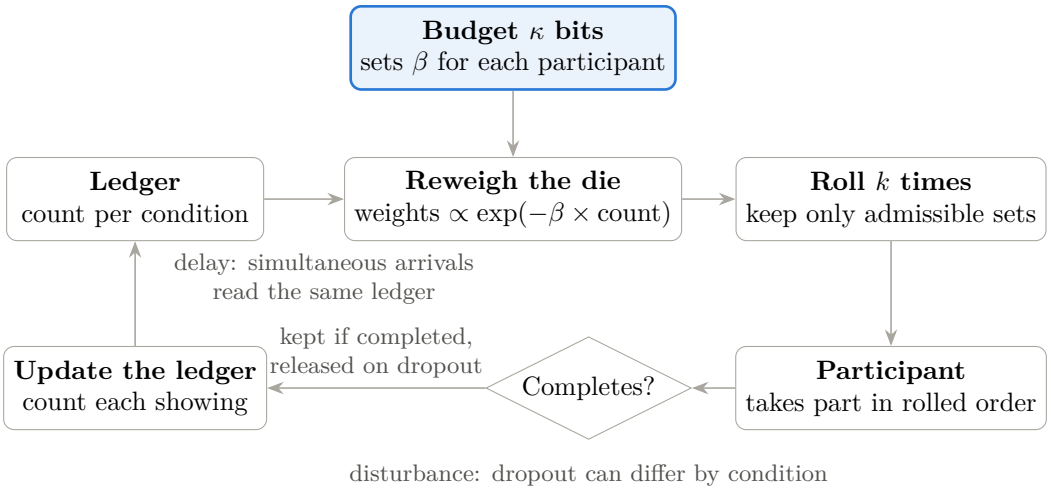


Figure 1: Allocation is a feedback loop, and only completed sessions keep their count. The budget is the only setting; the ledger, the die and the admissibility rule are fixed by the design.

Weakness 1: delay. When participants arrive faster than the ledger updates, several of them see the same counts, and a deterministic rule hands them all the same least-shown stimuli. A high-gain controller with a delay in its loop overshoots; Section 6 measures how badly.

Weakness 2: predictability. Anyone who knows the counts can foresee a deterministic allocation. In trials that lets recruiters steer which participant receives which condition, the selection bias that allocation concealment exists to prevent (Blackwell and Hodges, 1957; Schulz and Grimes, 2002). It also leaves randomization-based inference nothing to work with, and its pattern can line up with drifts in who enrolls when.

Stochastic minimisation softens the rule. We choose each admissible deck with probability

$$\pi_\beta(S) \;=\; \frac{\exp\big(-\beta \sum_{i\in S} c_i\big)}{\sum_{S'\in\mathcal{A}} \exp\big(-\beta \sum_{i\in S'} c_i\big)}, \qquad S \in \mathcal{A}. \tag{5}$$

Setting $\beta = 0$ gives complete randomization; letting β grow without bound gives deterministic minimisation with random tie-breaking. In control terms β is the gain of the feedback; in statistical physics $1/\beta$ is a temperature.

The single parameter is entropy, not gain. A given β is gentle when the counts are nearly equal and forceful when they are far apart, so it has no fixed meaning. We let the researcher fix the bits each allocation may give up, κ, and solve for β participant by participant:

$$\mathrm{KL}\big(\pi_{\beta_t} \,\big\|\, U\big) \;=\; \kappa \quad \text{for every participant } t. \tag{6}$$

The budget means the same thing in every design and at every point in a study, which is what a single default needs.

4 The algorithm: reweighted dice

The allocation law factorizes over the deck's members, so it can be sampled exactly by rolling one weighted die k times and keeping only admissible rolls.

Algorithm 1 Online stochastic trial minimisation, one allocation

Require: ledger $c \in \mathbb{N}^m$, admissible sets $\mathcal{A}$, deck size k, budget κ bits

$d_i \leftarrow c_i - \min_j c_j$ for $i = 1, \dots, m$

find $\beta \geq 0$ with $\mathrm{KL}(\pi_\beta \,\|\, U) = \kappa \ln 2$; if ties among the least-shown decks leave less to gain, set $\beta \leftarrow \infty$

$w_i \leftarrow \exp(-\beta d_i)$ for $i = 1, \dots, m$ ▷ reweigh the die

repeat

    roll $s_1, \dots, s_k$ independently with $\Pr(s_p = i) \propto w_i$

until $s_1, \dots, s_k$ are distinct and pairwise compatible

present the conditions in the order $s_1, \dots, s_k$

log the set, β and $\log \pi_\beta(\{s_1, \dots, s_k\})$

update the ledger (Section 7 discusses when a participant should count)

Why dice. The weight of a deck is a product over its members,

$$\exp\Big(-\beta \sum_{i \in S} c_i\Big) \;=\; \prod_{i \in S} \exp(-\beta c_i).$$

Roll k times a die whose face i has probability proportional to $\exp(-\beta c_i)$, keep the roll when its faces form an admissible deck, and the kept decks follow π_β exactly (Proposition 2). Because the same die is rolled for every slot, the presentation order is a uniformly random shuffle at no extra cost.

Reweighing after every participant. The die's weights are recomputed from the ledger before each allocation, and each showing divides a face's weight by e^β. Faces shown less often become heavier until the counts even out, and the solved temperature keeps each reweighing within the budget.

Calibrating β exactly. For designs of moderate size we enumerate $\mathcal{A}$ once: 2,682,432 sets in the example, in 107 ms. For each participant we tally N_E, the number of admissible decks whose offsets sum to $E = \sum_{i \in S} d_i$. The information spent is then a function of β alone (in nats; divide by $\ln 2$ for bits):

$$\kappa(\beta) \;=\; \log|\mathcal{A}| - \log \sum_E N_E\, e^{-\beta E} - \beta\, \bar{E}_\beta, \qquad \bar{E}_\beta \;=\; \frac{\sum_E E\, N_E\, e^{-\beta E}}{\sum_E N_E\, e^{-\beta E}}. \tag{7}$$

$\kappa(\beta)$ rises monotonically from 0 to $\log(|\mathcal{A}|/N_{E_{\min}})$, so bisection finds β to machine precision. One allocation, calibration included, takes 8–10 ms in Node.js, and about 18 ms at the very start of a study, when every deck ties.

Larger designs. When $\mathcal{A}$ is too large to enumerate, $\kappa(\beta)$ can be estimated from the dice themselves. With $\tilde{w}$ the die's normalized weights and $A(\beta)$ the probability that k rolls are admissible,

$$\mathrm{KL}(\pi_\beta \,\|\, U) \;=\; \mathbb{E}_{\pi_\beta}\Big[\sum_{p=1}^{k} \log\big(m\, \tilde{w}_{s_p}\big)\Big] - \log \frac{A(\beta)}{A(0)}. \tag{8}$$

Reusing the same random numbers across candidate values of β keeps the estimate smooth enough for bisection.

How many rolls. Under complete randomization 30% of rolls are admissible, and at budgets up to 2 bits the rate stayed between 24% and 31%, so a deck takes three or four rolls. At 4 bits it occasionally fell to 2% (about 50 rolls), when the die piled onto a few least-shown faces. If rolling keeps failing, the implementation draws directly from the enumerated list, which it also does whenever $\beta = \infty$.

Using the implementation. The JavaScript module has no dependencies and runs in Node.js or a browser:

```
import { hypercubeConflicts, createDesign, createAllocator } from './ostm.js';
const design = createDesign({ ...hypercubeConflicts(6, 2), k: 5 });
const allocator = createAllocator(design, { budgetBits: 1 });
const { deck, logProbSet } = allocator.allocate(counts);  // counts = the ledger
```

Each call returns the deck in presentation order and the log probability of its set of stimuli, for the analysis log.

5 Theory

Five results explain why the reweighted die is the right object: it is the maximum-entropy rule, it samples exactly, it maximizes a Jensen–Shannon gain, it solves an optimal-control problem exactly, and it keeps imbalance bounded. Throughout, $E(S) = \sum_{i \in S} c_i$ and U is the uniform law on $\mathcal{A}$. Proofs are in Appendix A.

Proposition 1 (Maximum entropy)**.** *Let $\beta > 0$. Every law π on $\mathcal{A}$ with $\mathrm{KL}(\pi \,\|\, U) \le \mathrm{KL}(\pi_\beta \,\|\, U)$ has $\mathbb{E}_\pi[E] \ge \mathbb{E}_{\pi_\beta}[E]$. Equivalently, among laws with the same expected total count, π_β has the most entropy.*

This is Jaynes's (1957) maximum-entropy principle, which Klotz (1978) applied to allocating single treatments.

Proposition 2 (Exact dice)**.** *Let $s_1, \dots, s_k$ be independent with $\Pr(s_p = i) \propto e^{-\beta c_i}$. Conditional on $\{s_1, \dots, s_k\}$ being an admissible set, the set has law π_β and its $k!$ orders are equally likely. The acceptance probability is $k! \sum_{S \in \mathcal{A}} \prod_{i \in S} \tilde{w}_i$.*

Proposition 3 (Jensen–Shannon identity)**.** *Let $P_t(i) = c_i(t)/(kt)$ be the ledger's distribution of showings, U_S the uniform distribution on the members of S, and $w = t/(t+1)$, so that adding S gives $P_{t+1} = wP_t + (1-w)U_S$ and*

$$\mathrm{JS}_w(P_t, U_S) \;=\; H(P_{t+1}) - w\,H(P_t) - (1-w)\,H(U_S).$$

(a) Maximizing $\mathrm{JS}_w(P_t, U_S)$ over decks is the same as maximizing the entropy of the updated ledger.

(b) Near balance, the minimisation score $-2E(S)$ equals $\frac{2(kt+k)^2}{m}\,\mathrm{JS}_w(P_t, U_S)$ plus a term that does not depend on S, up to third-order terms in the deviation of P_{t+1} from uniform.

(c) $H(P_t) = \mathrm{JS}(U_{S_1}, \dots, U_{S_t}) + \log k$: balanced exposure means decks that are as different from one another as possible.

(d) If every allocation spends at most κ bits, the mutual information between the history and the next deck is at most κ bits.

In the example after 200 participants, the score and the exact JS_w agree across admissible decks with correlation 0.9995.

Proposition 4 (Optimal control)**.** *Take the surplus e as the state, complete randomization as the passive dynamics, and the long-run average of $q(e_t) + \mathrm{KL}(\pi_t \,\|\, U)$ as the cost (Todorov, 2007, 2009). For fixed $\beta > 0$ and running cost*

$$q_\beta(e) \;=\; \log \mathbb{E}_{S\sim U}\Big[\exp\Big(-\beta \sum_{i\in S} e_i\Big)\Big],$$

the stationary policy π_β is optimal, with quadratic value function $v(e) = \frac{\beta}{2}\|e\|^2$.

The cost q_β is the cumulant-generating function of the deck's total surplus. It is quadratic near balance and linear far from it; with two arms and D the difference in counts it is $\log\cosh(\beta D/2)$, the hyperbolic-cosine family that Spencer (1977) used to bound online balancing.

Proposition 5 (Stability)**.** *Suppose every stimulus lies in the same number of admissible sets, and let $\lambda > 0$ be the smallest nonzero eigenvalue of the covariance of $x(S)$ under U. For any fixed $\beta > 0$, and for any per-participant budget $\kappa > 0$, the surplus process $e(t)$ is positive recurrent, and the long-run average of $\mathbb{E}\|e(t)\|$ is finite.*

The proof is a Foster–Lyapunov drift argument with $V(e) = \|e\|^2$, as in the analysis of covariate-adaptive designs by Hu and Hu (2012). In the example $\lambda = 0.055$.

Per-participant or average budget. A fixed β spends a different number of bits on each participant; a per-participant budget fixes the bits and lets β vary. With two arms the difference matters, because a per-allocation entropy floor reduces exactly to Efron's biased coin. Efron's $p = 2/3$ is a budget of 0.08 bits per allocation, and it needs about 47% more squared imbalance than the fixed-temperature rule at the same average entropy (Figure 2). With many stimuli the two versions coincide, because the spread of deck totals concentrates (Section 6.5). We use the per-participant budget because its guarantee holds for every participant, not just on average.

6 Simulation study

With a budget of 1 bit per participant, exposure stayed within about one showing of target for all 1,000 participants, while complete randomization drifted to a standard deviation of 8.4. The same budget tolerated simultaneous enrollment and, with a ledger of completions, corrected the unequal condition representation that differential attrition caused under complete randomization.

Design. We simulated the 64-vignette design of Section 2 under budgets from 0.25 to 8 bits, complete randomization (0 bits) and deterministic minimisation with random tie-breaking. Each run enrolled 1,000 participants, and each condition was replicated 10 times (6 times in the delay study, 8 in the attrition study). Exposure SD is averaged over participants 100 to 1,000 unless stated.

6.1 Balance versus budget

Any positive budget stops the drift. Under complete randomization the SD grew from 2.6 after 100 participants to 8.4 after 1,000, when the most- and least-shown vignettes differed by 42 showings. With 1 bit it stayed between 0.9 and 1.0 throughout, and the final range was 4.4 showings (Figure 3).

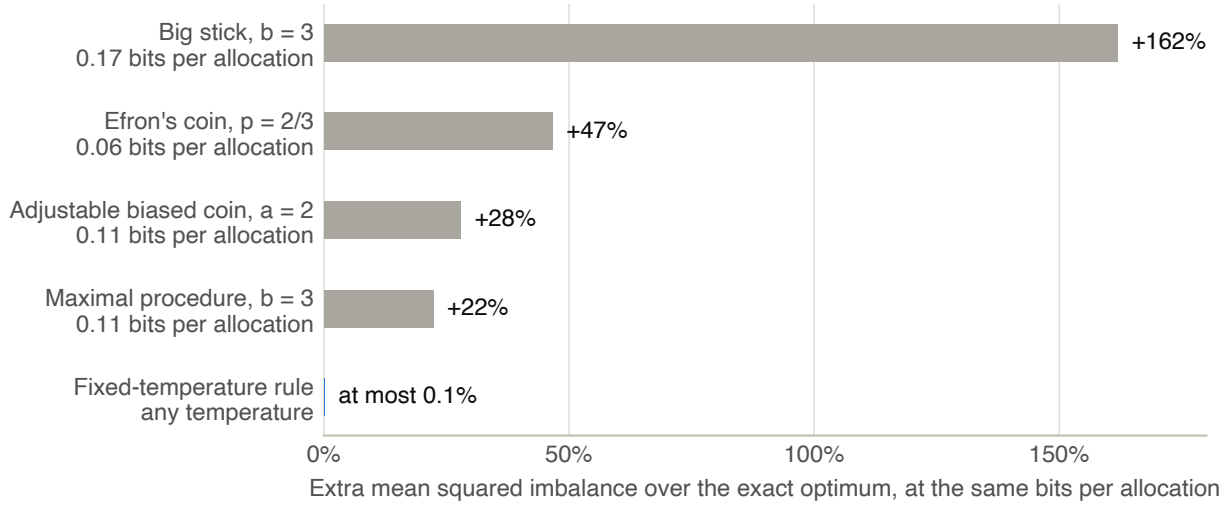


Figure 2: Classical two-arm designs need 22–162% more squared imbalance than the optimum at the same entropy rate; the fixed-temperature rule sits on the optimum. Excess mean squared imbalance against the exact KL-control optimum, at each design's usual setting. Across their parameter ranges the families stay above the optimum: Efron's coin by 6–65%, the adjustable biased coin by 6–53%, the maximal procedure by 17–27% and the big stick by 77–431%.

Returns diminish quickly. Each doubling of the budget cut the SD by roughly 30%, so the SD scales about as $1/\sqrt{\kappa}$ (Figure 4). Deterministic minimisation reached 0.40, near the floor that integer counts allow, while giving up 6.4 bits per participant on average.

Large budgets often cannot be spent. At 4 bits, 13% of participants met ties that left less than the budget to gain; at 8 bits, 63% did. Beyond about 2 bits the extra balance is small, while allocation probabilities become far more unequal (Section 6.4).

6.2 Simultaneous enrollment

Delay punishes high gain. When b participants share one stale ledger, deterministic minimisation went from an SD of 0.40 at $b = 1$ to 3.8 at $b = 16$ and 17.7 at $b = 64$, nearly three times worse than complete randomization (Figure 5). Stochastic budgets degraded gracefully: at 1 bit the SD went from 0.93 to 1.24 and 2.10.

The best budget shrinks as the delay grows. With up to 16 simultaneous participants, 2 bits did best; with 64, budgets of 0.5 to 1 bit did. Counting participants as soon as they are assigned shortens the delay and is the first remedy (Section 7).

6.3 Differential attrition

Counting completions restores equal representation of conditions; counting assignments does not. Each participant dropped out with probability 15% plus 6 points per vignette in their deck with factor A at its higher level, so dropout ranged from 15% to 45% and averaged about 30%. Under complete randomization, factor A made up 48.4% of completed showings instead of 50%, and the exposure SD among completers reached 9.0 by 1,400 enrollments (Figure 6).

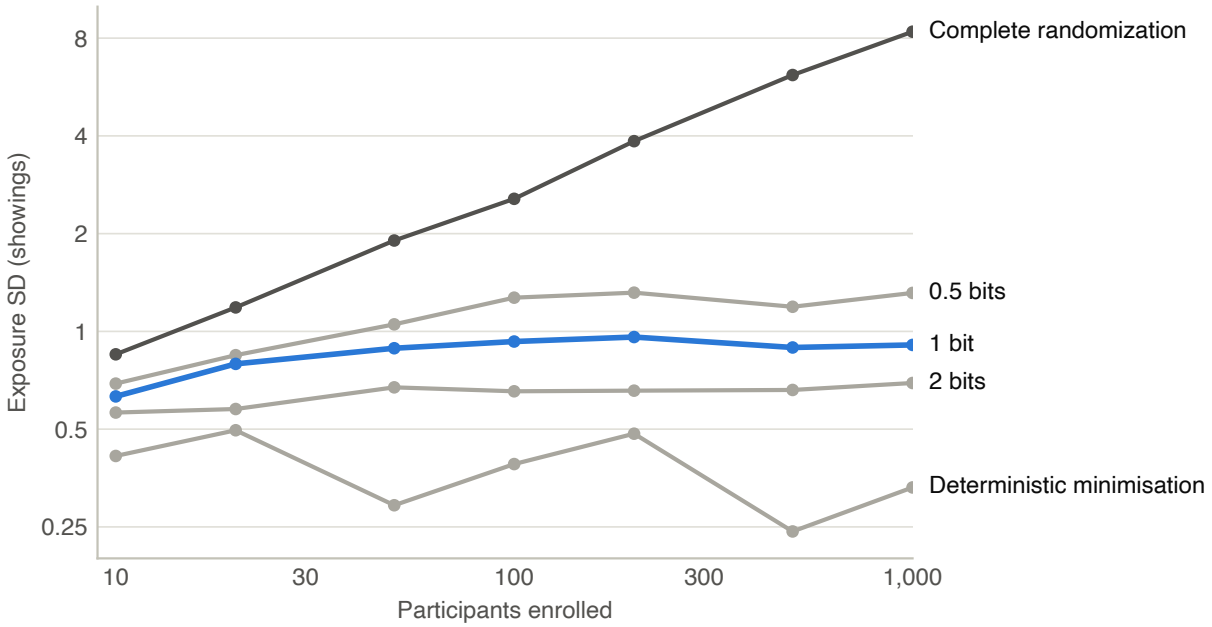


Figure 3: Randomization drifts like $\sqrt{n}$; 1 bit per participant holds exposure SD near 0.9. Exposure SD after each number of participants, mean of 10 runs, both axes on log scales.

The ledger decides what gets balanced. With 1 bit and a ledger of assignments, assignments were balanced but completed data were not: factor A made up 48.5% and the SD reached 5.3. With a ledger of completions, the same budget gave 49.9% and an SD of 0.90, the same balance as with no attrition.

Equal counts do not make the estimates unbiased. The simulation has no outcomes, so it shows which conditions completers saw, not who the completers are. When dropout depends on participants' characteristics, the completers in each condition can still differ, which is the selective attrition that Zhou and Fishbach (2016) documented, and the outcome analysis must still address it (Bell et al., 2013).

6.4 Order, factors and allocation probabilities

Order stays balanced in expectation. The G statistic for independence of vignette and position, divided by its degrees of freedom, lay between 0.98 and 1.06 for every budget, where 1 is what random order gives. Rolling one die for every slot shuffles the order, so no budget distorts it.

Factor levels are balanced too. After 1,000 participants (5,000 showings), the worst of the six factors showed one level 103 more times than the other under complete randomization. At 1 bit the worst gap was 10, and under deterministic minimisation 3.

The price is unequal allocation probabilities at any moment. After 500 participants at 1 bit, the most-favored vignette was 2.8 times as likely to be in the next deck as under complete randomization, and the least-favored 0.23 times as likely. These probabilities change with every participant, which is why the implementation logs them (Section 7).

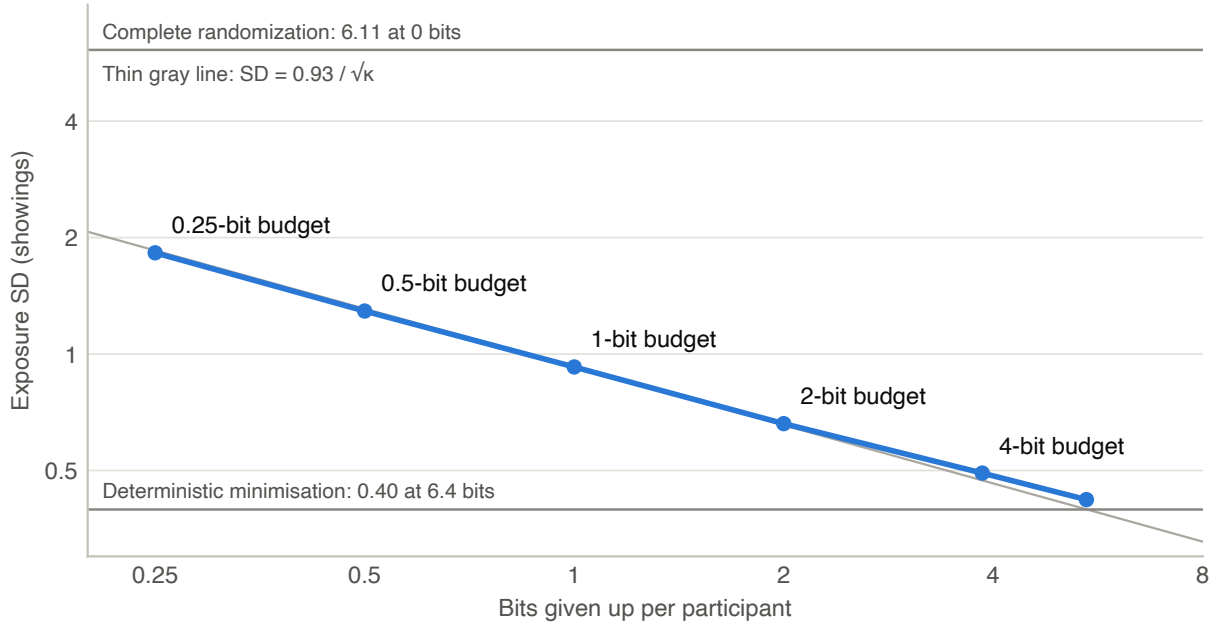


Figure 4: Each doubling of the budget cuts exposure SD by about 30%, to a floor near 0.4. Mean exposure SD over participants 100–1,000 against bits actually spent. The 4- and 8-bit budgets plot at the bits they spent, 3.9 and 5.4, because ties left nothing more to gain.

6.5 Per-participant budget or fixed temperature

In this design the two versions behave alike. We ran fixed temperatures, then per-participant budgets equal to the bits each temperature spent on average (Table 2). With 64 stimuli the spread of deck totals is stable, so under a fixed budget β barely moves: at 1 bit it stayed between 0.53 and 0.73 for 80% of participants.

Table 2: Exposure SD under a fixed temperature and under a per-participant budget that spends the same bits on average (6 runs each, 600 participants).

Fixed β	Bits spent on average	SD, fixed β	SD, per-participant budget
0.3	0.49	1.29	1.31
0.6	0.96	0.93	0.92
1.2	1.90	0.68	0.67

6.6 Runtime

One allocation took 8–10 ms in Node.js on a laptop, and about 18 ms at the start of a study, when every deck ties. Building the list of admissible decks took 107 ms and 13 MB of memory.

7 Practical recommendations

Use a budget of 1 to 2 bits per participant, count participants when they are assigned and release them if they drop out, log every allocation probability, and state the budget in the protocol.

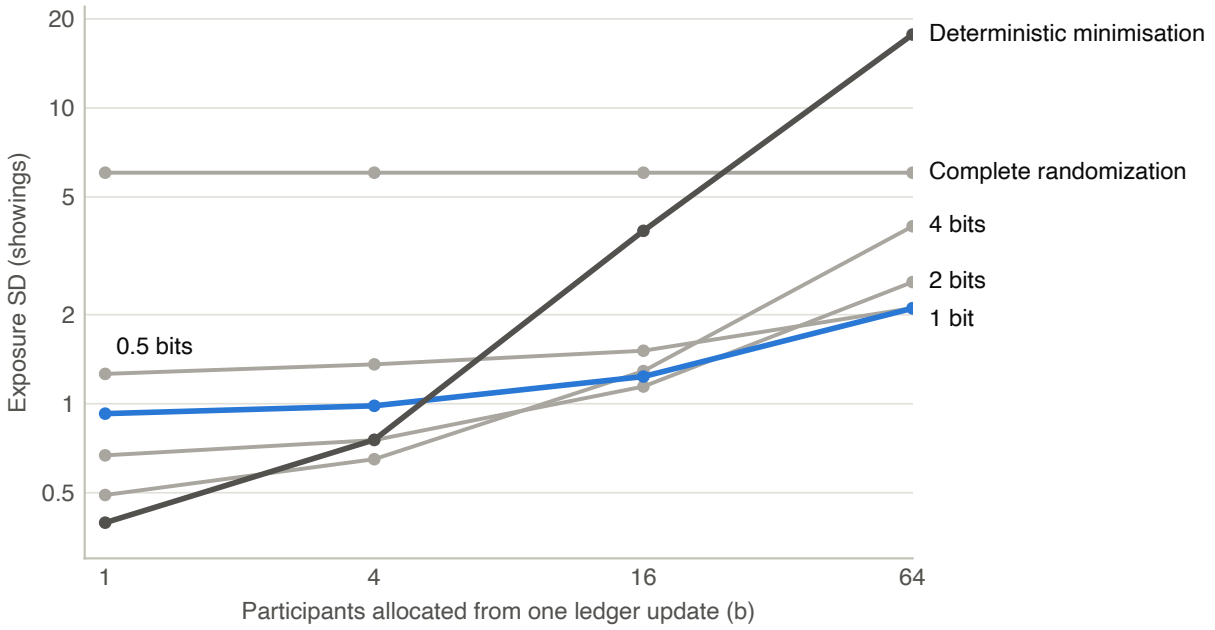


Figure 5: When participants share a stale ledger, deterministic minimisation overshoots. Mean exposure SD over participants 100–960 when b participants are allocated from one ledger update, 6 runs per condition, log scales. The ranking of budgets reverses as the delay grows: the classic trade-off between gain and delay in feedback control.

Choosing the budget. The figures in Section 6 refer to the 64-vignette example; the ordering of the recommendations carries over to other designs (Table 3). For a single treatment per participant the scale is different, because one allocation among two arms carries at most 1 bit. Efron's $p = 2/3$ corresponds to 0.08 bits, so budgets for two-arm trials sit well below 1 bit.

Table 3: Choosing the budget.

Situation	Budget per participant	Reason
Default: participants enroll one at a time	1 bit	Exposure SD about 0.9; each allocation as random as a uniform draw from half the admissible sets
Balance is paramount and enrollment is rarely simultaneous	2 bits	Exposure SD about 0.66; little gain beyond this
Tens of participants can share one stale ledger	0.5 to 1 bit	Lower gain overshoots less under delay
Staff who recruit could see the counts	as low as balance allows	The history reveals at most κ bits about the next allocation

When a participant counts. A ledger of assignments is fast but lets differential attrition skew which conditions completers saw, while a ledger of completions keeps completed showings balanced but reacts later. The better rule counts a participant at assignment and releases the count if they drop out or time out, as psiTurk does for sessions in progress (Gureckis et al., 2016). Define "completion" in the protocol, including any exclusions such as failed attention checks.

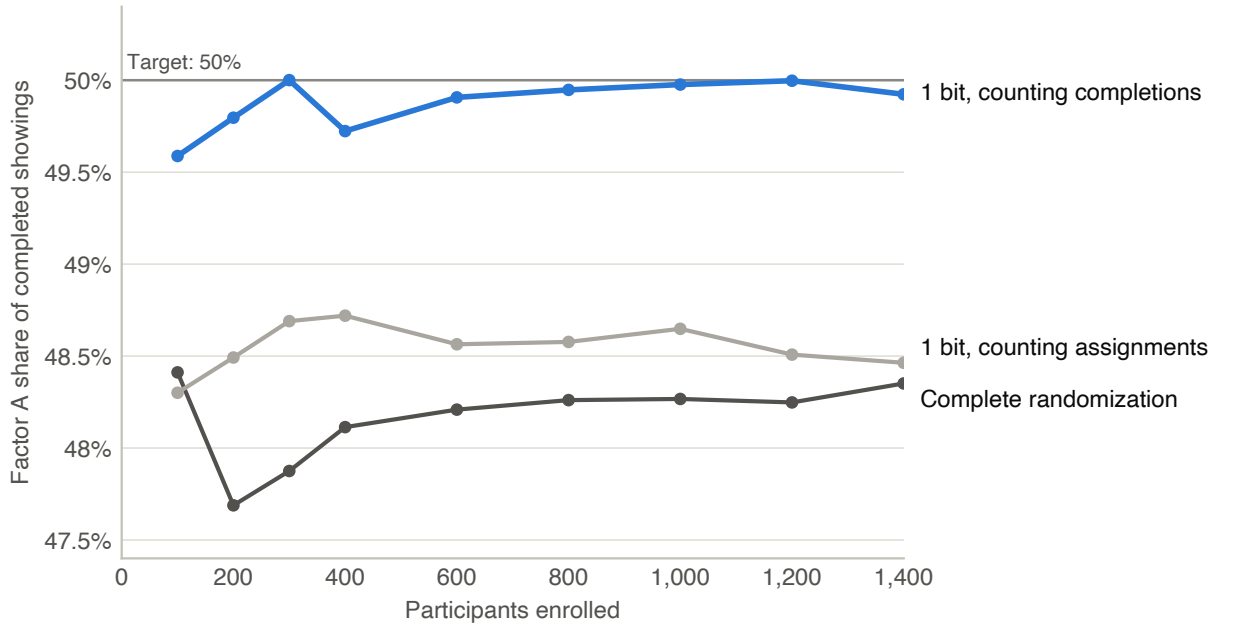


Figure 6: Only a ledger of completions keeps factor A at 50% of completed showings. Share of completed showings with factor A at its higher level, mean of 8 runs; dropout 15–45% by deck content. Deterministic minimisation with a ledger of completions also reached 50.0%, with an exposure SD of 0.42 among completers.

Report attrition by condition. Balanced completed showings do not make completers comparable across conditions, so report dropout by condition and analyze attrition as in any trial (Bell et al., 2013).

Run it where nobody can see the ledger. Compute allocations on a server and send each participant only their own conditions. This is the analogue of allocation concealment in trials (Schulz and Grimes, 2002), and it also prevents participants from reverse-engineering the design.

Log what the analysis will need. For each participant, store the deck, its order, β, the log probability of the set, the budget, the code version and the random seed. These make the allocation auditable and support inverse-probability weighting (Horvitz and Thompson, 1952) or randomization tests that replay the algorithm (Rosenberger and Lachin, 2016). Model-based analyses with participant and stimulus effects remain valid, because no allocation depends on that participant's own characteristics.

Report the allocation as a randomization method. State the admissibility rule, the budget in bits, the ledger policy and its timeout, and how ties were broken, as CONSORT 2025 (item 17b) and SPIRIT 2025 (item 21b) ask for minimisation, including whether a random element was used (Hopewell et al., 2025; Hróbjartsson et al., 2025). Because every allocation keeps a stated random element, the method meets the ICH E9 call for a random element in every dynamic allocation (ICH, 1998) and the EMA instruction that deterministic schemes be avoided (EMA, 2015). The budget makes that random element a number a reviewer can check.

8 Discussion

Minimisation for sets of conditions is a feedback controller whose one natural parameter is entropy. Most of its ingredients have precedents; what is new is applying them to constrained sets of conditions, enrolled online, with a budget that can be reported and checked.

Single-treatment allocation. Klotz (1978) derived exponential allocation probabilities by maximizing entropy under a balance constraint, re-solved for each patient, and Ryeznik and Sverdlov (2018) extended the rule to several arms with one parameter. The biased coins of Smith (1984) and Ball et al. (1993) are fixed-temperature rules of the same kind, and Wei's (1977) urn is, asymptotically, the unit-temperature rule on the ledger's Jensen–Shannon gain. We make the entropy budget itself the parameter, show what a per-allocation budget costs when there are few arms, and give the exact control identity.

Online and combinatorial balancing. Herding (Welling, 2009; Chen and Welling, 2016) and the self-balancing walk (Alweiss et al., 2021), which Arbour et al. (2022) adapted to online experiments, balance running sums with deterministic or randomized feedback. The Gram–Schmidt walk design trades robustness against balance with one parameter (Harshaw et al., 2024), and best-of-K rerandomization bounds the loss of robustness through $\log K$ (Banerjee et al., 2020). None of these allocates constrained sets per participant, and the best-of-K bound says nothing once the K draws are repeated for every participant; a per-participant entropy budget is the natural replacement.

Stale counts in trials. We found no prior analysis of minimisation when several sites allocate from the same out-of-date counts. Emergency trials sidestep the delay by computing each site's next allocation in advance (Zhao et al., 2010), and response-adaptive designs have been analyzed under delayed responses (Bai et al., 2002). Our delay results also bear on Taves's (2010) view that a random element is rarely needed: under simultaneous enrollment, removing it made minimisation worse than randomization.

Vignette, conjoint and choice designs. Standard practice builds decks offline, for instance with D-efficient or confounded factorial designs, and then randomizes their order (Atzmüller and Steiner, 2010; Dülmer, 2016; Su and Steiner, 2020). The two approaches combine naturally: an offline design's decks can serve as the admissible set, and the allocator then balances how often each deck is used as participants arrive.

Limitations.

1. Shannon entropy is not predictability. With two arms and an imbalance bound of 3, the maximal procedure has more entropy than the big stick (0.89 against 0.83 bits per allocation) but is easier to guess (0.625 against 0.583). Where guessing matters, report a guessing measure as well.

2. The ledger here counts conditions only. Balancing participants' covariates, positions or pairs of conditions needs a richer ledger; any cost that is linear in counts keeps the dice, but weighted costs need re-weighted dice to stay optimal.

3. Exact calibration needs the admissible sets enumerated. Larger designs need Monte Carlo calibration, whose accuracy we have not studied systematically.

4. Plain rejection stalls at very low temperatures, so the implementation falls back to drawing from the enumerated list.

5. The simulations cover one design. The $1/\sqrt{\kappa}$ scaling and the delay results should be checked in others.

6. Allocation probabilities vary from participant to participant. Design-based estimators should use the logged probabilities, and inference for within-participant effects under this design deserves its own treatment.

7. Balanced completed showings are not unbiased estimates. The attrition study shows only that the method corrects unequal condition representation caused by differential attrition; it simulates no outcomes, and when dropout depends on participants' characteristics the completers in each condition can still differ (Zhou and Fishbach, 2016).

Future work. Three directions follow. One is robustness bounds indexed by the per-participant budget; another is non-uniform targets and covariate-adaptive ledgers. The third is ledgers that predict in-progress sessions across sites, which would shorten the effective delay.

Code availability. The implementation and the simulation scripts are at `https://github.com/andrewbistras/min_sim`, in the `paper/` folder, and reproduce every number and figure in this paper. Running `sh run_all.sh` in `paper/sim` reruns all of them.

A Proofs

Throughout, $E(S) = \sum_{i \in S} c_i$ is a deck's total count, U is the uniform law on $\mathcal{A}$, and $\pi_\beta(S) \propto \exp(-\beta E(S))$.

Proposition 1. For any law π on $\mathcal{A}$, expanding $\mathrm{KL}(\pi \,\|\, \pi_\beta) \geq 0$ gives

$$\mathrm{KL}(\pi \,\|\, U) + \beta\, \mathbb{E}_\pi[E] \;\geq\; -\log \mathbb{E}_U\left[e^{-\beta E}\right] \;=\; \mathrm{KL}(\pi_\beta \,\|\, U) + \beta\, \mathbb{E}_{\pi_\beta}[E].$$

So any π that spends no more bits than π_β has an expected total count at least as large, for $\beta > 0$. □

Proposition 2. Every ordered roll of an admissible set S has probability $\prod_{i \in S} \tilde{w}_i$, which is proportional to $\exp(-\beta E(S))$ and so depends on the set alone. Conditioning on admissibility therefore gives the set law π_β and makes all $k!$ orders equally likely. □

Proposition 3. Adding deck S gives counts $c + x(S)$ over $k(t+1)$ showings, so $P_{t+1} = wP_t + (1-w)U_S$ with $w = t/(t+1)$. (a) In JS_w only $H(P_{t+1})$ depends on S. (b) Expanding $H(u+\delta) = \log m - \frac{m}{2}\|\delta\|^2 + O(\|\delta\|^3)$ around the uniform u, and using $\sum_i (c_i + x_i - T)^2 = \mathrm{const} + 2E(S)$ with $T = k(t+1)/m$, gives $\mathrm{JS}_w = \mathrm{const} - mE(S)/(kt+k)^2$. (c) P_t is the equal-weight mixture of $U_{S_1}, \ldots, U_{S_t}$, each with entropy $\log k$. (d) Let $\mathcal{H}$ be the history, $\pi_{\mathcal{H}}$ the law of the next set and $\bar{\pi}$ its average over histories. Then $\mathbb{E}[\mathrm{KL}(\pi_{\mathcal{H}} \,\|\, U)] = I(\mathcal{H}; S) + \mathrm{KL}(\bar{\pi} \,\|\, U) \geq I(\mathcal{H}; S)$, so a per-participant budget κ bounds $I(\mathcal{H}; S)$ by κ. □

Proposition 4. The Bellman equation of the average-cost problem (Todorov, 2009) is $c + v(e) = q(e) - \log \mathbb{E}_U[\exp(-v(e'))]$, with optimal policy $\pi^*(S \mid e) \propto \exp(-v(e'))$. With $v(e) = \frac{\beta}{2}\|e\|^2$ and $\sum_i e_i = 0$, the next value is $v(e') = v(e) + \beta \sum_{i \in S} e_i + \text{const}$, so $\pi^* = \pi_\beta$, and the equation holds for $q = q_\beta$ with the constants absorbed into c. With two arms and $D = n_A - n_B$ the surpluses are $\pm D/2$, and $\mathbb{E}_U[\exp(-\beta e_S)] = \cosh(\beta D/2)$. □

Proposition 5. Let $V(e) = \|e\|^2$; then $\mathbb{E}[\Delta V \mid e] = 2\,\mathbb{E}_\pi[X] + k(1 - k/m)$, with $X = \sum_{i \in S} e_i$. We proceed in three steps.

1. *A good deck always exists.* Under U, X has mean 0, because every stimulus lies in the same number of admissible sets. Its variance is at least $\lambda\|e\|^2$, and $X \le k\|e\|$. The Bhatia–Davis inequality then gives $\min_S X \le -(\lambda/k)\|e\|$.

2. *Fixed β.* The Gibbs law lies within $\log|\mathcal{A}|/\beta$ of the minimum, so $\mathbb{E}[\Delta V] \le -2(\lambda/k)\|e\| + 2\log|\mathcal{A}|/\beta + k$.

3. *Budget κ.* The mixture $\varepsilon\,\delta_{S^*} + (1 - \varepsilon)U$ that spends κ bits has $\mathbb{E}[X] = \varepsilon \min_S X$, and by Proposition 1 the budgeted rule does at least as well. So $\mathbb{E}[\Delta V] \le -2\varepsilon(\lambda/k)\|e\| + k$, where ε depends only on κ and $|\mathcal{A}|$.

In both cases the drift is negative outside a bounded set, and Foster's criterion gives positive recurrence and a finite long-run average of $\mathbb{E}\|e(t)\|$ (Meyn and Tweedie, 2009). □

B Admissible decks in the 2^6 example

The independent sets of the six-cube number 64, 1,824, 30,720, 342,400 and 2,682,432 for sizes 1 to 5. Of the admissible five-sets, 201,376 contain only vignettes with an even number of factors at their higher level, and 201,376 only odd ones; such single-parity sets satisfy the rule automatically.

Under complete randomization over admissible decks, two vignettes at Hamming distance 2 share a deck with probability 0.0060, against 0.0053 for distances 3 to 6. Context effects that depend on pairs should keep that 13% difference in mind. The covariance of the inclusion vector has eigenvalues from 0.055 to 0.117 on the sum-zero subspace, so the stability constant of Proposition 5 is $\lambda = 0.055$.

C Exact calibration by enumeration

Enumerate $\mathcal{A}$ once. For each participant, tally the histogram N_E of $E = \sum_{i \in S}(c_i - \min c)$ over $\mathcal{A}$, and evaluate $\kappa(\beta)$ from (7). Bisect on β until $\kappa(\beta)$ equals the budget.

If the budget is at least $\log(|\mathcal{A}|/N_{E_{\min}})$, set $\beta = \infty$ and draw uniformly among the minimum-total decks. For that draw, pick the r-th such deck in a second pass, with r uniform. Both passes cost one sweep over $\mathcal{A}$ (about 5 ms in the example).

D The candidate-pool approximation

A tempting implementation draws M admissible decks uniformly and picks one with probabilities proportional to $\exp(\gamma \times \text{score})$. It is sampling-importance-resampling (Smith and Gelfand, 1992) aimed at π_β, and it has two knobs, M and γ.

The two knobs are one. With a score that also counted positions (weight 0.075), pools of 10 to 1,000 decks and values of γ from 0.05 to 1 all fell within about 2% of the exact dice curve when indexed by bits spent. The entropy of the pool's softmax, H_{pool}, already measures the budget: $\ln M - H_{\text{pool}}$ was 0.39 nats where the true spend was 0.40.

Two add-ons behaved unexpectedly. A fixed bonus for chosen vignettes raised their counts by a constant of about 1.4 showings rather than by a proportion, 5.7% after 320 participants but 0.5% after 3,000. A small position term (weight 0.075) improved order balance by only about 8% over shuffling the order.

References

Altman, D. G. and Bland, J. M. (2005). Treatment allocation by minimisation. *BMJ*, 330(7495):843.

Alweiss, R., Liu, Y. P., and Sawhney, M. (2021). Discrepancy minimization via a self-balancing walk. In *Proceedings of the 53rd Annual ACM Symposium on Theory of Computing*, pages 14–20.

Arbour, D., Dimmery, D., Mai, T., and Rao, A. (2022). Online balanced experimental design. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pages 844–864. arXiv:2203.02025.

Atkinson, A. C. (2014). Selecting a biased-coin design. *Statistical Science*, 29(1):144–163.

Atzmüller, C. and Steiner, P. M. (2010). Experimental vignette studies in survey research. *Methodology*, 6(3):128–138.

Auspurg, K. and Hinz, T. (2015). *Factorial Survey Experiments*, volume 175 of *Quantitative Applications in the Social Sciences*. Sage.

Bai, Z. D., Hu, F., and Rosenberger, W. F. (2002). Asymptotic properties of adaptive designs for clinical trials with delayed response. *Annals of Statistics*, 30(1):122–139.

Ball, F. G., Smith, A. F. M., and Verdinelli, I. (1993). Biased coin designs with a Bayesian bias. *Journal of Statistical Planning and Inference*, 34(3):403–421.

Banerjee, A. V., Chassang, S., Montero, S., and Snowberg, E. (2020). A theory of experimenters: Robustness, randomization, and balance. *American Economic Review*, 110(4):1206–1230.

Bell, M. L., Kenward, M. G., Fairclough, D. L., and Horton, N. J. (2013). Differential dropout and bias in randomised controlled trials: When it matters and when it may not. *BMJ*, 346:e8668.

Blackwell, D. and Hodges, J. L. (1957). Design for the control of selection bias. *Annals of Mathematical Statistics*, 28(2):449–460.

Charness, G., Gneezy, U., and Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization*, 81(1):1–8.

Chen, Y. and Welling, M. (2016). Herding as a learning system with edge-of-chaos dynamics. In Hazan, T., Papandreou, G., and Tarlow, D., editors, *Perturbations, Optimization, and Statistics*, pages 73–125. MIT Press. arXiv:1602.03014.

Ciolino, J. D., Palac, H. L., Yang, A., Vaca, M., and Belli, H. M. (2019). Ideal vs. real: A systematic review on handling covariates in randomized controlled trials. *BMC Medical Research Methodology*, 19:136.

Clifford, S., Sheagley, G., and Piston, S. (2021). Increasing precision without altering treatment effects: Repeated measures designs in survey experiments. *American Political Science Review*, 115(3):1048–1065.

Collins, L. M. (2018). *Optimization of Behavioral, Biobehavioral, and Biomedical Interventions: The Multiphase Optimization Strategy (MOST)*. Springer.

Dülmer, H. (2016). The factorial survey: Design selection and its impact on reliability and internal validity. *Sociological Methods & Research*, 45(2):304–347.

Efron, B. (1971). Forcing a sequential experiment to be balanced. *Biometrika*, 58(3):403–417.

EMA (2015). European medicines agency: Guideline on adjustment for baseline covariates in clinical trials (EMA/CHMP/295050/2013). `https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-adjustment-baseline-covariates-clinical-trials_en.pdf`.

Gureckis, T. M. et al. (2016). psiTurk: An open-source framework for conducting replicable behavioral experiments online. *Behavior Research Methods*, 48(3):829–842.

Harshaw, C., Sävje, F., Spielman, D. A., and Zhang, P. (2024). Balancing covariates in randomized experiments with the Gram–Schmidt walk design. *Journal of the American Statistical Association*, 119(548):2934–2946.

Hopewell, S. et al. (2025). CONSORT 2025 explanation and elaboration: Updated guideline for reporting randomised trials. *BMJ*, 389:e081124.

Horiuchi, Y., Markovich, Z., and Yamamoto, T. (2022). Does conjoint analysis mitigate social desirability bias? *Political Analysis*, 30(4):535–549.

Horvitz, D. G. and Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 47(260):663–685.

Hróbjartsson, A. et al. (2025). SPIRIT 2025 explanation and elaboration: Updated guideline for protocols of randomised trials. *BMJ*, 389:e081660.

Hu, Y. and Hu, F. (2012). Asymptotic properties of covariate-adaptive randomization. *Annals of Statistics*, 40(3):1794–1815.

ICH (1998). International council for harmonisation, E9: Statistical principles for clinical trials. `https://database.ich.org/sites/default/files/E9_Guideline.pdf`.

Jaynes, E. T. (1957). Information theory and statistical mechanics. *Physical Review*, 106(4):620–630.

Johnson, F. R., Lancsar, E., Marshall, D., et al. (2013). Constructing experimental designs for discrete-choice experiments: Report of the ISPOR conjoint analysis experimental design good research practices task force. *Value in Health*, 16(1):3–13.

Kahan, B. C. and Morris, T. P. (2012). Reporting and analysis of trials using stratified randomisation in leading medical journals: Review and reanalysis. *BMJ*, 345:e5840.

Kim, J. P. and Yang, H.-J. (2024). A novel experimental vignette methodology: SMART vignettes. *Methodological Innovations*, 17(2):111–118.

Klotz, J. H. (1978). Maximum entropy constrained balance randomization for clinical trials. *Biometrics*, 34(2):283–287.

Kuhn, J., Sheldrick, R. C., Broder-Fingert, S., Chu, A., Fortuna, L., Jordan, M., Rubin, D., and Feinberg, E. (2019). Simulation and minimization: Technical advances for factorial experiments designed to optimize clinical interventions. *BMC Medical Research Methodology*, 19:239.

Meyn, S. and Tweedie, R. L. (2009). *Markov Chains and Stochastic Stability*. Cambridge University Press, 2nd edition.

Moore, R. T. and Moore, S. A. (2013). Blocking for sequential political experiments. *Political Analysis*, 21(4):507–523.

Mummolo, J. and Peterson, E. (2019). Demand effects in survey experiments: An empirical assessment. *American Political Science Review*, 113(2):517–529.

Peabody, J. W., Luck, J., Glassman, P., Dresselhaus, T. R., and Lee, M. (2000). Comparison of vignettes, standardized patients, and chart abstraction: A prospective validation study of 3 methods for measuring quality. *JAMA*, 283(13):1715–1722.

Pocock, S. J. and Simon, R. (1975). Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics*, 31(1):103–115.

Rosenberger, W. F. and Lachin, J. M. (2016). *Randomization in Clinical Trials: Theory and Practice*. Wiley, 2nd edition.

Ryeznik, Y. and Sverdlov, O. (2018). A comparative study of restricted randomization procedures for multiarm trials with equal or unequal treatment allocation ratios. *Statistics in Medicine*, 37(21):3056–3077.

Schulz, K. F. and Grimes, D. A. (2002). Allocation concealment in randomised trials: Defending against deciphering. *Lancet*, 359(9306):614–618.

Senn, S. (2002). *Cross-over Trials in Clinical Research*. Wiley, 2nd edition.

Smith, A. F. M. and Gelfand, A. E. (1992). Bayesian statistics without tears: A sampling–resampling perspective. *American Statistician*, 46(2):84–88.

Smith, R. L. (1984). Sequential treatment allocation using biased coin designs. *Journal of the Royal Statistical Society, Series B*, 46(3):519–543.

Spencer, J. (1977). Balancing games. *Journal of Combinatorial Theory, Series B*, 23(1):68–74.

Su, D. and Steiner, P. M. (2020). An evaluation of experimental designs for constructing vignette sets in factorial surveys. *Sociological Methods & Research*, 49(2):455–497.

Taves, D. R. (1974). Minimization: A new method of assigning patients to treatment and control groups. *Clinical Pharmacology & Therapeutics*, 15(5):443–453.

Taves, D. R. (2010). The use of minimization in clinical trials. *Contemporary Clinical Trials*, 31(2):180–184.

Todorov, E. (2007). Linearly-solvable Markov decision problems. In *Advances in Neural Information Processing Systems 19*, pages 1369–1376. MIT Press.

Todorov, E. (2009). Efficient computation of optimal actions. *Proceedings of the National Academy of Sciences*, 106(28):11478–11483.

Wei, L. J. (1977). A class of designs for sequential clinical trials. *Journal of the American Statistical Association*, 72(358):382–386.

Welling, M. (2009). Herding dynamical weights to learn. In *Proceedings of the 26th International Conference on Machine Learning*, pages 1121–1128.

Zhao, W., Ciolino, J., and Palesch, Y. (2010). Step-forward randomization in multicenter emergency treatment clinical trials. *Academic Emergency Medicine*, 17(6):659–665.

Zhou, H. and Fishbach, A. (2016). The pitfall of experimenting on the web: How unattended selective attrition leads to surprising (yet false) research conclusions. *Journal of Personality and Social Psychology*, 111(4):493–504.