capstract:2611.00072v1 [med.PS] 20 Nov 2026

Original research

Floors descended and the time to a better response after a standardised administrative rebuke: a randomised, assessor-blinded trial of exit route

Marguerite O. Oyelaran-Voss[1,†], Ingrid S. Haldorsen[2], Dev R. Chaudhary[3], Lucía M. Ferreyra-Boyd[1], Anselm Ketteridge[1,2]

[1] Department of Clinical and Health Psychology, Aldermoor University, Edinburgh, UK
[2] Centre for Ambulatory Assessment in Psychiatry, Halvorsen Institute for Mental Health Research, Bergen, Norway
[3] Medical Statistics Unit, Tarnside University, Leeds, UK

Correspondence: Dr Marguerite O. Oyelaran-Voss, Department of Clinical and Health Psychology, Aldermoor University, Edinburgh, UK (m.oyelaran-voss@aldermoor.example)

Abstract

Background: The better response to an interpersonal rebuke is commonly reported to occur to people on the staircase as they leave. The French idiom for this experience has been used for two and a half centuries as a figure of speech, and whether the staircase has any causal role has not been tested. We randomised the route by which people left a standardised rebuke.

Methods: In this seven-arm, parallel-group, assessor-blinded trial, 420 adults received a scripted 90-second rebuke from an administrator in an office on level 5 of a university building. They were then randomised (1:1:1:1:1:1:1) to wait in a room on the same level (corridor arm) or in a room 1, 2 or 4 levels below, reached by lift or by the dogleg staircase. Participants pressed a handheld logger when a better thing to have said occurred to them. The primary outcome was time to first press within 30 minutes. Secondary outcomes were return to the office to deliver the response and rated adequacy. The trial is registered with the Behavioural Trials Registry, BTR-2025-1182.

Findings: A first press occurred in 146 of 180 participants (81.1%) in the stairs arms, 105 of 180 (58.3%) in the lift arms and 39 of 60 (65.0%) in the corridor arm. The median time to first press was 1.9 minutes by stairs, 16.9 by lift and 16.2 in the corridor arm. The hazard ratio for stairs against lift was 1.90 (95% CI 1.48 to 2.56); lifts did not differ from the corridor arm (0.83, 0.55 to 1.28). There was no gradient across 1, 2 and 4 levels (hazard ratio per level 1.06, 0.92 to 1.24). The press rate was 48.2 per 100 person-minutes during descent of the staircase and 5.2 afterwards. The odds of returning to the office fell by 31% for each level below it at the time of the press (odds ratio 0.69, 0.56 to 0.85), and the proportion of randomised participants who delivered a response was 28.3% after one level of stairs and 26.7% in the corridor arm. No serious adverse events occurred.

Interpretation: Descending a staircase brought forward the better response, with the effect reached by a single level, and the gain in delivery was offset by the cost of climbing back. Ascent was not studied.

Funding: Hallam Trust for Behavioural Health Research.

Keywords: post-event processing; interpersonal conflict; randomised controlled trial; locomotion; event boundaries; time to event

Introduction

The better response to an unkind remark is generally said to occur to people on the way out. The experience has an idiom in French, *l’esprit de l’escalier*, a usage attributed to Diderot [1], and the idiom is conventionally read as figurative: the staircase stands for the interval between the exchange and the reflection. We are not aware of any study that has treated the staircase as an exposure. The experience has been quantified in an ambulatory study of 225 adults, in which the hazard of the better response was 3.5 times higher in the shower and 4.6 times higher in bed than during ordinary time at home, and in which stairs were not among the coded contexts [2]. Both of those contexts are low in external demand, which makes them poor candidates for an account of what a staircase might do. Descending stairs is a moderately demanding motor task with a regular rhythm and, in most buildings, at least one reversal of heading.

Two literatures suggest that something might happen on a staircase. In event segmentation theory, passing from one event to another updates the working model of the current situation and makes material from the preceding event available for retrieval [3], and walking through a doorway has been reported to cause forgetting of items from the room that was left [4]. If a change of level behaves as an event boundary, the arrival of a rejoinder after a rebuke might be a retrieval consequence of crossing one. Separately, walking has been reported to increase the production of novel responses on divergent thinking tasks relative to sitting [5], and stair descent is a form of walking.

These accounts predict different things. A boundary account predicts that any change of level should raise the hazard of the better response, whether it is made on foot or in a

lift, and that more levels should give more boundaries. A locomotion account predicts an effect confined to the stairs, with the hazard rising during the descent. A third possibility, that the idiom has no causal content and the staircase is merely where people are when the interval has run out, predicts no difference between routes. We tested these in a randomised trial in which the only manipulated feature was the route by which participants left a standardised rebuke and the number of levels they descended. We also recorded whether participants went back to deliver the response, because the staircase that brings a response forward also increases the distance between the participant and the person who should hear it.

Methods

Trial design and oversight

This was a single-centre, seven-arm, parallel-group, randomised trial with 1:1:1:1:1:1:1 allocation and blinded outcome assessment, done at Aldermoor University, Edinburgh, UK. The trial was registered with the Behavioural Trials Registry (BTR-2025-1182) before enrolment, and the protocol and statistical analysis plan were finalised before the database was locked. We report it according to CONSORT 2010 [6]. The pre-specified hypotheses were that stairs would produce a higher hazard of a first better response than lifts (H1), that the hazard would increase with the number of levels descended (H2), and that, within the staircase arms, the hazard would be higher at half-landings, where the direction of travel reverses, than on straight flights (H3). The participants were told that the study concerned form-completion performance and recovery from interruption. The deception was approved by the ethics committee and was resolved at debriefing.

Participants

Adults aged 18 to 65 years were recruited through university and community notice boards and a research volunteer register between 13 January and 12 June 2025. Exclusion criteria were a mobility limitation that precluded stair use, a vestibular or balance disorder by self-report, and previous participation in a study using the same task. Participants who gave written informed consent were paid £25 on completion.

Setting and procedure

The trial was done in a seven-storey university building with a passenger lift and a single dogleg stairwell with two half-flights of 11 steps and a half-landing at each level (floor-to-floor height 3.6 m). Participants arrived at an administrative office on level 5 and completed a two-page form (about 6 minutes). An administrator, one of two trained confederates working to a script, then delivered the Standardised Administrative Rebuke: three statements in a flat tone over 90 seconds, asserting that the participant had used a superseded version of the form, that the reference field should have been checked, and that the form should not be left incomplete in future. The form was the current version and was complete. Rebukes of this kind are mild in comparison with the standard laboratory social stressor [7, 8]; in a pilot of 24 volunteers, the mean state anger rating (0 to 10) after the rebuke was 3.0. The time at which the administrator said that was all was taken as time zero.

At time zero the participant was handed a handheld logger and told to go to a waiting room on a named level and to take the lift or the stairs, as allocated. The logger had one button, a barometric pressure sensor with a resolution of about 0.3 m and a three-axis accelerometer. Participants were instructed to press the button whenever it occurred to

them that there was something better they could have said, and the logger recorded the time and pressure altitude of every press. The location of each press was classified as leaving the office, in the lift, on a straight flight, at a half-landing or in the waiting room from the altitude and accelerometer traces, and was checked against stairwell door sensors. The waiting rooms were identical in furnishing and contents. Participants stayed in the waiting room until 30 minutes after time zero.

After the first press, the display on the logger asked whether the participant wished to go back to the office and say it. The administrator stayed at the desk for the whole observation period and for 10 minutes after it. Participants who answered yes were met by a research assistant and could return by either route. Return was counted when the participant reached the office and delivered the response. At 30 minutes the logger was collected, and a research assistant who had no knowledge of the allocation recorded the text of the response and a state anger rating, and then debriefed the participant.

Interventions

In the corridor arm (control), the waiting room was on level 5, 40 metres from the office on the same corridor. In the lift arms the waiting room was on level 4, 3 or 1 (1, 2 or 4 levels descended), reached by the lift. In the stairs arms the same rooms were reached by the stairwell. In all arms the walk from the office to the lift or stairwell door was about 40 seconds.

Outcomes

The primary outcome was time from time zero to the first press within 30 minutes, with participants who did not press censored at 30 minutes. Secondary outcomes were the press rate in each phase of the route, return to the office to deliver the response, and

the adequacy of the response. Adequacy was rated from the recorded text by two clinical psychology trainees who were blind to arm and location of press, on a 7-point scale of pragmatic adequacy (1, would have worsened the exchange or been irrelevant to it; 7, relevant, proportionate and relationship-preserving, had it been said at the time). The mean of the two ratings was used, and ratings of 5 or more were classed as adequate. Agreement was assessed with the intraclass correlation for the mean of two raters [9]. Adverse events were recorded.

Randomisation and masking

Allocation used permuted blocks of 7 and 14, generated by a statistician with no contact with the participants, and was held in sequentially numbered opaque envelopes, opened by the research assistant at time zero. Participants could not be masked to the route but were not told the hypotheses. The administrators delivered the rebuke before allocation was known to them. Outcome data came from the logger, and the raters were masked.

Sample size

We assumed a cumulative incidence of a first press at 30 minutes of 60% in the control arm. With 60 participants in each of two arms, a hazard ratio of 2.0 between them can be detected with 80% power at a two-sided α of 0.05, from the Schoenfeld approximation $d = (z_{1-\alpha/2} + z_{1-\beta})^2 / \{p_1 p_2 (\ln \mathrm{HR})^2\}$ with $p_1 = p_2 = 0.5$, which gives about 65 events. The primary contrast, pooled stairs arms against pooled lift arms (180 against 180), required about 102 events to detect a hazard ratio of 1.74.

Statistical analysis

All analyses were by allocated arm. Time to first press was analysed with a Cox model, and the three arms were summarised by Kaplan–Meier curves [10–12]. With F_i the number of levels descended, S_i and L_i indicators of stairs and lift, and the corridor arm as the reference,

$$h_i(t) = h_0(t)\exp\left\{\theta_S S_i + \theta_L L_i + \delta_S S_i F_i + \delta_L L_i F_i\right\}. \tag{1}$$

H1 was tested by $\exp(\theta_S + \delta_S \bar{F}) / \exp(\theta_L + \delta_L \bar{F})$ over the pooled arms, and H2 by $\exp(\delta_S)$ and $\exp(\delta_L)$. Hazard ratios are reported as the ratio of cumulative hazards at 30 minutes, which equals the hazard ratio under proportional hazards, because of the small number of events in some arms, with 95% confidence intervals from a nonparametric bootstrap of 2,000 resamples [13]. H3 was assessed by the ratio of press rates per person-minute at risk on half-landings and on straight flights. Rates are reported with Poisson confidence intervals that ignore clustering by participant, since each participant contributes at most one first press.

Return to the office was analysed among participants who pressed, with a logistic model for the probability of return given the number of levels D_i between the office and the participant at the time of the press,

$$\operatorname{logit}\Pr(R_i = 1) = \alpha + \rho D_i. \tag{2}$$

Analyses used R version 4.3.2.

Results

Of 503 people assessed, 83 were excluded (47 ineligible and 36 declined), and 420 were randomised, 60 to each arm (Figure 1). Three participants did not receive the

allocated route: the lift was out of service for two participants (one each in the lift arms for 2 and 4 levels), and the alarm on the stairwell door sounded for one in the stairs arm for 1 level. All three were analysed as allocated. No logger data were lost. The mean age was 29.6 years (SD 9.0), and 247 participants (58.8%) were women (Table 1). The mean brooding score on the Ruminative Responses Scale was 10.7 [14]. The mean state anger rating after the rebuke was 3.4 (SD 2.1), and the arm means ranged from 3.2 to 3.6.

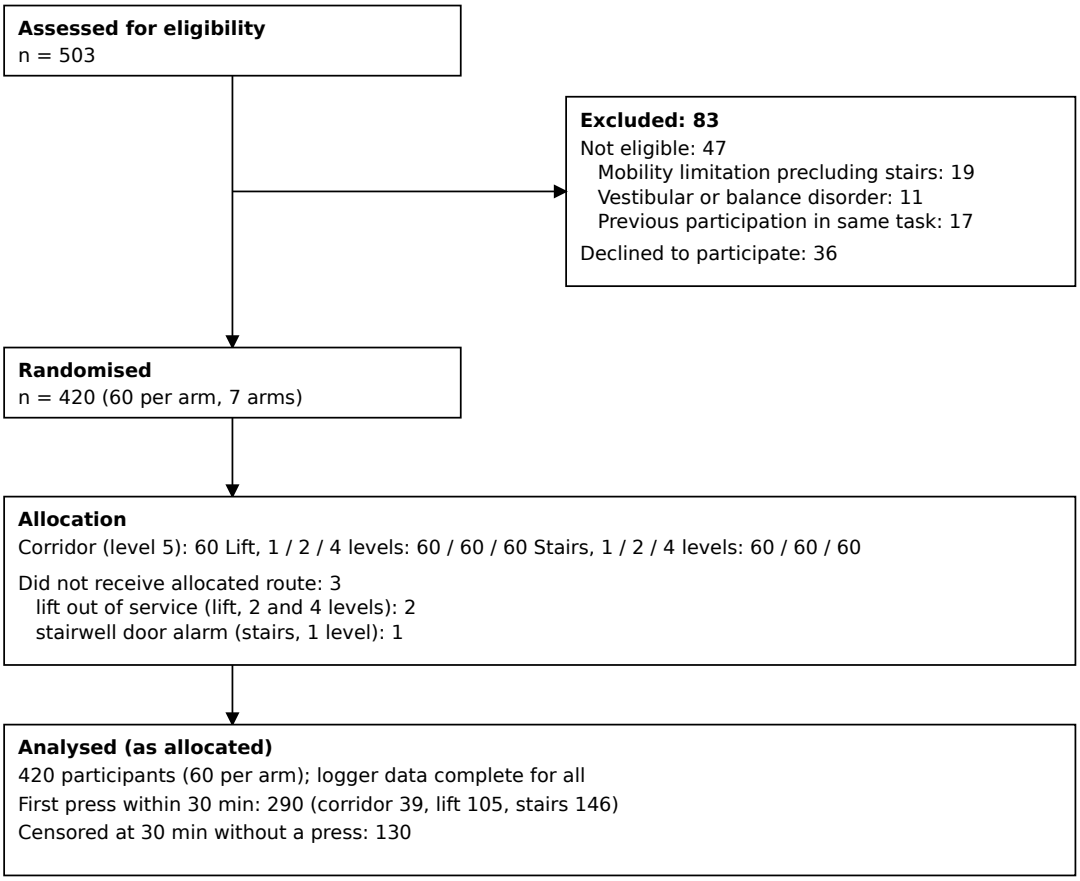


Figure 1. Participant flow.

Table 1. Characteristics of participants by route. RRS = Ruminative Responses Scale, brooding subscale (range 5 to 25).

	Corridor (n=60)	Lift (n=180)	Stairs (n=180)	All (n=420)
Age, years, mean (SD)	29.8 (8.7)	29.5 (9.1)	29.7 (8.9)	29.6 (9.0)
Women	35 (58.3%)	106 (58.9%)	106 (58.9%)	247 (58.8%)
RRS brooding, mean (SD)	10.9 (3.3)	10.6 (3.4)	10.8 (3.5)	10.7 (3.4)
Usual stair use, flights per day, median (IQR)	4 (2–7)	4 (2–7)	4 (2–8)	4 (2–7)
Confederate A	31 (51.7%)	91 (50.6%)	90 (50.0%)	212 (50.5%)

A first press within 30 minutes occurred in 290 of the 420 participants (69.0%): 39 of 60 (65.0%) in the corridor arm, 105 of 180 (58.3%) in the lift arms and 146 of 180 (81.1%) in the stairs arms (Figure 2, panel a). The median time to first press was 16.2 minutes in the corridor arm, 16.9 minutes in the lift arms and 1.9 minutes in the stairs arms. The hazard ratio for stairs against lift was 1.90 (95% CI 1.48 to 2.56; H1). The stairs arms had a higher hazard than the corridor arm (1.59, 1.09 to 2.42), and the lift arms did not (0.83, 0.55 to 1.28). Within each level of descent, stairs against lift gave hazard ratios of 1.92 (1.21 to 3.21) for one level, 2.01 (1.26 to 3.42) for two levels and 1.81 (1.16 to 2.99) for four levels.

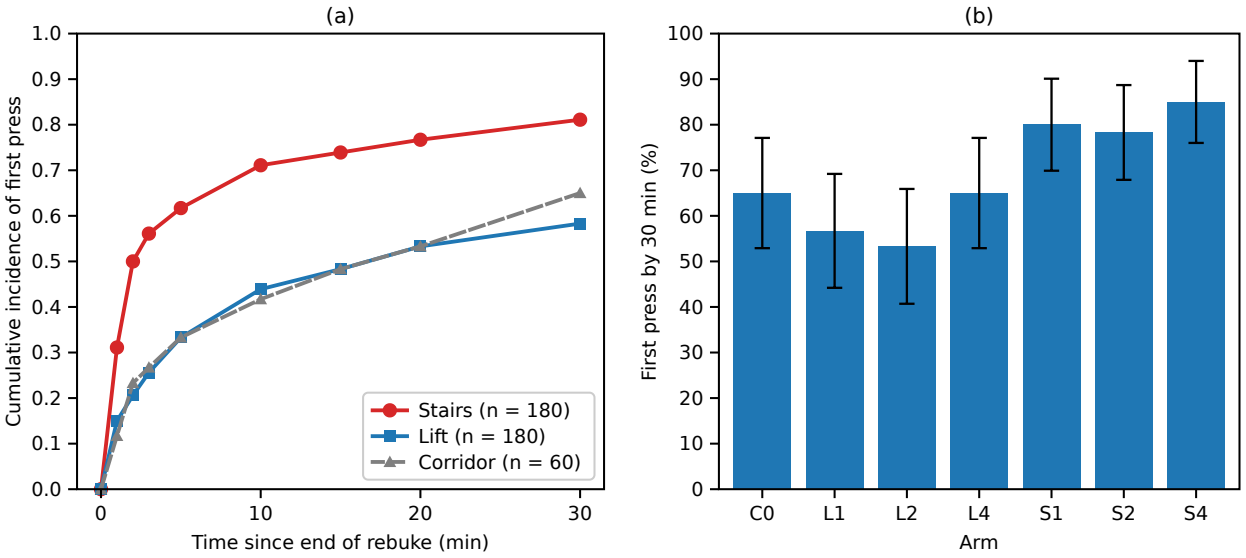


Figure 2. Time to first better response. (a) Cumulative incidence of a first press by route, pooled over levels descended. (b) Percentage of participants with a first press by 30 minutes, by arm, with 95% confidence intervals. C0 = corridor arm; L1, L2 and L4 = lift, 1, 2 and 4 levels; S1, S2 and S4 = stairs, 1, 2 and 4 levels. n = 60 per arm.

The hazard did not rise with the number of levels descended (H2). Per level, the hazard ratio was 1.06 (0.92 to 1.24) in the stairs arms and 1.09 (0.93 to 1.28) in the lift arms. The cumulative incidence at 30 minutes was 80.0%, 78.3% and 85.0% after 1, 2 and 4 levels of stairs, and 56.7%, 53.3% and 65.0% after 1, 2 and 4 levels by lift (Table 2).

Table 2. Outcomes by arm. Press and return are counts of participants (percentage of randomised). Median time is to the first press, in minutes. Adequacy is the mean rating (1 to 7) among participants who pressed.

Arm	Levels descended	First press by 30 min	Median time to first press, min	Returned and delivered	Adequacy, mean
Corridor	0	39 (65.0%)	16.2	16 (26.7%)	3.35
Lift	1	34 (56.7%)	17.8	14 (23.3%)	3.25
Lift	2	32 (53.3%)	27.9	11 (18.3%)	3.47
Lift	4	39 (65.0%)	10.0	7 (11.7%)	2.99
Stairs	1	48 (80.0%)	1.6	17 (28.3%)	3.07
Stairs	2	47 (78.3%)	2.5	12 (20.0%)	3.35
Stairs	4	51 (85.0%)	1.8	13 (21.7%)	3.11

The press rate varied with the phase of the route (Figure 3, panel a). In the stairs arms, the rate while descending was 48.2 per 100 person-minutes at risk (43 presses in 89.2 person-minutes), compared with 25.8 while leaving the office (28 in 108.6) and 5.2 in the waiting room after arrival (75 in 1,440.6). In the lift arms the corresponding rates were 17.7 while leaving the office (20 in 113.2), 10.7 in the lift or waiting for it (14 in 131.4) and 2.6 in the waiting room (71 in 2,754.1). Of the 146 first presses in the stairs arms, 43 (29.5%) were made on the staircase, 29 on straight flights and 14 at half-landings, 28 (19.2%) while leaving the office and 75 (51.4%) after arrival. The rate at half-landings was 72.4 per 100 person-minutes (14 in 19.4) and on straight flights 41.5 (29 in 69.8), a ratio of 1.74 (95% CI 0.92 to 3.29; H3). The rate in the waiting room was 2.0 times higher after stairs than after a lift (95% CI 1.46 to 2.80), which was not constant across the observation period (Figure 2, panel a).

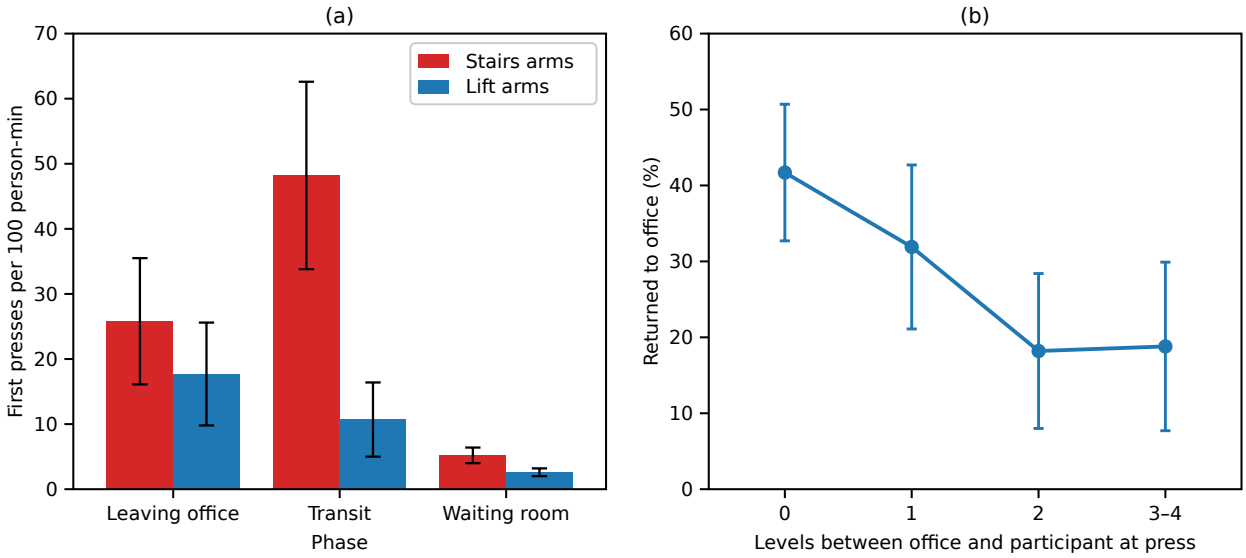


Figure 3. Press rate by phase and return to the office. (a) First presses per 100 person-minutes at risk, by phase of the route, with 95% Poisson intervals. Transit is the staircase in the stairs arms and the lift or waiting for it in the lift arms. (b) Percentage of participants who returned to deliver the response, by the number of levels between the office and the participant at the time of the press, with 95% confidence intervals. n = 115, 72, 55 and 48 participants who pressed.

Of the 290 participants who pressed, 90 (31.0%) went back and delivered the response: 16 of 39 (41.0%) in the corridor arm, 32 of 105 (30.5%) in the lift arms and 42 of 146 (28.8%) in the stairs arms. The proportion who returned fell with the distance from the office at the time of the press (Figure 3, panel b): 48 of 115 (41.7%) at the same level, 23 of 72 (31.9%) one level below, 10 of 55 (18.2%) two levels below and 9 of 48 (18.8%) three or four levels below. Each level reduced the odds of return by 31% (odds ratio 0.69, 95% CI 0.56 to 0.85; equation (2)). As a proportion of those randomised, delivery was 26.7% in the corridor arm, 17.8% in the lift arms and 23.3% in the stairs arms, and by arm it ranged from 11.7% (lift, 4 levels) to 28.3% (stairs, 1 level). The difference between the stairs arm with one level and the corridor arm was 1.7 percentage points (95% CI −14.4 to 17.7).

The mean adequacy rating was 3.21 (SD 0.88), and the intraclass correlation between raters was 0.78 (95% CI 0.73 to 0.82). Ten of the 290 responses (3.4%) were rated adequate, and one of these was delivered. Responses first pressed on the staircase were rated lower than those first pressed in the waiting room (2.93 against 3.37; difference −0.44, 95% CI −0.73 to −0.15, unadjusted). No falls, injuries or other adverse events were recorded. Two participants rated their anger after the rebuke at 8 or above and were seen by the principal investigator at debriefing.

Discussion

Participants who left a rebuke by the stairs reached a better response sooner than those who left by lift. The median time to a first press was 1.9 minutes on the stairs and 16.9 minutes by lift, and the lift did not differ from staying on the same level. The effect

was present after one level and did not increase with two or four. The press rate was highest while participants were on the staircase and remained about twice as high as in the lift arms for some time after arrival.

The absence of an effect in the lift arms argues against a simple event-boundary account, at least as it applies to a change of level. A lift crosses the same boundaries as the stairs, including a door, a change of room and a change of floor, and in these data it did nothing. The doorway effect on forgetting was reported for walking through doors in a room-to-room task [4], and our lift participants walked through two doors without a change in the hazard. The features that distinguish a staircase from a lift are sustained rhythmic locomotion and, in this building, a reversal of heading at each half-landing. The locomotion account fits the pattern during descent. It fits the persistence afterwards less well, since the walking literature reports effects during and shortly after walking [5] and our elevated rate continued for several minutes. The point estimate for half-landings was higher than for straight flights, as the heading account would predict, but the interval included 1 and H3 is not supported. We cannot separate descent from the effort of descent, the vertical displacement itself, or the change of heading, because every stairs participant had all of them.

The lack of a gradient across levels was not what we expected. A staircase effect that was saturated at one level suggests either a threshold or an effect that is largely over by the first half-landing, and in either case the idiom’s usual picture of a long staircase is no better a setting than a short one. This is a result about the building. A different stairwell, with steeper flights, different lighting or a different number of reversals, might give a different gradient.

The delivery result is the one that bears most on the experience the idiom describes. Responses that occurred further from the office were less likely to be delivered, by about a third per level. The stairs brought the response forward by about 15 minutes in the median and moved the participant a level or more away from the person it was meant for at the moment it arrived. The result was that delivery as a proportion of those randomised was no higher in the stairs arm with one level than in the corridor arm, and was lower after four levels by lift than by any other route. Adequacy was low in every arm, and the responses that arrived on the staircase were rated lower than later ones, which is in line with our earlier ambulatory finding that adequacy rises with latency [2].

The trial has limitations. It was done in one building with one staircase, using a mild standardised rebuke by two confederates, and the participants were volunteers, so the findings may not generalise to rebukes that matter to the person receiving them. Participants could not be masked to route. The press is a report of a private event, and participants in the stairs arms may have pressed more readily if they believed the staircase was the point, although they were not told the hypotheses and the press rate on leaving the office, before the route had been used, was already higher in the stairs arms than in the lift arms (25.8 and 17.7 per 100 person-minutes), which would be compatible with anticipation or with chance. Only descent was studied. The idiom’s direction of travel is down, but the trial has nothing to say about whether the effect would be seen on ascent, and no arm controlled for stepping without a change of level. The observation period was 30 minutes, and in our earlier study the median latency was 5.8 hours, so the responses we measured are the early part of a distribution. Finally, the

return option was binary and the administrator stayed at the desk, which is an unusual degree of availability.

Ethics statement

The trial was approved by the Aldermoor University School of Health and Behavioural Sciences Research Ethics Committee (reference AU-SHBS-2025-0112), including the use of deception. All participants gave written informed consent before enrolment. Participants were fully debriefed, were told the purpose of the study, and were offered the chance to withdraw their data, which none did. The administrators were trained in the delivery of the rebuke and in stopping it if a participant became distressed.

Funding

This work was supported by the Hallam Trust for Behavioural Health Research (grant HTB-2024-031). The funder had no role in the design of the trial, the collection, analysis or interpretation of data, or the writing of the report.

Data sharing

De-identified participant-level logger data, with a data dictionary and the analysis code, will be available from the corresponding author on reasonable request after publication. Recorded text of responses cannot be shared. The protocol and statistical analysis plan are available from the trial registry.

Contributors

MOV and ISH conceived the trial. MOV, ISH and DRC designed the protocol and the logger software. LMF-B and AK delivered the rebuke, and LMF-B rated the responses. DRC wrote the statistical analysis plan and did the analysis, and MOV and DRC accessed and verified the data. MOV wrote the first draft. All authors revised the manuscript and approved the final version.

Declaration of interests

We declare no competing interests.

Acknowledgments

We thank the participants, the building services staff who maintained the lift and stairwell sensors, and the two administrators who delivered the rebuke.

References

1. Diderot D. Paradoxe sur le comédien. Paris: Sautelet; 1830.
2. Oyelaran-Voss MO, Haldorsen IS, Chaudhary DR, Ferreyra-Boyd LM, Ketteridge A. Time to generation of the better rejoinder after interpersonal conflict and its relation to trait brooding: a 21-day ecological momentary assessment study. Behaviour Research and Therapy. 2025;188:104711.
3. Zacks JM, Speer NK, Swallow KM, Braver TS, Reynolds JR. Event perception: a mind-brain perspective. Psychological Bulletin. 2007;133(2):273–293.
4. Radvansky GA, Krawietz SA, Tamplin AK. Walking through doorways causes forgetting: further explorations. Quarterly Journal of Experimental Psychology. 2011;64(8):1632–1645.
5. Oppezzo M, Schwartz DL. Give your ideas some legs: the positive effect of walking on creative thinking. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2014;40(4):1142–1152.
6. Schulz KF, Altman DG, Moher D. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340:c332.
7. Kirschbaum C, Pirke K-M, Hellhammer DH. The Trier Social Stress Test: a tool for investigating psychobiological stress responses in a laboratory setting. Neuropsychobiology. 1993;28(1-2):76–81.

8. Dickerson SS, Kemeny ME. Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. Psychological Bulletin. 2004;130(3):355–391.
9. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychological Bulletin. 1979;86(2):420–428.
10. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. Journal of the American Statistical Association. 1958;53(282):457–481.
11. Cox DR. Regression models and life-tables. Journal of the Royal Statistical Society: Series B. 1972;34(2):187–220.
12. Therneau TM, Grambsch PM. Modeling Survival Data: Extending the Cox Model. New York: Springer; 2000.
13. Efron B, Tibshirani RJ. An Introduction to the Bootstrap. New York: Chapman & Hall; 1993.
14. Treynor W, Gonzalez R, Nolen-Hoeksema S. Rumination reconsidered: a psychometric analysis. Cognitive Therapy and Research. 2003;27(3):247–259.