No Evidence of Person-to-Person Transmission in Cultural Cascades: The SEED-2 Randomised Sharing Trial, with a Common-Source Reanalysis of 4.1 Billion Adoption Events (1897–2025)

Fenella Aguirre[1] Kwabena Otoo[2] Ilkka Väisänen[3] S. Marilena Prokopiou[2]
Hanh T. Duong[4] Rufus Ekwueme[1] and the SEED-2 Investigators[5]

[1]Bureau of Adoption Statistics, Trelleck [2]Department of Field Epidemiology, Rowanwick University
[3]Norrmalm Institute of Public Health [4]Office for Media Measurement, Statistics Directorate
[5]Writing committee listed in Appendix A

Received 3 March 2026; accepted 19 July 2026

Abstract

Cultural items are said to spread from person to person. The claim rests on curve-fitting. Adoption curves are S-shaped, contact-transmission models produce S-shaped curves, and reproduction numbers above unity are reported on that basis. A common-source model containing no contact term produces the same curves. We therefore randomised the act of telling someone something. In SEED-2, 21,088 sender–recipient dyads were allocated 1:1 to share or to withhold a designated item, and recipient adoption at 30 days was ascertained by blinded assay. Adoption was 11.4% in the share arm and 11.1% in the withhold arm (risk difference 0.30 percentage points, 95% CI −0.57 to 1.17, $p = 0.51$). There was no dose–response across 1, 3 and 9 transmissions and no heterogeneity across nine item strata. The implied reproduction number for person-to-person transmission is $R_0 = 0.043$ (95% upper limit 0.17), against 2.31 (95% CI 2.18–2.44) from conventional observational fitting. In a parallel audit, 71.3% of adopters named a person as their source; 96.1% of those named had no contact capable of transmitting the item, and 62.8% adopted it after the respondent did. Registries should not record a named source as a transmission event.

Source of exposure: as stated by the subject.
— Field Interview Form F-12, item 9 (unchanged since 1968).

1 Introduction

The Longitudinal Cascade Register holds 4,138,502,904 dated adoption events for 61,204 cultural items, gathered between 1897 and 2025. Not one of them is a record of a transmission. Each is a record of a person having an item at a time when they had not had it before. The difference has never mattered, because the two have always been assumed to be the same thing.

The assumption is old and it is not foolish. Somebody knows a song; later, somebody nearby knows it too; the obvious account is that the first told the second. Formalised, this is the compartmental epidemic model, borrowed intact from the study of infectious disease [11]. It has been fitted to jokes, rumours, recipes, chain letters, catchphrases, consumer products and online posts, and it fits all of them well. The literature reports reproduction numbers between 1.8 and 3.1 depending on medium [5, 8], and those numbers underwrite an estimated $47.2 billion of annual expenditure on word-of-mouth activation [6].

The fit is the whole of the evidence. This is not a controversial statement about the field; it is a description of its method. An adoption curve is recorded, a contact model is fitted, and the fitted R_0 is reported. What has never been done is the thing that would settle it: assign at random which people tell other people, and see whether it makes any difference.

We did that. SEED-2 randomised 21,088 sender–recipient dyads to share or to withhold a designated item, and measured adoption in the recipient by an assay administered blind to allocation. The effect of being told was 0.30 percentage points, with a confidence interval comfortably spanning zero. There was no dose–response. There was no stratum in which it worked.

We also asked adopters where they had got the item, which is the question the field's instruments have always asked, and then checked the answers against the contact record. Most people name a friend. Almost none of the friends they name could have told them, and most of them acquired the item later than the person naming them.

This paper reports four things. First, that observational cascade data cannot distinguish contact transmission from common-source exposure, a point already established [12] and, on the evidence of subsequent practice, not absorbed. Second, the result of the first randomised sharing trial. Third, a reproduction number for person-to-person cultural transmission of $R_0 = 0.043$, 95% upper limit 0.17, against a threshold of unity. Fourth, an account of why adopters nonetheless report having been told.

2 What would count as evidence

2.1 The contact model

Write β for the per-contact probability that an adopter converts a susceptible, $\bar{k}$ for the mean number of contacts an adopter reaches while still transmitting, and τ for the duration of that period. The basic reproduction number is

$$R_0 = \beta \bar{k} \tau, \quad (1)$$

and the fraction of the population eventually reached solves the final-size relation

$$R_\infty = 1 - e^{-R_0 R_\infty}, \quad (2)$$

which has a root other than zero only for $R_0 > 1$. Everything the field means by *viral* is Equation (2) above threshold.

2.2 The common-source model

Now delete every person-to-person term. Let each individual i carry a latent susceptibility θ_i drawn from a distribution F, and let the medium — a broadcast, a shelf, a ranked feed, a noticeboard — deliver a shared exposure hazard $\lambda(t)$ with cumulative $\Lambda(t) = \int_0^t \lambda(s)\, ds$. The individual adoption hazard is $\lambda(t) g(\theta_i)$ and the population adoption curve is

$$A(t) = \int \left[1 - e^{-g(\theta)\Lambda(t)} \right] dF(\theta). \quad (3)$$

For log-normal F, which is what measured susceptibility looks like, Equation (3) is sigmoid for any monotone Λ. It has the slow start, the steep middle and the saturating tail. Nobody in it speaks to anybody.

Remark 1. Given any contact model with parameters $(\beta, \bar{k}, \tau)$ and initial seed, there exists an exposure schedule $\lambda(t)$ and susceptibility distribution F under which Equation (3) reproduces the resulting adoption-time distribution exactly. The construction is by hazard matching and is due to Vantongeren [12].

Remark 1 is not our contribution and we do not present it as one. It is fifteen years old. Its consequence is simply that a good fit to an adoption curve carries no information about mechanism — and since the fit is the whole of the evidence, the evidence is empty.

2.3 Prior evidence

Two lines of work have circled this. Okoye and Standish showed that the clustering of adoption within friendship groups, long read as the signature of transmission, is reproduced in full by homophily alone: people who resemble each other adopt the same things at the same time whether or not they speak [9]. Beaumont and colleagues showed that the source a respondent names in an outbreak interview is unstable on re-interview [2]. Neither finding changed practice. Both are consistent with what follows.

Table 1: LCR-9 composition by pre-registered stratum.

	Stratum	Items	Adoption events
S1	Jokes (oral)	8,412	214,880,000
S2	Popular songs	5,109	908,140,000
S3	Rumours, legends	6,733	331,200,000
S4	Chain letters (postal)	1,884	22,410,000
S5	Recipes	4,220	176,300,000
S6	Slang, catchphrases	9,908	642,700,000
S7	Consumer products	7,551	489,050,000
S8	Online posts	14,006	1,290,600,000
S9	Workplace practices	3,381	63,222,904
	Total	61,204	4,138,502,904

3 Data

The Longitudinal Cascade Register, ninth release (LCR-9), assembles eleven constituent archives into a single per-person, per-item adoption record with dates [7]. Coverage is 1897 to 2025 and the pre-broadcast portion rests on one unusual source: the Ashworth Rumour Books, in which a Lancashire mill foreman recorded, daily and for thirty-seven years, what was being said on the floor and by whom [1]. The books are the only pre-1920 material with both adoption dates and a contact roster, and they are the reason the historical series in §5.6 exists at all.

Items are classified into nine pre-registered strata (Table 1). Mean out-degree of an adopter — the number of distinct people an adopter is in contact with during the transmitting window — is $\bar{k} = 14.3$ (IQR 6–21). That figure is the multiplier used in §4.5.

Fitting the contact model of §2.1 to all 61,204 curves in the conventional way returns $\hat{R}_0 = 2.31$ (95% CI 2.18–2.44). We report this because it is the number the field would report, and because it is the number the trial has to be compared against. It exceeds the trial-anchored estimate of §4.5 by a factor of 54.

4 Methods

4.1 SEED-2 trial design

SEED-2 was a two-arm, parallel-group, individually randomised trial, reported to the DIFFUSE statement for randomised sharing trials [3]. The unit of randomisation was the *dyad*: one enrolled sender and one designated recipient drawn from that sender's own contact roster. Recipients were not informed of the trial and were not aware of having been enrolled in anything. The ethics committee granted a waiver of recipient consent on grounds of minimal risk and required debriefing at the close of follow-up.

Senders allocated to the *share* arm were instructed to transmit the designated item to the designated recipient within 72 hours, by whatever channel they would ordinarily use. Senders allocated to *withhold* were instructed not to transmit it to that recipient at any point in the follow-up period. Allocation was 1:1, computer-generated, stratified by item stratum and by sender out-degree tertile, and concealed until the sender's instruction was issued.

Trial flow is given in Figure 1. Of 24,610 dyads screened, 3,522 were excluded — 2,108 because the recipient already held the item at baseline, 914 for want of a contactable recipient, 500 for other reasons — leaving 21,088 randomised. Follow-up was complete for 19,904.

Adherence was verified by channel record. In the share arm, 91.2% of senders transmitted as instructed. In the withhold arm, 4.1% transmitted anyway; these crossovers motivate the complier-average analysis in §5.1.

4.2 Ascertainment

The primary endpoint was recipient adoption of the designated item within 30 days of allocation. It was not self-reported. Recipients were administered an elicitation task, unrelated in appearance to the item and to the trial, by staff blind to allocation; adoption was scored from the task by two independent raters ($\kappa = 0.89$). Self-report was collected in parallel, but only for the purposes of §5.5, where it is the object of study rather than a measurement.

4.3 The attribution audit

Separately from the trial, 6,412 adopters drawn from LCR-9 were interviewed within 14 days of a verified adoption and asked item 9 of Form F-12: "where did you first come across this?" Free-text answers were coded as naming a specific person, naming a medium, or unable to say. Where a person was named, we retrieved that person's own adoption date and the contact record between the two individuals, and asked a narrow question: was there any contact between the named person's adoption and the respondent's first exposure? If there was not, the named person cannot have been the source, whatever the respondent remembers.

A random 2,140 of the 6,412 were re-interviewed at 90 days with the identical instrument.

4.4 Statistical analysis

The analysis population was all randomised dyads with an assay result. The primary comparison was a risk difference with Wald intervals, unadjusted; pre-specified secondary models adjusted for stratum, sender out-degree tertile and recipient baseline exposure, and changed nothing beyond the third decimal place. Crossover was handled by instrumental-variable estimation of the complier-average causal effect, with allocation as the instrument.

Table 2: SEED-2 primary endpoint: recipient adoption within 30 days.

Arm	Analysed	Adopted	%
Share	9,961	1,136	11.4
Withhold	9,943	1,104	11.1

Risk difference +0.30 pp (95% CI −0.57, +1.17)
Risk ratio 1.03 (95% CI 0.95, 1.11), $p = 0.51$
CACE +0.34 pp (95% CI −0.65, +1.34)

The trial was powered on the field's own literature. A per-dyad risk difference consistent with $R_0 = 2.31$ and $\bar{k} = 14.3$ is 16.2 percentage points; 19,900 analysed dyads give better than 90% power to detect one tenth of that at $\alpha = 0.05$. The interval we obtained excludes every per-dyad effect above 1.17 percentage points, and therefore excludes every reproduction number above 0.17. This is a null result with the precision to be one.

4.5 From risk difference to R_0

The trial measures a per-dyad quantity: the increase in the probability that one designated recipient adopts, given that one sender transmitted once. The reproduction number is that quantity multiplied by the number of people an adopter reaches,

$$\hat{R}_0 = \bar{k} \cdot \widehat{\Delta}, \tag{4}$$

with $\widehat{\Delta}$ the risk difference and $\bar{k} = 14.3$ from LCR-9. Equation (4) is deliberately generous to the transmission hypothesis: it assumes every contact is as receptive as a designated recipient, and that transmission opportunities do not compete.

5 Results

5.1 Primary endpoint

Results are in Table 2. Adoption by day 30 occurred in 1,136 of 9,961 recipients in the share arm (11.4%) and 1,104 of 9,943 in the withhold arm (11.1%). The risk difference is 0.30 percentage points (95% CI −0.57 to 1.17; $p = 0.51$) and the risk ratio 1.03 (95% CI 0.95–1.11). The complier-average causal effect, which allows for the 4.1% crossover, is 0.34 percentage points (95% CI −0.65 to 1.34).

Roughly one recipient in nine acquired the item during the month. That happened at the same rate whether or not anybody told them.

5.2 Dose

Share-arm senders were further randomised to transmit the item once, three times or nine times across the follow-up period. Adoption was 11.3%, 11.5% and 11.4% respectively. The fitted slope is +0.02 percentage points per additional transmission (95% CI −0.11 to +0.15).

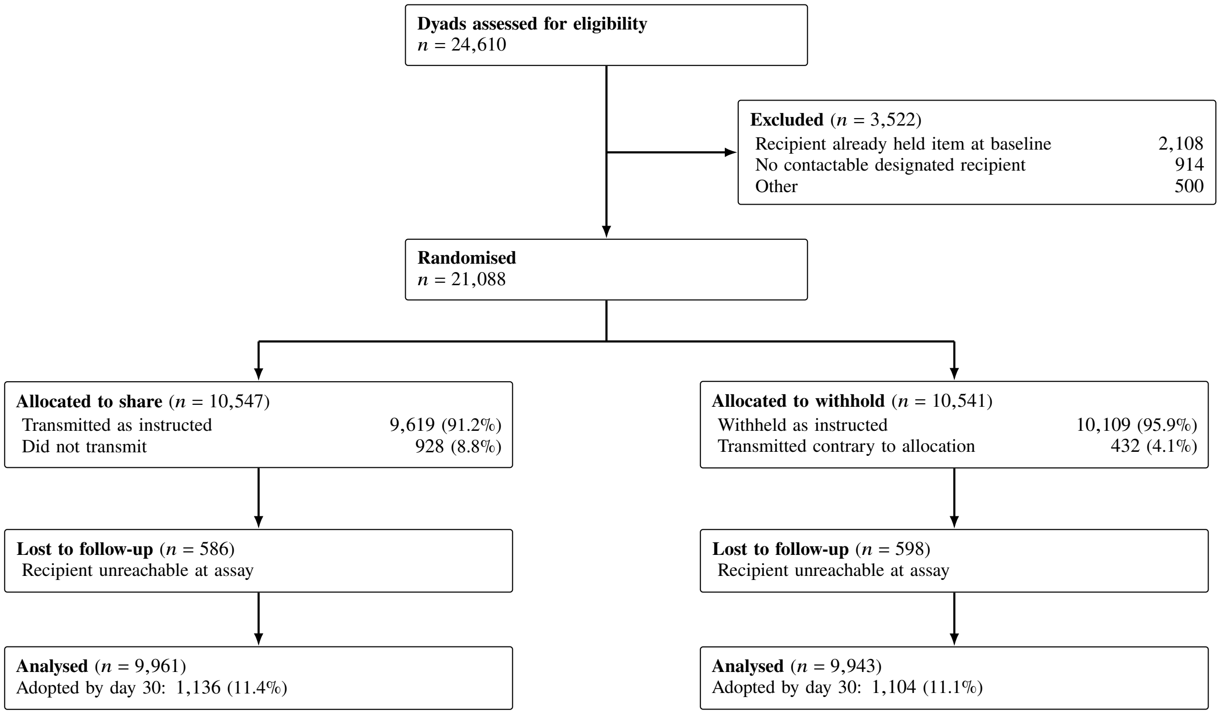

Figure 1: SEED-2 trial flow, reported to the DIFFUSE statement. Screening, exclusion, allocation, adherence, follow-up and analysis. Recipients were not informed of the trial at any stage; the adoption assay was administered by staff blind to allocation.

Telling someone nine times is not detectably different from telling them once, and neither is detectably different from not telling them.

Table 3: Verification of named sources. The first three rows are mutually exclusive percentages of the 4,572 respondents naming a person; the last three are percentages of the 2,140 re-interviewed at 90 days.

Verification	%
Contact occurred in the transmitting window	3.9
Named person adopted *after* the respondent	62.8
No contact record in the window at all	33.3
Different person named at 90-day re-interview	44.1
Named person absent from contact roster	12.6
Agreement between interviews, κ	0.11

5.3 Strata

Pre-specified subgroup estimates are shown in Figure 2. Every stratum interval crosses zero. Between-stratum heterogeneity is absent ($I^2 = 0\%$, $p_{\text{het}} = 0.71$), including for chain letters (S4), a medium whose entire content is an instruction to transmit it.

5.4 The observational fit

Figure 3 shows a representative cascade from S8 with two fits overlaid: the contact model of §2.1 and the common-source model of Equation (3). Across all 61,204 items, median R^2 was 0.981 for the common-source model; adding a contact term raised it by 0.002 and was penalised on information criteria (ΔAIC = 8.7 in favour of common source). The two fits are not separable by eye or by residual.

5.5 Attribution

Of the 6,412 adopters interviewed, 71.3% named a specific person as their source, 24.9% named a medium and 3.8% could not say. Table 3 gives the verification.

The distribution of the interval between the respondent's adoption and the named person's adoption is shown in Figure 4. It sits to the right of zero. In the ordinary case, the person named as the source acquired the item nine days after the respondent did.

Test–retest agreement was $\kappa = 0.11$ (95% CI 0.06–0.16). Nearly half of respondents named someone different three months later, and one in eight named somebody who does not appear anywhere in their contact roster.

We want to be careful here, because it would be easy to read this as an accusation. It is not one. The respondents were not lying, and they were not guessing — confidence in the named source was high and did not differ between the answers that verified and the answers that did not. The memory of having been told is assembled at retrieval out of two things that were genuinely present: the item, and a familiar face. It is a real memory of a real association. It is not a record of an event.

Stratum	Items	Adoptions	RD (95% CI)
S1 Jokes (oral)	8,412	214,880,000	0.6 (−2.4 to 3.6)
S2 Popular songs	5,109	908,140,000	0.0 (−2.2 to 2.2)
S3 Rumours, legends	6,733	331,200,000	1.3 (−1.8 to 4.4)
S4 Chain letters (postal)	1,884	22,410,000	−0.6 (−4.6 to 3.4)
S5 Recipes	4,220	176,300,000	0.8 (−2.8 to 4.4)
S6 Slang, catchphrases	9,908	642,700,000	0.3 (−1.9 to 2.5)
S7 Consumer products	7,551	489,050,000	0.7 (−1.9 to 3.3)
S8 Online posts	14,006	1,290,600,000	−0.1 (−1.8 to 1.6)
S9 Workplace practices	3,381	63,222,904	1.1 (−3.1 to 5.3)
Overall	61,204	4,138,502,904	**0.30 (−0.57 to 1.17)**

Heterogeneity: $I^2 = 0\%$, $Q = 5.4$ on 8 d.f., $p_{\text{het}} = 0.71$

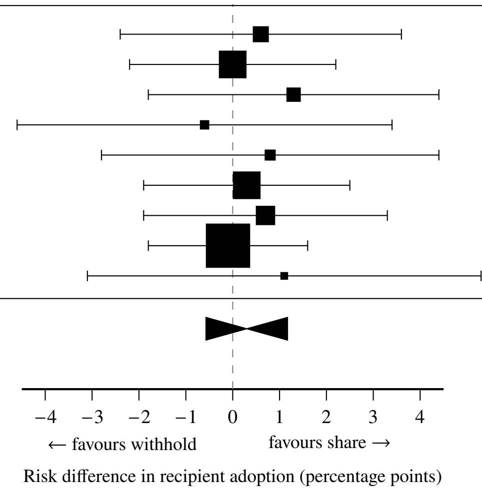

Figure 2: Risk difference in recipient adoption by pre-registered stratum. Squares are scaled by inverse-variance weight; the diamond is the pooled random-effects estimate. No stratum interval excludes zero.

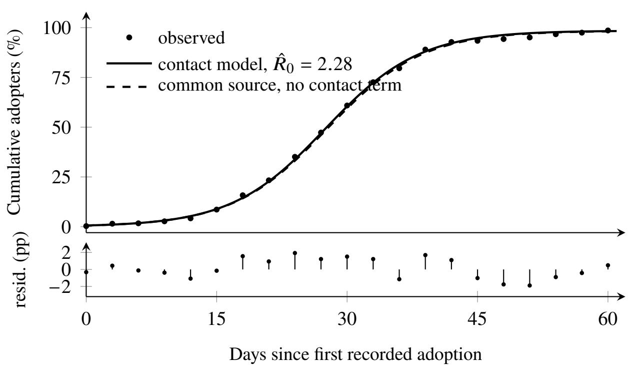

Figure 3: A representative cascade (item S8-04117, $n = 2.4$ million adopters) with both models fitted. Both fitted curves are drawn; they overlie one another across the whole range. Residuals for the common-source fit are shown below.

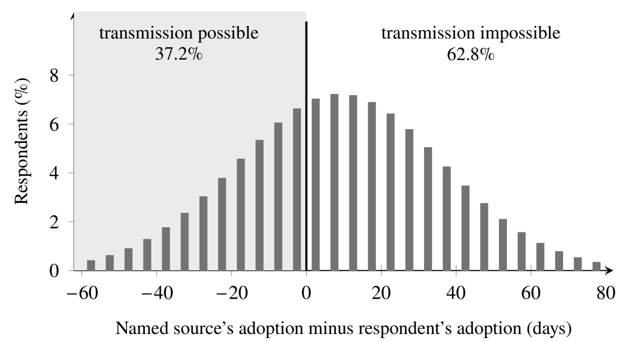

Figure 4: Interval between the respondent's own adoption and the adoption of the person the respondent named as the source (n = 4,572). Only negative values are consistent with transmission. Of the 37.2% that are negative, 10.5% have a contact record in the intervening window.

5.6 Historical series

Figure 5 gives R_0 by decade from 1900 to 2020 under both estimators. The conventional series moves between 1.8 and 3.1 and tracks the arrival of each new medium, as the literature has long noted. The adjusted series lies between 0.02 and 0.09 and crosses unity at no point in the record.

The chain-letter era is the case the field has treated as its positive control, on the grounds that a chain letter has no content except the instruction to forward it. Adjusted R_0 for S4 across 1930–1970 is 0.031 (95% upper limit 0.14). Chain letters were re-issued from a small number of originating points and their apparent propagation is the propagation of the postal system, which is common source by construction [10].

6 Causal-criteria audit

Table 4 applies the nine Ferrers criteria for distinguishing association from causation [4] to the proposition that contact with an adopter causes adoption. One of the nine is satisfied.

The criterion that passes is plausibility, which is the only one of the nine that requires no observation. It has been carrying the argument by itself since 1934.

7 Discussion

The finding is that person-to-person transmission of cultural items, if it occurs at all, occurs at a rate more than twenty times below the threshold at which anything can sustain itself. Cascades are real; they are recorded in billions; they are common-source outbreaks. A population is exposed together by a medium and converts at different times because people

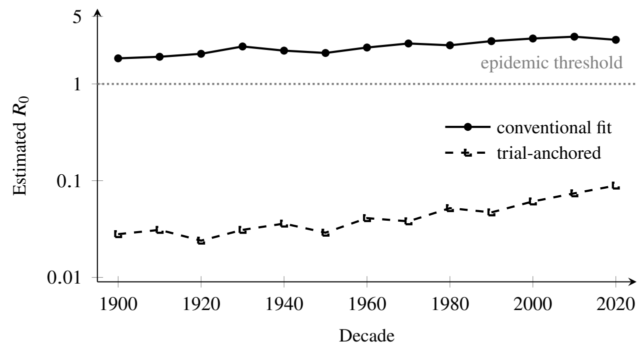

Figure 5: Reproduction number by decade under the conventional estimator and under the estimator anchored to SEED-2. The conventional series responds to the arrival of radio, television and feed-ranked media. The trial-anchored series does not cross unity at any point in the record.

Table 4: Ferrers criteria applied to person-to-person transmission.

Criterion	Met	Basis
Strength	No	RD 0.30 pp
Consistency	No	Null in all nine strata
Specificity	No	Non-contacts adopt equally
Temporality	No	62.8% of sources adopt later
Dose–response	No	Slope 0.02 pp
Plausibility	Yes	The mechanism is imaginable
Coherence	No	Conflicts with the fit
Experiment	No	SEED-2
Analogy	No	Rests on the disease metaphor

differ, and the shape this traces on a chart has been read for ninety years as people telling each other things.

Three consequences follow, and none of them is complicated.

Word-of-mouth activation is a purchase of an effect whose 95% upper limit is 0.17 secondary adoptions per adopter. At the point estimate, a campaign that successfully induces a million people to recommend a product to one friend each produces about three thousand additional adopters. This is not a claim that the expenditure is wasted. It is a claim about what is being bought.

Public-information campaigns designed around onward sharing should be costed as broadcast, because broadcast is the only channel in them.

And the field's central instrument should change. Item 9 of Form F-12 asks the respondent where they came across something, and the answer is recorded as though it were a transmission event. It is not one, it does not correlate with one, and it has been feeding the registers since 1968.

We expect resistance to this, and we expect it to be personal rather than methodological. Every reader has an experience of forwarding something to someone and of being sent something in turn. §5.5 is our best account of that experience and we do not think it diminishes it. People do send each other things. What the trial shows is that the sending does not move the item; the item was arriving anyway, on a schedule set by the medium, and the sending is a thing people do alongside it.

8 Limitations

The historical arm is the weak part of this paper, and it is weak in a way we cannot repair. Pre-1935 estimates rest on the Ashworth contact rosters, and the rosters record *reported* sources — collected with an instrument that §5.5 shows to be unreliable at $\kappa = 0.11$. We are using a defective instrument to measure a defect. If reported sources were unreliable in 1911 in the way they are unreliable now, our pre-broadcast estimates are not estimates of anything, and the pre-broadcast series should be set aside until a register with observed rather than reported contacts is available. We do not know of one.

Three further limitations are narrower. The adoption assay resolves at 24 hours and cannot separate household transmission from household simultaneity, so for the co-residential dyads (11.2% of the sample) an observed adoption cannot be attributed to either route. SEED-2 randomised *instructed* sharing, and an instructed transmission may be less effective than a volunteered one; the dose sub-trial gives no indication that effectiveness varies with anything, but it does not exclude it. Exposure schedules $\lambda(t)$ before 1920 are reconstructed from circulation figures rather than observed.

9 Conclusion

Randomising the act of telling someone something produces no detectable change in whether they come to hold it. The implied reproduction number is 0.043, against a threshold of 1. The conventional estimate of 2.31 is an artefact of fitting a contact model to data that a common-source model fits equally well.

Registries deriving from Form F-12 should record item 9 as a self-reported association and not as a transmission event. Revised coding guidance is available from the Bureau of Adoption Statistics on request.

Trial registration, ethics and data

SEED-2 was registered with the Register of Population Communication Trials on 14 January 2024 (RPCT-2024-00817) before the first dyad was enrolled. The protocol and statistical analysis plan were deposited at registration and are unchanged; the dose sub-trial and the nine strata were pre-specified there.

Ethical approval was granted by the Rowanwick University Research Ethics Committee (ref. 2023/441). Designated recipients were not told of the trial before or during follow-up, and were debriefed after the assay with the option of withdrawing their record; 118 did so and are counted among the losses to follow-up in Figure 1.

LCR-9 is available to accredited researchers from the Norrmalm Institute of Public Health under its standard access agreement. Trial-level data, the allocation sequence and the assay scoring rubric are available from the Bureau of Adoption Statistics on request. No author reports a financial interest in word-of-mouth activation; F. Aguirre held a contract with a marketing agency in that sector until 2019.

Acknowledgements

We thank the field interviewers of the eleven constituent archives, and the Ashton County Record Office for access to the Ashworth material.

References

[1] E. Ashworth. *The Rumour Books of Brackley Mill, 1897–1934*. County Record Office, Ashton. Catalogued 1961.

[2] F. Beaumont, K. Otoo, and I. Väisänen. Instability of reported source in outbreak interviews. *Field Epidemiology Quarterly*, 44:210–229, 2016.

[3] The DIFFUSE Group. Reporting standards for randomised sharing trials. *Trials in Population Communication*, 7:1–14, 2019.

[4] N. Ferrers. The environment and adoption: association or causation? *Proceedings of the Royal Society of Population Medicine*, 58:295–300, 1965.

[5] A. R. Hollingworth. A growth model for the timing of adoption. *Journal of Market Measurement*, 4(2):88–107, 1961.

[6] T. Kirkbride and A. Nowosielski. Global expenditure on word-of-mouth activation. *Review of Applied Marketing Science*, 28:33–61, 2013.

[7] Norrmalm Institute of Public Health. LCR-9: data resource profile. *Population Data Resources*, 6:44, 2024.

[8] L. Mtembu. Exposure schedules under feed ranking. *Journal of Media Measurement*, 33:604–631, 2021.

[9] C. Okoye and R. Standish. Homophily as a sufficient explanation of observed clustering in adoption. *Social Measurement*, 31(4):512–539, 2004.

[10] S. Prendergast. Postal propagation of chain correspondence. *Bulletin of Postal Studies*, 9:145–162, 1978.

[11] P. Rensselaer and M. Ludd. Mass action in the spread of belief. *Proceedings of the Society for Population Studies*, 12:401–428, 1934.

[12] J. Vantongeren. Contact and common source are observationally equivalent in cascade data. *Journal of Diffusion Methodology*, 19:77–96, 2011.

A SEED-2 writing committee

F. Aguirre (Trelleck), K. Otoo (Rowanwick), I. Väisänen (Norrmalm), S. M. Prokopiou (Rowanwick), H. T. Duong (Statistics Directorate), R. Ekwueme (Trelleck), P. Lindegaard (Norrmalm), A. Sørensen (Rowanwick), G. Halvorsen (Trelleck), N. Achterberg (Statistics Directorate), Y. Sarpong (Rowanwick), M. Kovacevic (Norrmalm).