Collapsing Waste Fractions Without Reducing Sorting Accuracy

Anton P. Wrisberg (anton.wrisberg@lucs.lu.se)[1], Annika Wallin[1], & Philip Pärnamets[2]
[1]Division of Cognitive Science, Dept. of Philosophy, Lund University
[2]Division of Psychology, Dept. of Clinical Neuroscience, Karolinska Institutet

Abstract

Waste sorting signs present an inviting avenue for comparison of categorisation theories as they often contain both a rule and either an abstract prototype or visual exemplars. Focus in waste sorting studies is often on increasing the number of waste fractions, but some public settings might require fewer-than-ordinary fractions. This paper investigates two approaches to reducing a waste sorting system's granularity at a music festival: explicit combining waste fractions by merging waste signs and implicitly combining fractions by removing alternatives to 'residual waste'. The digital sorting results are modelled as a probit regression allowing direct interpretation in terms of signal detection theory's senstitivy and criterion parameters. The paper further compares four alternative waste sorting signs including visual exemplars to improve waste sorting accuracy for the implicitly combined waste fraction. The results suggest that explicitly combining waste fractions outperform implicit combinations – even with visual exemplars added, despite resulting in more false positives.

Keywords: categorisation; prototype theory; exemplar theory; signal detection theroy; conceptual spaces; waste sorting; applied research

Background

Waste sorting is no trivial cognitive task, but a naturalistic categorisation task under changing task rules. To be successful, one must: i) identify the waste item in question, ii) retrieve the categories and governing rules of the current waste sorting scheme (e.g., the ones presented in Figure 1), iii) retrieve attributes of the item at a level corresponding to the scheme's governing rules, and iv) judge if the retrieved attributes of the item sufficiently match the rules of the waste fraction in question. While people are good at identifying objects and materials (see e.g., Grill-Spector & Kanwisher, 2005; Sharan et al., 2014), ambiguity can arise when evaluating attributes against incomplete or conflicting sorting scheme knowledge. Especially as waste-sorting schemes vary across contexts and over time (Krishnan et al., 2026).

This categorisation challenge can be conceptualised as a signal detection task (for a brief introduction, see Macmillan, 2001). Doing so provides a principled account of participants' ability to discriminate (their *sensitivity*) items correctly belonging to a fraction (*signal* trials) from items not belonging to a fraction (*noise* trials) and the evidence participants require for accepting an item as belonging in a fraction (their *criterion*). Signal detection theory has long been used to quantify cognitive performance (see e.g., Tanner & Swets, 1954) and been subject to many qualified extensions (see e.g., Erev, 1998; Jang et al., 2009). In domains neighbouring the one of this paper, signal detection theory has recently been used to compare text and pictograms for recognition and categorisation of food and kitchenware (Butorova et al., 2022) and to quantify the effects of waste bin colour changes on detection in urban scenes (Abdel Sater et al., 2023).

In event settings and public spaces, pragmatics may restrict the number of waste fractions practically possible. Sorting into fewer fractions could immediately appear less cognitively taxing. But if participants are used to more granular sorting schemes, reduced granularity leads to bigger intra-fraction variance, which, in signal detection theoretic terms, leads to noisier signals. This can impact both the collapsed and the remaining fractions. In practice, downstream waste handling systems decide the preference for either improving hit rates and/or reducing false alarm rates. Understanding how collapsing of fractions affect participants sensitivity and criterion is thus instrumental to ensuring that infrastructure pragmatics does not unnecessarily restrict material recovery.

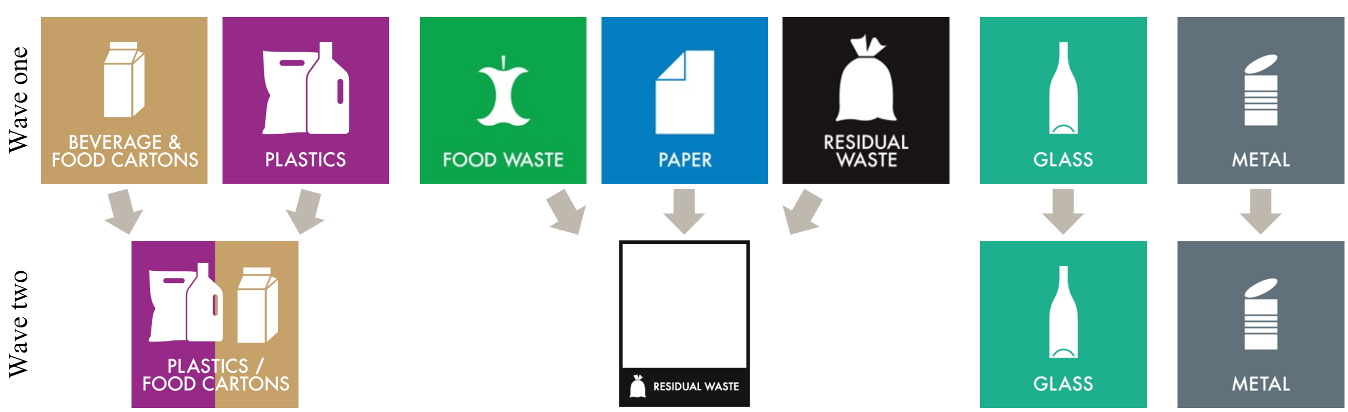

Figure 1 Overview of waste fractions sorted into in waves one and two of the experiment. The fractions 'beverage & food cartons' and 'plastics' where explicitly collapsed into the fraction 'plastics / food cartons' in wave two. The fractions 'food waste' and 'paper' where covertly absorbed into 'residual waste' with space for exemplars in wave two.

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

The present work is part of a research programme investigating the underlying cognitive mechanisms of why exemplars on signs improve waste sorting accuracy (also previously found by e.g., Austin et al., 1993; Cakanlar et al., 2025; Miller et al., 2016; Rosenthal & Linder, 2021; Wu et al., 2018). Here we focus on the collapsing of fractions – meaning when multiple existing fractions are simplified into a less granular categorisation system (see Figure 1). In this paper, we investigate how two options for collapsing fractions affect sorting accuracy and participants' sensitivity and criterion. We evaluated this digitally at Roskilde Festival (DK) in two waves using the EUpicto waste pictogram system (EUpicto, 2020) used in Denmark. EUpicto represents waste fractions with the waste fraction's name (the governing rule) below a stylised exemplar/prototype as a white pictogram on a distinctly coloured quadrant (see examples in Figure 1). Collapsing option one is explicitly mixing fractions by using a combined sign, as seen in Figure 1; from the two individual fractions 'beverage & food cartons' and 'plastics' to the mixed 'plastics / food cartons' fraction sign. Option two is a covert combination of fractions by removing individual fractions whose members, by virtue of the logic of the waste sorting system, then belongs in 'residual waste'. We additionally compare alternative ways of adding high-fidelity exemplars in the form of visual images of actual waste items to the residual waste sign as a way of boosting sorting performance.

Method

Participants

We recruited 766 participants at the camping grounds of Roskilde Festival in two waves in two consecutive years. 404 participants (194 female, 201 male, 9 undisclosed, age $M = 24.9$, $SD = 7.0$) in wave one and 362 participants (160 female, 188 male, 14 undisclosed, age $M = 28.1$, $SD = 10.3$) in wave two. Excluding wave one participants who sorted less than 50% correctly ($n = 8$) and wave two participant who failed any one attention checks ($n = 89$) left us with 669 participants (311 female, 340 male, 18 undisclosed, age $M = 26.2$, $SD = 8.6$).

Procedure

In a setup designed for and accessed by participants on their smartphones, we showed participants a waste item below one of the Danish waste fraction signs (see Figure 2). Participants were instructed to swipe right (left) to accept (reject) the item for this waste fraction. After swiping the item left or right, a new item occurred while the fraction sign stayed.

Wave one participants initially judged their own sorting ability on a 7-point Likert-scale. Afterwards, they sorted as many items as possible within a 30 second timeframe for five of the nine most frequent fractions sorted into in Denmark ('cardboard', 'food waste', 'beverage & food cartons', 'glass', 'hazardous waste', 'metal', 'paper', 'plastics', and 'residual waste'; see Figure 1, top row). The participant-specific selection was randomly ordered but weighted so that independent of fraction, each item would appear approximately equally often in signal trials between participants.

Wave two participants saw only the fractions used at the festival site ('glass', 'metal', 'plastics / food cartons', and 'residual waste'; see Figure 1, bottom row), in that order. For the first three fraction, participants sorted 20 items with the same sign style as in wave one (see Figure 1). For 'residual waste', the sign allowed space for exemplars (see Figure 2) and participants sorted 40 items under one of five conditions: no exemplars, five positive easy-to-sort exemplars, five positive difficult-to-sort exemplars, five negative easy-to-sort exemplars, or five negative difficult-to-sort exemplars. The exemplars were drawn at random among items identified in wave one as having either a high (easy-to-sort) or low (difficult-to-sort) likelihood of being accepted into the right fraction (correlating with prototypicality, as shown by Wrisberg et al., 2026). Negative exemplars appeared beneath a red barred circle (see Figure 2, right). For 'metal' and 'residual waste', attention checks occurred asking participants to sort it as either a signal or a noise trial.

After sorting, participants answered a series of background questions (see Online Supplements: osf.io/e2rdj).

Stimuli

To increase ecological validity (often neglected in categorisation studies; see e.g., Ashby et al., 1994; Hu & Nosofsky, 2021; Rouder & Ratcliff, 2006), we used a subset ($n = 536$) of the Lund Everyday Waste Image Set (Wrisberg et al., 2026) selected to ensure that the items would likely appear at the festival site (see examples in Figure 2).

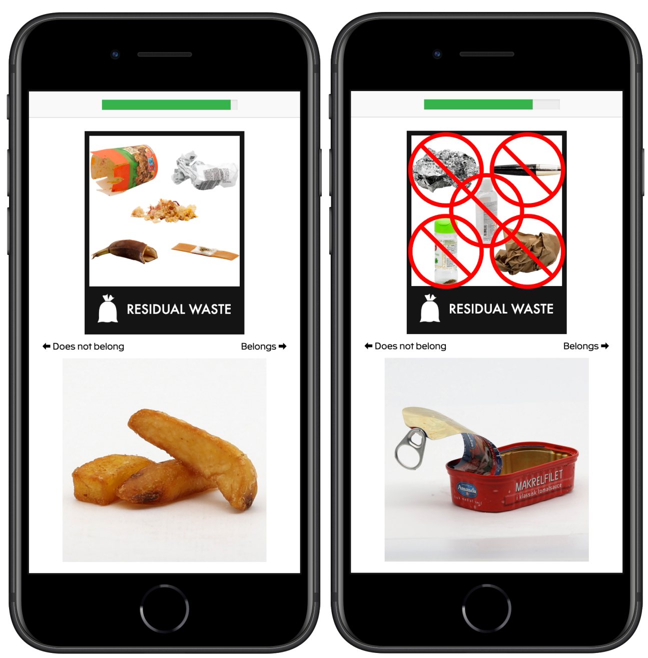

Figure 2 Display presented to participants. Left: a 'food waste' item correctly belonging in the covertly combined 'residual waste' fraction in wave two with the sign showing positive, easy-to-sort exemplars. Right: a 'metal' item not belonging to 'residual waste' in wave two with the sign showing negative difficult-to-sort exemplars.

Analysis

Before modelling, we cleaned the data for attention check trials and unbalanced items (1,619 trials in total; see Online Supplements: osf.io/e2rdj), leaving 63,810 trials for analysis.

We modelled accept/reject-responses (coded accept = 1, reject = 0) as a probit-linked Bernoulli-distribution using Bürkner's (2017) brms-package in R. The main fixed effect was whether the item truly belonged in the fraction, the 'signal' (coded signal = 0.5, noise = -0.5). This design allowed the model's estimates to be interpreted in terms of signal detection theory parameters of an equal variance signal detection theoretic model (for the mathematics, see DeCarlo, 1998; for practical implementation, see Vuorre, 2017). With this parameterisation, the model's intercept equals the negative signal detection theoretic criterion (-*c*), meaning that an increase of the intercept equals a more liberal criterion, while the coefficient for belonging equals the sensitivity (*d'*) as it indicates the estimated difference between accepting true positives (signal) and true negatives (noise).

As fixed effects, we also included the combined fraction:wave factor and its interaction with signal. We added fixed effects for condition and its interaction with the signal for trials of 'residual waste' from wave two only. We added random intercepts for items, participants, and observed participant-fraction:wave-combinations all with random slopes for signal. All factors (item, participant, and fraction:wave) were sum-zero contrast coded.

For the collapsed fractions in wave two ('plastics / food cartons' and 'residual waste'), we allowed participants responses to signal trials to vary by the individual fractions ('plastics' and 'beverage & food cartons' and 'food waste', 'paper', and 'residual waste', respectively)[1]. For 'residual waste' trials of wave two, acceptance was further allowed to vary by condition. Lastly, we added a random slope for these subcategory acceptance variations covarying by participant.

As priors, we used standard normal distributions for intercepts and condition-coefficients, standard half-normal distributions for group-level standard deviations, and a Cholesky prior of 5 for correlation matrices.

For model estimates, we report posterior median values (*b*), posterior standard errors (*SE*) with equal-tailed 95% credible intervals (*CrI*), and effects on the natural scale in percentage points (p.p.). Note that in signal trials the item should be accepted, thus acceptance reflect accuracy. In noise trials the item should be rejected, thus an increase in acceptance equals a decrease in accuracy. Furthermore, although the model estimates the negative criterion (-*c*) as its intercept on the probit scale, we report *c*, such that higher values correspond to a more conservative decision criterion.

When comparing 'plastics / food cartons' between waves one and two, we use the mean of estimates for 'beverage & food cartons' and 'plastics' for wave one.

While data for 'cardboard' and 'hazardous waste' collected in wave one was part of the modelling, we do not report on these, as they have no relevance for our conclusions (apart from strengthening the model's estimation abilities).

Results

Unchanged Fractions

Figure 3, Panel A, shows the model estimates for fractions that were sorted separately in both waves one and two ('glass' and 'metal'). Including these allowed us to estimate the effect of collapsing fraction on the sorting accuracy, sensitivity, and criterion of seemingly unaffected fractions. Acceptance estimates for neither signal nor noise trials differed reliably between waves for 'glass', nor did the estimates of sensitivity and criterion. For 'metal', signal trials were slightly more likely to be accepted (b_{diff} = 0.27, *SE* = 0.1, *95 CrI* = [0.08, 0.47]), driven by a small but reliable increase in sensitivity (b_{diff} = 0.32, *SE* = 0.13, *95 CrI* = [0.06, 0.58]) and an unreliable move towards a more liberal criterion (b_{diff} = -0.11, *SE* = 0.07, *95 CrI* = [-0.25, 0.03]). This equals a ~5 p.p. (*SE* = 0.02, *95 CrI* = [0.01, 0.1]) increase in accuracy in signal trials.

In sum, collapsing fractions did not induce a global shift in accuracy, criterion or sensitivity for unaffected fractions.

Explicitly Combined Fractions

We next compared the modelled acceptance rates for 'plastics' and 'beverage & food cartons' in wave one with 'plastics / food cartons' in wave two (see Panel B in Figure 3). This allowed us to estimate the difference in performance between collapsed and non-collapsed fractions when the collapsing was explicit. We observed an acceptance rate of signal trials for the combined fraction in wave two not differing reliably (b_{diff} = 0.1, *SE* = 0.13, *95 CrI* = [-0.14, 0.36]) from the average of the individual fractions in wave one (marked as shaded distribution in Figure 3, Panel B). However, for noise trials, the model estimated a reliable increase in acceptance (b_{diff} = 0.38, *SE* = 0.07, *95 CrI* = [0.24, 0.51]) corresponding to a ~7 p.p. decrease (*SE* = 0.01, *95 CrI* = [0.04, 0.1]) in accuracy for noise trials. This change was driven by a reliably more liberal criterion (b_{diff} = -0.24, *SE* = 0.07, *95 CrI* = [-0.38, -0.1]) combined with an unreliable decrease in sensitivity (b_{diff} = -0.27, *SE* = 0.14, *95 CrI* = [-0.55, 0.01]).

Although the estimates of the individual fractions in wave two must be interpreted with caution, the model estimates further suggested that the explicitly combined fraction increased acceptance for 'plastics' in wave two compared to wave one (b_{diff} = 0.25, *SE* = 0.09, *95 CrI* = [0.09, 0.43]) equalling a ~4 p.p. increase (*SE* = 0.01, *95 CrI* = [0.01, 0.06]) in accuracy for signal trials, reflecting a more liberal criterion (b_{diff} = -0.28, *SE* = 0.06, *95 CrI* = [-0.39, -0.17]) and unchanged sensitivity (b_{diff} = -0.04, *SE* = 0.1, *95 CrI* = [-0.25,

[1] These variances are only allowed for signal trials as trials that are noise in relation to either subcategory are noise trials to all. This challenges the direct mapping between our model's estimates and signal detection theoretic terms slightly, and interpretations of the criterion and sensitivity estimates should be interpreted cautiously. However, with the variance sum-zero contrast coded, the subcategories vary around a shared mean, allowing, at least, for correct interpretations of criterion and sensitivity for the combined fraction.

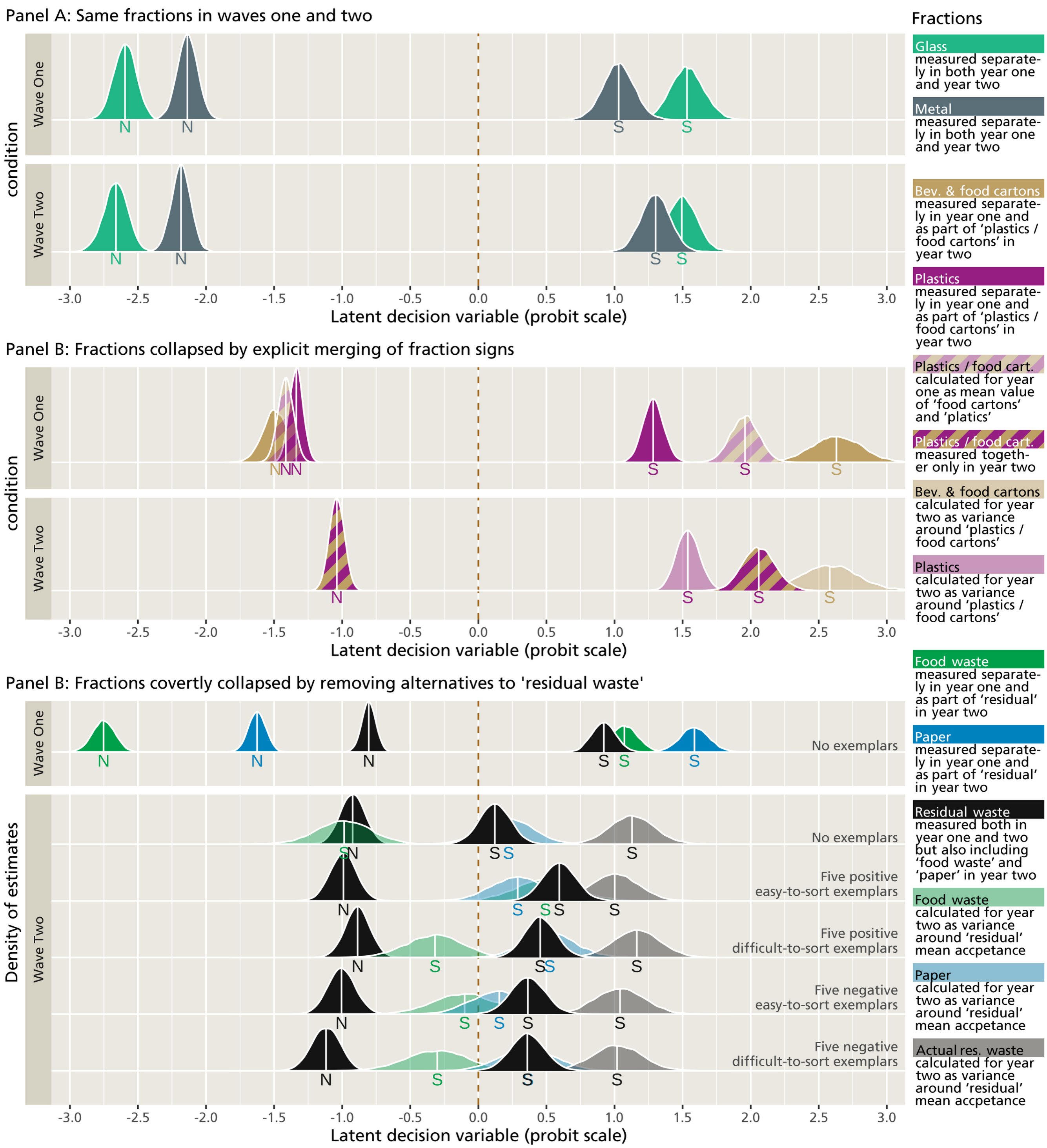

Figure 3 Distributions of model estimates of latent decision variable on probit scale for signal trials (marked 'S') and noise trials (marked 'N'). Distances between distribution medians (vertical white lines) equal the signal detection theoretic sensitivity parameter, *d'*. Halfway between them: the negative criterion, -*c*. Panel A: 'Glass' and 'metal' sorted separately in both waves one and two. Panel B: 'Beverage & food cartons' and 'plastics' sorted separately in wave one and explicitly combined into 'plastics / food cartons' in wave two. Shaded combined distribution in wave one is calculated as mean between the measured distributions. Shaded distribution in wave two shows fraction-specific acceptance-variance in signal trials around mean (striped, opaque distribution). Panel C: 'Food waste', 'paper', and 'residual waste' sorted separately in wave one and all in 'residual waste' in wave two. Shaded distribution in wave two shows fraction-specific acceptance-variance in signal trials around mean (opaque black distribution). Wave two distributions are further divided by condition.

0.16]). For 'food cartons' the model estimated no difference in acceptance rates for signal trials, but the increased acceptance rate for the combined fraction in wave two lead to a lower sensitivity estimate (b_{diff} = -0.5, *SE* = 0.24, *95 CrI* = [-0.98, -0.02]) with a slightly, but unreliably, more liberal criterion (b_{diff} = -0.2, *SE* = 0.13, *95 CrI* = [-0.45, 0.04]).

In sum, explicitly combining fractions led to a slight decrease in accuracy noise trials leading to more false positives, fuelled by a more liberal criterion.

Covertly Combined Fractions

We next compared the modelled acceptance rates for the fractions in wave two ('food waste', 'paper', and 'residual waste') that covertly combined by virtue of removing the former two of the alternatives and expecting participants to sort them into 'residual waste'.

Comparing the modelled acceptance rate for 'residual waste' signal trials between waves one and two, we saw a serious decline in acceptance (b_{diff} = -0.8, *SE* = 0.13, *95 CrI* = [-1.06, -0.55]) equalling a ~27 p.p. decrease (*SE* = 0.05, *95 CrI* = [-0.37, -0.18]) in accuracy for 'residual waste' signal trials while the noise trials showed no reliable change (b_{diff} = -0.12, *SE* = 0.1, *95 CrI* = [-0.31, 0.07]). The decline for signal trials appeared to be mostly driven by even larger declines for 'food waste' (b_{diff} = -2.06, *SE* = 0.2, *95 CrI* = [-2.45, -1.68]) and 'paper' (b_{diff} = -1.37, *SE* = 0.19, *95 CrI* = [-1.74, -1]), equalling reductions of accuracy of ~69 p.p. (*SE* = 0.05, *95 CrI* = [-0.78, -0.59]) and ~35 p.p. (*SE* = 0.07, *95 CrI* = [-0.5, -0.23]), respectively. By comparison, the acceptance rate for items that would also belong in 'residual waste' under full fraction granularity increased slightly, albeit not reliably (b_{diff} = 0.21, *SE* = 0.16, *95 CrI* = [-0.11, 0.53]).

In signal detection theoretic terms, this was driven by a decrease in sensitivity (b_{diff} = -0.68, *SE* = 0.15, *95 CrI* = [-0.99, -0.38]) and a move towards a more conservative criterion (b_{diff} = 0.46, *SE* = 0.09, *95 CrI* = [0.29, 0.63]). While again interpreting the signal detection theoretic parameter estimates for the subcategories with caution, for 'food waste', sensitivity plummeted (b_{diff} = -3.89, *SE* = 0.22, *95 CrI* = [-4.33, -3.45]) while the criterion did not reliably change (b_{diff} = 0.11, *SE* = 0.12, *95 CrI* = [-0.11, 0.34]). Indeed, as can be seen by the shaded distribution of signal 'food waste' acceptance aligning with the distribution for noise 'residual waste' trials in Panel C in Figure 3, sensitivity for 'food waste' in the covertly combined 'residual waste' fraction is just around zero (b_{diff} = -0.06, *SE* = 0.22, *95 CrI* = [-0.49, 0.36]) in wave two. For 'paper', sensitivity likewise decreased (b_{diff} = -2.07, *SE* = 0.2, *95 CrI* = [-2.47, -1.66]), while the criterion also became more conservative (b_{diff} = 0.33, *SE* = 0.11, *95 CrI* = [0.11, 0.55]). The collapsing of other fractions did not appear to affect items normally belonging in 'residual waste' with no reliable changes in neither sensitivity (b_{diff} = 0.33, *SE* = 0.18, *95 CrI* = [-0.03, 0.68]) nor criterion (b_{diff} = -0.05, *SE* = 0.1, *95 CrI* = [-0.24, 0.15]).

In sum, covertly combining fractions by removing some alternatives and expecting people to re-assign these to the "next logical alternative" severely harmed sorting accuracy for the removed fractions while not reliably affecting the fraction they were absorbed into. This effect was driven by a plummeting sensitivity and potentially more conservative criterion for the removed fractions.

Effect of Exemplars

In wave 2 we additionally manipulated how many exemplars and which were shown on the 'residual waste' sign. We found that positive exemplars reliably increased the acceptance rate for 'residual waste' signal trials with both easy-to-sort (b_{diff} = 0.47, *SE* = 0.15, *95 CrI* = [0.18, 0.78]) and difficult-to-sort (b_{diff} = 0.33, *SE* = 0.15, *95 CrI* = [0.05, 0.62]) positive exemplars and showed positive, albeit unreliable, tendencies for easy-to-sort (b_{diff} = 0.25, *SE* = 0.16, *95 CrI* = [-0.07, 0.55]) and difficult-to-sort (b_{diff} = 0.24, *SE* = 0.16, *95 CrI* = [-0.07, 0.55]) negative exemplars compared to the no-exemplar condition. However, the main driver appeared to be the items normally belonging to 'food waste' and 'paper', as trials with items normally belonging to 'residual waste' showed no reliable change in acceptance rate under any of the conditions compared to the no-exemplar condition.

Indeed, in signal trials with items normally belonging to 'food waste' the model estimated reliable positive effects of both positive easy-to-sort (b_{diff} = 1.48, *SE* = 0.28, *95 CrI* = [0.94, 2.04]), positive difficult-to-sort (b_{diff} = 0.66, *SE* = 0.27, *95 CrI* = [0.13, 1.19]), negative easy-to-sort (b_{diff} = 0.89, *SE* = 0.28, *95 CrI* = [0.32, 1.44]), and negative difficult-to-sort (b_{diff} = 0.68, *SE* = 0.28, *95 CrI* = [0.13, 1.24]) exemplars, compared to the no-exemplar condition. The effect of positive easy-to-sort exemplars further reliably increased acceptance rate in signal trials compared to both positive (b_{diff} = 0.81, *SE* = 0.29, *95 CrI* = [0.26, 1.39]) and negative difficult-to-sort (b_{diff} = 0.8, *SE* = 0.3, *95 CrI* = [0.2, 1.41]) exemplars and, although not reliably, compared with negative easy-to-sort exemplars (b_{diff} = 0.59, *SE* = 0.3, *95 CrI* = [0, 1.2]).

For signal trials with items normally belonging to 'paper', the model estimated no reliable difference between any of the exemplar conditions compared to the no-exemplar condition. Neither did the model estimate any reliable effect on acceptance for noise trials between any of the exemplar conditions and the no-exemplars condition.

For the combined 'residual waste' fraction of wave two, sensitivity was reliably increased by both positive easy-to-sort exemplars (b_{diff} = 0.54, *SE* = 0.18, *95 CrI* = [0.19, 0.91]) and negative difficult-to-sort exemplars (b_{diff} = 0.43, *SE* = 0.19, *95 CrI* = [0.07, 0.81]), with positive easy-to-sort exemplars almost reliably causing a more liberal criterion (b_{diff} = -0.2, *SE* = 0.1, *95 CrI* = [-0.41, 0]) and negative difficult-to-sort exemplars not producing a meaningful change in criterion (b_{diff} = -0.02, *SE* = 0.11, *95 CrI* = [-0.23, 0.19]). Neither positive difficult-to-sort nor negative easy-to-sort exemplars affected sensitivity or criterion reliably. Again, these effects appear to be mostly driven by the effect on 'food' items.

In sum, adding exemplars increased the acceptance of some of the covertly collapsed fractions into the absorbing fraction, with positive easy-to-sort exemplars driving the biggest increase.

Discussion

In this paper, based on responses from 669 participants collected at a major Danish music festival, we compared how explicitly and covertly collapsing waste fractions affected participants' sorting accuracy, sensitivity, and criterion.

For 'beverage & food cartons' and 'plastics' explicitly merged into 'plastics / food cartons' in wave two, we saw a similar acceptance of items in signal trials as the mean acceptance of the individual fractions in wave one. For signal trials in wave two, we further saw a slight increase of acceptance of items that would normally belong in 'plastics'. For noise trials, we saw an increase in acceptance when comparing the explicitly merged fraction in wave two with the mean of the individual fractions in wave one. In signal detection theoretic terms, this reflects a more liberal criterion for 'plastics' with no change in sensitivity and a lower sensitivity estimate (that is: worse detection) for 'food cartons' without a change in criterion. Despite the slightly higher false positive rate, explicitly collapsing fractions appear to communicate well the individual fractions.

One potential explanation of why explicitly combining waste fractions lead to too liberal inclusion criteria might be found by regarding the waste sorting task in terms of the separation of a conceptual waste space (for an introduction to conceptual spaces, see e.g., Gärdenfors, 1990, 2004). If you accept that areas in the conceptual space governing a single category could have graded borders, the combination of two neighbouring categories might increase the border space now accepted. For example, imagine two overlapping circles indicating membership with an opaque colour, with transparency increasing away from their centres, and how, where they overlap, the opaqueness combines, increasing the overall opaque area. Whether the explicit combination of two fractions by merging their signs do indeed require participants to form a mixed understanding of the fraction or whether it just directs two individual fractions into one bin, is a question for future work.

When covertly combining fractions (as for 'food waste' and 'paper' into 'residual waste' in wave two), we, overall, saw a decline in acceptance in 'residual waste' signal trials driven by a big decline in acceptance of 'food waste' and 'paper' items in wave two compared to when the fractions were sorted individually in wave one. In fact, the acceptance rate of 'food waste' items in signal trials was as low as in noise trials for 'residual waste' overall. Together, this contrasts sharply with explicit collapsing and indicates that removing alternatives does not by itself induce a broadened category representation.

When participants were provided exemplars, this pattern changed. The acceptance of 'food waste' into 'residual waste' was reliably increased by all exemplar conditions, while items normally belonging to 'paper' were unaffected. Interestingly, even negative exemplars, i.e. showing items not belonging to the fraction with a red barred circle on top (see Figure 2, right), helped participants to accept 'food waste' into 'residual waste' in wave two. This suggests that part of the effect of exemplars might be a "reminder" to participants that 'residual waste' was atypical to what they were accustomed to. For 'food waste', positive easy-to-sort exemplars even increased acceptance more than the other exemplar conditions, suggesting that fraction-typical exemplars might best convey an understanding of the now combined fraction.

That items normally belonging in 'food waste' appears to be more affected by being collapsed into 'residual waste' than items normally belonging to 'paper' and similarly are affected more by the exemplars could be explained by the conceptual overlap of the three fractions. To investigate that, we multidimensionally scaled items' likelihood of ending up in any of the nine fractions in wave one into a two-dimensional space (for such a plot, see Online Supplements: osf.io/e2rdj). In this space, the clouds of items correctly belonging in 'residual waste' and 'paper' show a bigger overlap than any of them do with the cloud of items belonging in 'food waste'. This could suggest that participants' mental representation of 'paper waste' is closer to 'residual waste' than any of them are to 'food waste'. Whether this relationship between overlapping mental representations, sorting accuracy, and exemplar effects generalises to other fractions is a potential avenue for further research.

Interestingly, we did not observe any effect of any of the exemplar conditions on the acceptance of items normally belonging in 'residual waste'. This is surprising given existing studies showing exemplars supporting sorting accuracy in natural settings (see e.g., Austin et al., 1993; Cakanlar et al., 2025; Miller et al., 2016). It is especially surprising as our own ongoing studies suggest that the effect of exemplars is attenuated for 'residual waste'

Overall, explicitly collapsing waste fractions by including their names, icons, and colours on a single sign appears better at promoting an understanding of the now combined fraction than covertly removing alternatives and expecting participants to automatically re-assign items normally belonging to the removed fractions. Even when given exemplars. Whether this is due to abstracted icons in themselves better supporting combined conceptual understandings, due to overlearning these specific icons found in everyday life, or due to the inclusion of both fraction names on the sign, remains to be determined. Lastly, given that 'metal' and 'glass' (which remained unchanged between waves one and two) showed no negative effect when other fractions were collapsed suggests that collapsing fractions likely has no negative spillover effects. Thus, for now, the practical recommendation following this study is: when simplifying an otherwise established waste sorting system for a given context, be explicit about which fractions now belong together, thereby trading more liberal decision criteria for preserved discriminability.

Acknowledgements

This work was supported by the Margit Stiernswärd Fund for Environmental Management Research (Ref. nr: 2024/1987 and 2025/2580), the Kamprad Family Foundation (Ref. nr: 20230166), and the Circular Laboratory – a partnership of Roskilde Festival and the Tuborg Foundation. Generative AI has assisted us in optimising the analysis' model structure.

References

Abdel Sater, R., Mus, M., Wyart, V., & Chevallier, C. (2023). A zero-cost attention-based approach to promote cleaner streets: A Signal Detection Theory approach in Parisian streets. *PLOS ONE*, *18*(4), e0284272. https://doi.org/10.1371/journal.pone.0284272

Ashby, F. G., Boynton, G., & Lee, W. W. (1994). Categorization response time with multidimensional stimuli. *Perception & Psychophysics*, *55*(1), 11–27. https://doi.org/10.3758/BF03206876

Austin, J., Hatfield, D. B., Grindle, A. C., & Bailey, J. S. (1993). Increasing Recycling in Office Environments: The Effects of Specific, Informative Cues. *Journal of Applied Behavior Analysis*, *26*(2), 247–253. https://doi.org/10.1901/jaba.1993.26-247

Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Butorova, A. S., Sergeev, A. P., & Tarasov, D. A. (2022). Psychophysical Model of the Signal Detection Theory in the Categorization Task of Textual and Pictographic Visual Stimuli. *2022 Ural-Siberian Conference on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT)*, 228–231. https://doi.org/10.1109/USBEREIT56278.2022.9923330

Cakanlar, A., Hunter, M., & Nenkov, G. Y. (2025). Recycle right: How to decrease recycling contamination with informational point-of-disposal signage. *Journal of the Academy of Marketing Science*, *53*(3), 781–803. https://doi.org/10.1007/s11747-024-01058-1

DeCarlo, L. T. (1998). Signal detection theory and generalized linear models. *Psychological Methods*, *3*(2), 186–205. https://doi.org/10.1037/1082-989X.3.2.186

Erev, I. (1998). Signal detection by human observers: A cutoff reinforcement learning model of categorization decisions under uncertainty. *Psychological Review*, *105*(2), 280–298. https://doi.org/10.1037/0033-295X.105.2.280

EUpicto. (2020). *EUpicto*. https://www.eupicto.com

Gärdenfors, P. (1990). Induction, Conceptual Spaces and AI. *Philosophy of Science*, *57*(1), 78–95. https://www.jstor.org/stable/187622

Gärdenfors, P. (2004). Conceptual Spaces as a Framework for Knowledge Representation. *Mind and Matter*, *2*(2), 9–27.

Grill-Spector, K., & Kanwisher, N. (2005). Visual Recognition: as soon as you know it is there, you know what it is. *Psychological Science*, *16*(2), 152–160. https://doi.org/10.1111/j.0956-7976.2005.00796.x

Hu, M., & Nosofsky, R. M. (2021). Exemplar-Model Account of Categorization and Recognition When Training Instances Never Repeat. *Journal of Experimental Psychology: Learning Memory and Cognition*, *48*(12), 1947–1969. https://doi.org/10.1037/xlm0001008

Jang, Y., Wixted, J. T., & Huber, D. E. (2009). Testing signal-detection models of yes/no and two-alternative forced-choice recognition memory. *Journal of Experimental Psychology: General*, *138*(2), 291–306. https://doi.org/10.1037/a0015525

Krishnan, K., Singh, S., Trushna, T., Kalyanasundaram, M., Sabde, Y., Atkins, S., Sahoo, K. C., Lundborg, C. S., Rousta, K., & Diwan, V. (2026). Standardizing waste classification in low and middle income countries: addressing terminology gaps to improve urban segregation behavior. *International Journal of Environmental Science and Technology*, *23*(3), 185. https://doi.org/10.1007/s13762-025-06994-1

Macmillan, N. A. (2001). Signal Detection Theory. In *International Encyclopedia of the Social & Behavioral Sciences* (Vol. 28, pp. 14075–14078). Elsevier. https://doi.org/10.1016/B0-08-043076-7/00677-X

Miller, N. D., Meindl, J. N., & Caradine, M. (2016). The Effects of Bin Proximity and Visual Prompts on Recycling in a University Building. *Behavior and Social Issues*, *25*(1), 4–10. https://doi.org/10.5210/bsi.v25i0.6141

Rosenthal, S., & Linder, N. (2021). Effects of bin proximity and informational prompts on recycling and contamination. *Resources, Conservation and Recycling*, *168*, 105430. https://doi.org/10.1016/j.resconrec.2021.105430

Rouder, J. N., & Ratcliff, R. (2006). Comparing Exemplar- and Rule-Based Theories of Categorization. *Current Directions in Psychological Science*, *15*(1), 9–13. https://doi.org/10.1111/j.0963-7214.2006.00397.x

Sharan, L., Rosenholtz, R., & Adelson, E. H. (2014). Accuracy and speed of material categorization in real-world images. *Journal of Vision*, *14*(9), 12–12. https://doi.org/10.1167/14.9.12

Tanner, W. P., & Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological Review*, *61*(6), 401–409. https://doi.org/10.1037/h0058700

Vuorre, M. (2017). *Estimating Signal Detection Models with regression using the brms R package*. https://vuorre.com/posts/sdttregression/.

Wrisberg, A. P., Wallin, A., & Pärnamets, P. (2026). *LEWIS: Lund Everyday Waste Image Set of Naturally Sourced, Normalised, and Evaluated Waste Images*. PsyArXiv. https://doi.org/10.31234/osf.io/b48gv_v1

Wu, D. W. L., Lenkic, P. J., DiGiacomo, A., Cech, P., Zhao, J., & Kingstone, A. (2018). How does the design of waste disposal signage influence waste disposal behavior? *Journal of Environmental Psychology*, *58*, 77–85. https://doi.org/10.1016/j.jenvp.2018.07.009