A Materials

The mappings from abstract frames to their corresponding instantiations for both our replication of White et al. 2018 (Table 1) and the MegaAcceptability dataset (Table 2) can be found below.

Abstract Frame	Instantiated Frame
NP ____ed	Someone ____ed.
NP ____ed NP	Someone ____ed something.
NP ____ed NP NP	Someone ____ed someone something.
NP ____ed NP S	Someone ____ed someone something happened.
NP ____ed NP S[−tense]	Someone ____ed someone something happen.
NP ____ed NP VP	Someone ____ed someone do something.
NP ____ed NP about NP	Someone ____ed someone about something.
NP ____ed NP that S	Someone ____ed someone that something happened.
NP ____ed NP that S[−tense]	Someone ____ed someone that something happen.
NP ____ed NP to VP	Someone ____ed someone to do something.
NP ____ed S	Someone ____ed something happened.
NP ____ed VPing	Someone ____ed doing something.
NP ____ed WH S	Someone ____ed why something happened.
NP ____ed WH to VP	Someone ____ed why to do something.
NP ____ed about NP	Someone ____ed about something.
NP ____ed for NP to VP	Someone ____ed for someone to do something.
NP ____ed if S	Someone ____ed if something happened.
NP ____ed if S[−tense]	Someone ____ed if something happen.
NP ____ed it that S	Someone ____ed it that something happened.
NP ____ed it that S[−tense]	Someone ____ed it that something happen.
NP ____ed so	Someone ____ed so.
NP ____ed that S	Someone ____ed that something happened.
NP ____ed that S[−tense]	Someone ____ed that something happen.
NP ____ed there to VP	Someone ____ed there to be a particular thing in a particular place.
NP ____ed to	Someone ____ed to.
NP ____ed to NP that S	Someone ____ed to someone that something happened.
NP ____ed to NP that S[−tense]	Someone ____ed to someone that something happen.
NP ____ed to VP	Someone ____ed to do something.
NP was ____ed that S	Someone was ____ed that something happened.
NP was ____ed that S[−tense]	Someone was ____ed that something happen.
NP was ____ed to VP	Someone was ____ed to do something.
S, I ____	Something happened, I ____.
S, NP ____ed	Something happened, someone ____ed.
It ____ed NP WH S	It ____ed someone why something happened.
It ____ed NP WH to VP	It ____ed someone why to do something.
It ____ed NP that S	It ____ed someone that something happened.
It ____ed NP that S[−tense]	It ____ed someone that something happen.
It ____ed NP to VP	It ____ed someone to do something.

Table 1: Abstract frames and corresponding instantiated frames used in our replication of White et al. 2018 (Section 3).

Abstract Frame	Instantiated Frame
NP ____ed	Someone ____ed.
NP ____ed NP	Someone ____ed something.
NP ____ed NP VP	Someone ____ed someone do something.
NP ____ed NP VPing	Someone ____ed someone doing something.
NP ____ed NP that S	Someone ____ed someone that something happened.
NP ____ed NP that S[+future]	Someone ____ed someone that something would happen.
NP ____ed NP that S[−tense]	Someone ____ed someone that something happen.
NP ____ed NP to NP	Someone ____ed something to someone.
NP ____ed NP to VP[+eventive]	Someone ____ed someone to do something.
NP ____ed NP to VP[−eventive]	Someone ____ed someone to have something.
NP ____ed NP whether S	Someone ____ed someone whether something happened.
NP ____ed NP whether S[+future]	Someone ____ed someone whether something would happen.
NP ____ed NP whichNP S	Someone ____ed someone which thing happened.
NP ____ed S	Someone ____ed something happened.
NP ____ed VPing	Someone ____ed doing something.
NP ____ed about NP	Someone ____ed about something.
NP ____ed about whether S	Someone ____ed about whether something happened.
NP ____ed for NP to VP	Someone ____ed for something to happen.
NP ____ed so	Someone ____ed so.
NP ____ed that S	Someone ____ed that something happened.
NP ____ed that S[+future]	Someone ____ed that something would happen.
NP ____ed that S[−tense]	Someone ____ed that something happen.
NP ____ed to NP that S	Someone ____ed to someone that something happened.
NP ____ed to NP that S[+future]	Someone ____ed to someone that something would happen.
NP ____ed to NP that S[−tense]	Someone ____ed to someone that something happen.
NP ____ed to NP whether S	Someone ____ed to someone whether something happened.
NP ____ed to NP whether S[+future]	Someone ____ed to someone whether something would happen.
NP ____ed to VP[+eventive]	Someone ____ed to do something.
NP ____ed to VP[−eventive]	Someone ____ed to have something.
NP ____ed whether S	Someone ____ed whether something happened.
NP ____ed whether S[+future]	Someone ____ed whether something would happen.
NP ____ed whether to VP	Someone ____ed whether to do something.
NP ____ed whichNP S	Someone ____ed which thing happened.
NP ____ed whichNP to VP	Someone ____ed which thing to do.
NP was ____ed	Someone was ____ed.
NP was ____ed S	Someone was ____ed something happened.
NP was ____ed about NP	Someone was ____ed about something.
NP was ____ed about whether S	Someone was ____ed about whether something happened.
NP was ____ed so	Someone was ____ed so.
NP was ____ed that S	Someone was ____ed that something happened.
NP was ____ed that S[+future]	Someone was ____ed that something would happen.
NP was ____ed that S[−tense]	Someone was ____ed that something happen.
NP was ____ed to VP[+eventive]	Someone was ____ed to do something.
NP was ____ed to VP[−eventive]	Someone was ____ed to have something.
NP was ____ed whether S	Someone was ____ed whether something happened.
NP was ____ed whether S[+future]	Someone was ____ed whether something would happen.
NP was ____ed whether to VP	Someone was ____ed whether to do something.
NP was ____ed whichNP S	Someone was ____ed which thing happened.
NP was ____ed whichNP to VP	Someone was ____ed which thing to do.
S, I ____	Something happened, I ____.

Table 2: Abstract frames and corresponding instantiated frames in the MegaAcceptability dataset (Section 4; see also White & Rawlins 2016).

All of the verbs used in the MegaAcceptability dataset can be found below.

a abhor, absolve, accept, acclaim, accredit, acknowledge, add, address, admire, admit, admonish, adore, advertise, advise, advocate, affect, affirm, afford, affront, aggravate, aggrieve, agitate, agonize, agree, aim, alarm, alert, allege, allow, alter, amaze, amuse, analyze, anger, anguish, annotate, announce, annoy, answer, anticipate, apologize, appall, appeal, appear, appease, applaud, apply, appoint, appraise, appreciate, approach, approve, argue, arouse, arrange, articulate, ascertain, ask, assert, assess, assign, assume, assure, astonish, astound, attempt, attest, audit, authorize, awe

b babble, back, badger, baffle, bandy, banter, bargain, bark, be, beam, bear, befuddle, beg, begin, believe, belittle, bellow, beseech, bet, bewilder, bicker, bitch, blame, blare, blast, bleat, bless, blog, bluff, bluster, boast, boggle, bore, bother, brag, brainstorm, bribe, brief, broadcast, brood, bug, bullshit, bully, bury, buy

c cackle, cajole, calculate, calibrate, call, calm, care, carp, catch, categorize, cause, caution, cease, celebrate, censor, censure, certify, challenge, change, chant, characterize, charge, charm, chasten, chastise, chat, chatter, check, cheer, cherish, chide, chime, chirp, choose, chronicle, chuckle, circulate, claim, clarify, classify, clear, cloud, coach, coax, coerce, come, come around, come out, comfort, command, commence, commend, comment, commission, communicate, compel, compete, complain, compliment, comprehend, compromise, compute, conceal, concede, conceive, concern, conclude, concur, condemn, condone, confess, confide, configure, confirm, confound, confuse, congratulate, conjecture, connect, consent, consider, console, conspire, constrain, consult, contact, contemplate, contend, content, contest, continue, contract, contribute, contrive, control, convey, convince, correct, corroborate, cough, counsel, counter, cover, crack, crave, credential, cringe, criticize, croak, croon, crow, crush, cry, curse

d dare, daunt, daydream, daze, debate, deceive, decide, declare, decline, decree, decry, deduce, deem, defend, define, deject, delete, deliberate, delight, delineate, delude, demand, demean, demonstrate, demoralize, demystify, denounce, deny, depict, deplore, depress, deride, derive, describe, design, designate, desire, despair, despise, detail, detect, determine, detest, devastate, devise, diagnose, dictate, dig, direct, disagree, disallow, disappoint, disapprove, disbelieve, discern, discipline, disclose, disconcert, discourage, discover, discriminate, discuss, disgrace, disgruntle, disgust, dishearten, disillusion, dislike, dismay, dismiss, disparage, dispatch, dispel, dispirit, display, displease, disprefer, disprove, dispute, disquiet, disregard, dissatisfy, dissent, distract, distress, distrust, disturb, dither, divulge, document, doubt, draw, drawl, dread, dream, drone, dub, dupe

e educate, elaborate, elate, elect, electrify, elucidate, email, embarrass, embellish, embitter, embolden, emphasize, employ, enchant, encourage, end, endorse, endure, energize, enforce, engage, enjoy, enlighten, enlist, enrage, ensure, enthrall, enthuse, entice, entreat, envision, envy, establish, estimate, evaluate, evidence, examine, exasperate, excite, exclaim, excuse, exhibit, exhilarate, expect, experience, explain, exploit, explore, expose, expound, express, extrapolate

f fabricate, face, fake, fancy, fantasize, fascinate, fax, faze, fear, feel, feign, fess up, feud, fight, figure, figure out, file, find, find out, finish, flatter, flaunt, flip out, floor, fluster, flutter, fool, forbid, force, forecast, foresee, foretell, forget, forgive, forgo, formulate, frame, freak out, fret, frighten, frown, frustrate, fuel, fume, function, fuss

g gab, gall, galvanize, gamble, gasp, gather, gauge, generalize, get, giggle, gladden, glare, glean, glimpse, gloat, glorify, go, gossip, grant, grasp, gratify, grieve, grill, grimace, grin, gripe, groan, grouse, growl, grumble, grunt, guarantee, guess, guide, gurgle, gush

h haggle, hallucinate, handle, hanker, happen, harass, hasten, hate, hear, hearten, hedge, hesitate, highlight, hinder, hint, hire, hold, holler, hoot, hope, horrify, hound, howl, humble, humiliate, hunger, hurt, hush up, hustle

i identify, ignore, illuminate, illustrate, imagine, imitate, impede, impel, implore, imply, impress, incense, incite, include, indicate, indict, induce, infer, influence, inform, infuriate, initiate, inquire, inscribe, insert, insinuate, insist, inspect, inspire, instigate, instruct, insult, insure, intend, intercept, interest, interject, interpret, interrogate, interview, intimate, intimidate, intrigue, investigate, invigorate, invite, irk, irritate, isolate

j jabber, jade, jar, jeer, jest, joke, judge, jump, justify

k keep, kid, know

l label, lament, laud, laugh, lead, leak, learn, lecture, legislate, license, lie, like, lisp, listen, loathe, lobby, log, long, look, love, lust

m madden, mail, maintain, make, make out, malign, mandate, manipulate, manufacture, mark, marvel, mean, measure, meditate, meet, memorize, mention, miff, mind, minimize, misinform, misjudge, mislead, miss, mistrust, moan, mock, monitor, mope, mortify, motivate, mourn, move, mumble, murmur, muse, mutter, mystify

n name, narrate, nauseate, need, negotiate, nonplus, note, notice, notify

o object, obligate, oblige, obscure, observe, obsess, offend, offer, okay, omit, operate, oppose, ordain, order, outline, outrage, overestimate, overhear, overlook, overwhelm

p pain, panic, pant, pardon, pause, perceive, permit, perplex, persuade, perturb, pester, petition, petrify, phone, pick, picket, picture, piece together, pine, pinpoint, pity, placate, plan, plead, please, plot, point out, ponder, pontificate, portend, portray, posit, post, pout, praise, pray, preach, predict, prefer, prejudge, prepare, present, press, pressure, presume, presuppose, pretend, print, probe, proclaim, procrastinate, prohibit, promise, prompt, prophesy, propose, protest, prove, provoke, publicize, publish, punt, pursue, puzzle

q qualify, quarrel, query, question, quibble, quip, quiz, quote

r radio, raise, rankle, rant, rap, rationalize, rave, read, reaffirm, realize, reason, reason out, reassert, reassess, reassure, rebuke, recall, recap, reckon, recognize, recollect, recommend, reconsider, reconstruct, record, recount, recruit, rediscover, reevaluate, reexamine, regard, register, regret, regulate, reiterate, reject, relate, relax, relay, relearn, relieve, relish, remain, remark, remember, remind, reminisce, renegotiate, repeat, repent, reply, report, represent, repress, reprimand, reproach, request, require, research, resent, resolve, respect, respond, restate, result, resume, retort, retract, reveal, review, revolt, ridicule, rile, ring, rouse, rue, rule, ruminate, rush

s sadden, sanction, satisfy, say, scare, schedule, scheme, scoff, scold, scorn, scowl, scramble, scrawl, scream, screech, scribble, scrutinize, see, seek, seem, select, send, sense, serve, set, set about, set out, settle, shame, shape, share, shatter, shock, shoot, shout, show, showcase, shriek, shut up, sicken, sigh, sign, sign on, sign up, signal, signify, simulate, sing, sketch, skirmish, slander, smell, smile, smirk, snap, sneer, snicker, snitch, snivel, snort, snub, sob, sober, soothe, sorrow, speak, specify, speculate, spellbind, splutter, spook, spot, spout, spread, spur, sputter, squabble, squawk, squeal, stagger, stammer, stand, start, start off, startle, state, steer, stereotype, stew, stifle, stimulate, stipulate, stop, store, strain, stress, struggle, strut, study, stump, stun, stupefy, stutter, subdue, submit, suffer, suggest, sulk, summarize, summon, suppose, surmise, surprise, survey, suspect, swear, sweat, swoon

t tackle, take, talk, tantalize, tap, tape, taste, taunt, teach, tease, televise, tell, tempt, terrify, terrorize, test, testify, thank, theorize, think, thirst, threaten, thrill, tickle, torment, torture, tout, track, train, transmit, traumatize, trick, trigger, trouble, trust, try, turn out, tutor, tweet, type

u uncover, underestimate, underline, underscore, understand, undertake, unnerve, unsettle, update, uphold, upset, urge, use, utter

v venture, verify, vex, videotape, view, vilify, visualize, voice, volunteer, vote, vow

w wager, wallow, want, warn, warrant, watch, weep, weigh, welcome, wheeze, whimper, whine, whisper, whoop, will, wish, witness, wonder, worry, worship, wound, wow, write

y yawn, yearn, yell, yelp

B Validation normalization

To normalize the acceptability judgments collected in the replication experiment (Section 3), we fit an ordinal (linked logit) mixed effects model to the ratings from both datasets, with fixed effects for VERB, FRAME, and their interaction and random unconstrained cutpoints for each participant (for further background on ordinal models, see Gelman & Hill 2006; Agresti 2012). This model is implemented in `tensorflow` (Abadi et al. 2015.

This procedure is analogous to the more familiar (within linguistics) approach of z-scoring by participant, then taking the mean of the scores for a particular verb-frame pair. The main difference between the two methods is in how they model the way that participants make responses on the basis of some "true" continuous acceptability. Both methods associate each participant with a different way of binning the continuous acceptability scale (usually modeled as isomorphic to the real values) to produce an ordinal response—the first bin corresponding to a 1 rating, the second corresponding to a 2 rating, etc. They differ in that z-scoring assumes that these bins are of equal size (for a particular participant)—the inverse of which is generally estimated via the standard deviation of the raw ordinal ratings (viewed as interval data)—whereas an ordinal model with unconstrained cutpoints (for each participant), assumes the bins can be of varying sizes.

We select the particular normalization method we use on the basis of empirical findings presented in White et al. 2018 (the paper whose data we validate against in Section 3). White et al. compare the fit to their data of six different possible ordinal models, varying in 3 respects: (i) whether the bins corresponding to each rating are of constant size or vary in size; (ii) whether the bins are centered around 0 for all participants or each participant has a different center (*additive* participant effects); and (iii) whether the size of the bins stays constant across participants or can be expanded or contracted depending on the participant (*multiplicative* participant effects). They point out that z-scoring corresponds to the model wherein the bins are of constant size but where there are both additive and multiplicative participant effects.

They fit each of these models with fixed effects for VERB, FRAME, and their interaction—effectively, each pairing of a verb v and a frame f is associated with some continuous acceptability value $a_{vf} = \beta_v + \beta_f + \beta_{vf}$, which is jointly optimized with parameters representing the bins.[1] They find that, even after penalizing for model complexity using both the Akaike Information Criterion (Akaike 1974: AIC;) and the Bayesian Information Criterion (BIC; Schwarz 1978), the model with varying bin sizes and additive and multiplicative participant effects fits the data substantially better than any other model, including the one corresponding to the assumptions of z-scoring (constant bin sizes and additive and multiplicative participant effects). We thus use a normalization method that assumes varying bin sizes.

We parameterize this method by assuming that each pairing of a verb v and a frame f is associated with some true real-valued acceptability a_{vf} (as described above) and that each participants p is associated with a way of binning these real-valued acceptability judgments, where each bin corresponds to a particular scale rating. These bins are defined by cutpoints $\mathbf{c}_p$ for each participants p, where the bin corresponding to the worst rating—in

[1] Steps must be taken to ensure identifiability, but how this is done is not important for current purposes.

our case, 1—is to the interval $(-\infty, c_{p1}]$ and the bin corresponding to the best ratings—in our case, 7—is the interval (c_{p6}, ∞). For all other ratings i, the corresponding bin for participant p is $(c_{p(i-1)}, c_{pi}]$. Alternatively, we say that $c_{p0} = -\infty$ and $c_{p7} = \infty$ for all participants p.

Similar to a binary logistic regression, which one can think of as having just two bins defined by a single cut point, we define the probability of a particular participant p giving a response r_{pvf} to verb v and frame f (assuming true acceptability a_{vf}) based on these cutpoints. First, we define the cumulative density function.

$$\mathbb{P}(r_{pvf} \leq i) = \text{logit}^{-1}\left(c_{pi} - a_{vf}\right)$$

Then, from the cumulative density function, we can reconstruct the probability for each response i.

$$\mathbb{P}(r_{pvf} = i) = \mathbb{P}(r_{pvf} \leq i) - \mathbb{P}(r \leq (i-1))$$

From this, the (log-)likelihood of the data immediately follows. This likelihood is the measure we use as a measure of variability in the main text, since the lower this likelihood is for a particular verb-frame pair, the less able the model is to "explain" the participants' responses using a single value a_{vf}, even after adjusting for differences in how the participant bins the scale.

We estimate the true acceptabilities $\mathbf{A}$ for all verb-frame pairs and the cutpoints for all participants $\mathbf{C}$ by using gradient descent to maximize the sum of the likelihood of the data, an Exponential prior on the distance between the cutpoints (thereby making this a mixed effects model), and a small smoothing term, under the constraint that the mean of the third cutpoint is locked to zero, thus making the parameters identifiable. All analyses use the resulting acceptabilities $\mathbf{A}$.

A reader may still wonder if there are empirical consequences to this choice of normalization method in contrast to z-scoring, even if this normalization is better theoretically and empirically motivated. In Appendix C we briefly explore this further, and show that using z-scoring produces scores that are highly correlated with the ordinal model-based method in the data at issue here.

C MegaAcceptability normalization

As for our replication of White et al.'s dataset, to measure interannotator agreement, we compute the Spearman rank correlation between the responses for each pair of participants that did the same list. This yields a

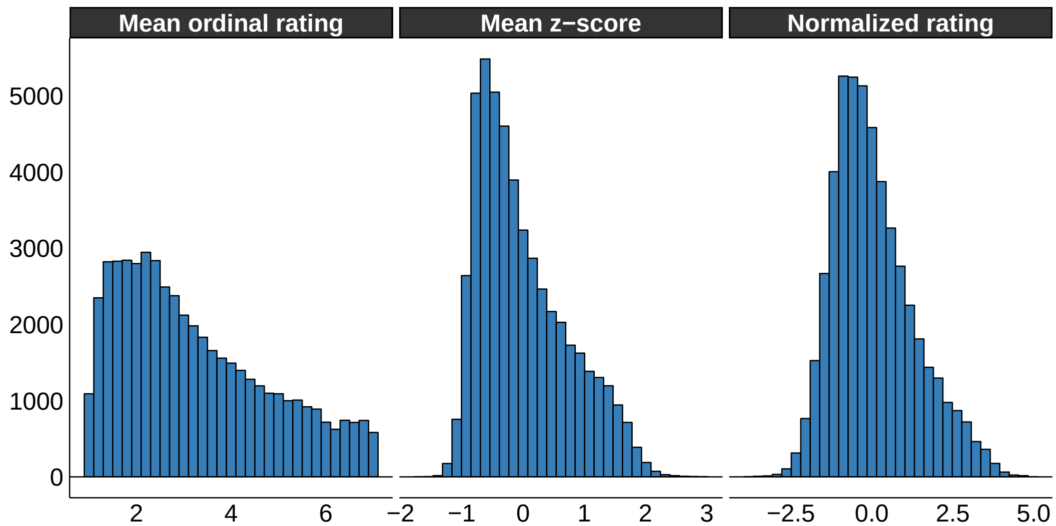

Figure 1: Marginal distribution across all verb-frame pairs of different acceptability scores.

mean correlation of 0.416 (95% CI: [0.413, 0.419]), which is more than 10 points lower than the agreement obtained in the replication.

Part of the reason for this is likely that White et al.'s—and consequently, our replication—contained mostly high frequency verbs, whereas the MegaAttitude dataset contains many low frequency verbs that participants are likely less certain about.[2] Another source of this low agreement is likely a higher rate of poor participants in these data. This is evidenced by the fact that the agreement scores have nontrivial left skew, with a median correlation of 0.455 (95% CI: [0.451, 0.458]).

To mitigate the effect of poor participants, we downweight the influence of those participants' responses in constructing the normalized acceptability for each verb-frame pair. Our approach amounts to using the ordinal model-based normalization described in Section 3, but weighting the likelihood of

[2] This reasoning is supported by an additional validation experiment we conducted investigating the 30 verbs discussed in White et al. 2018 in a majority of the frames used in MegaAcceptability. We find that agreement among participants was similar to that in the validation experiment reported in Section 3 ($\rho = 0.56$; 95% CI = [0.53, 0.59]). The two authors additionally annotated all the items in this validation themselves. Computing agreement by list, we agree with participants at $\rho = 0.55$ (95% CI = [0.52, 0.58]), averaging across lists, and with each other at 0.70 (95% CI = [0.62, 0.78]), averaging across lists.

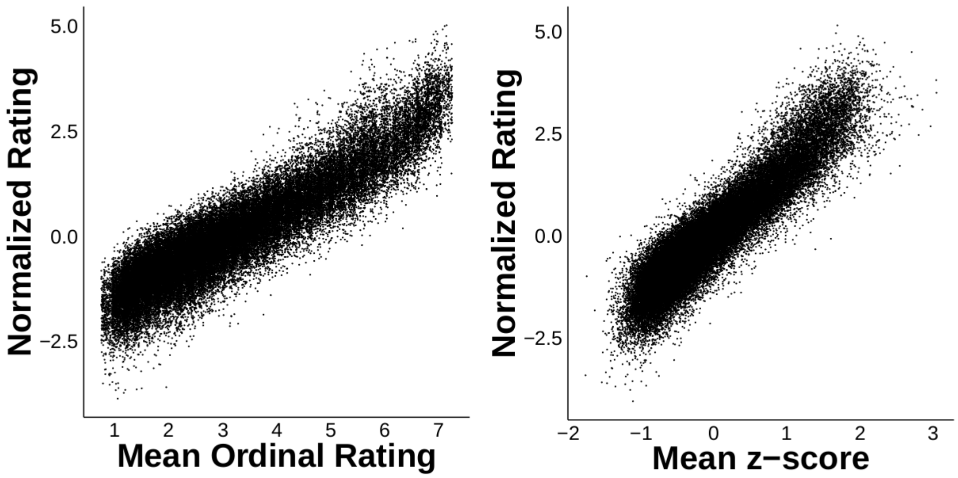

Figure 2: Relationship between mean ordinal responses (viewed as interval data) and normalized ratings produced by ordinal mixed model for particular verb-frame pairs (left) and relationship between mean of responses z-scored by participant and normalized ratings produced by ordinal mixed model for particular verb-frame pairs (right). Each point corresponds to a verb-frame pair.

the ordinal model by participant quality scores on [0, 1] deri____ed from pairwise agreement between participants.[3]

One simple way of deriving such a score would be to take the mean interannotator agreement for all pairs an participant occurs in and then normalize those means to lie on [0, 1]. This simple approach is problematic, however, since most participants only rate one list and so, if a good participant rates a list rated by mostly bad participants, that participant will be assigned a low quality score.

To address this issue, we derive a participant quality score by first fitting a linear mixed effects model with random intercepts for participant and list to the Spearman rank correlations—using `lme4` (Bates et al. 2015)—then extracting the Best Linear Unbiased Predictors for the participant intercepts. We then z-score these scores and squash them to [0, 1] using the normal cumulative distribution function. This participant quality score is thus high

[3] This procedure differs from the procedure used by White & Rawlins (2016) for the same dataset in that they filter participants with agreement under a particular threshold. Our approach can be seen as a soft version of their thresholding approach, wherein the influence of participants' responses drops off smoothly as a function of their overall agreement with other participants.

when an participant tends to show high agreement with other participants, adjusting for the effect of the particular list.

We combine these log-likelihoods into single variability score by computing their mean, weighted by the participant quality score of the participant who provided the rating.

Figure 1 shows the marginal distribution of ratings using the above method as well as two other common methods: (i) taking the mean of the ordinal responses (viewed as interval data) for each verb-frame pair (*mean ordinal rating*); and (ii) taking the mean of the ratings z-scored by participant for each verb-frame pair.

Figure 2 plots the corresponding joint distributions—i.e. the relationship between the resulting normalized value for each verb-frame pair and the mean of the ordinal responses for that pair (left) as well as the mean of the responses z-scored by participant (right). The Pearson correlation between the normalized value for each verb-frame pair and the mean of the ordinal responses (viewed as interval data) for that pair is 0.92, and the correlation between the normalized value for each verb-frame pair and the mean of the responses z-scored by participant is 0.95.

D Method for adding verbs

Seven verbs—*manage, fail, neglect, refuse, help, opt, deserve*—were unintentionally excluded from our large-scale experiment due to a coding error. We do not include these verbs in the analyses presented in the body of the paper because it is nontrivial, within the method described above, to build lists that include them without reconducting a large portion of the study.

Because we would like to have data about these verbs for future work, we instead evaluate an alternative method for adding missing verbs to our dataset. In this method, we test a single verb in all of the frames of interest within the same list.

To evaluate how this method compares to to a method wherein verbs are intermixed, we constructed a list for each of the 30 pilot verbs from Section 3 paired with each of the 50 frames from the MegaAcceptability data (Section 4). We find that the average pairwise agreement by list is actually higher in this experiment than in our original replication, with a median Spearman rank correlation of 0.65 (95% CI = [0.63 , 0.67]). This higher agreement is due to a few annotators who did many lists showing high agreement with each other, since when we fit the linear mixed effects model described in Appendix C to these correlations, we find an expected correlation of 0.54,

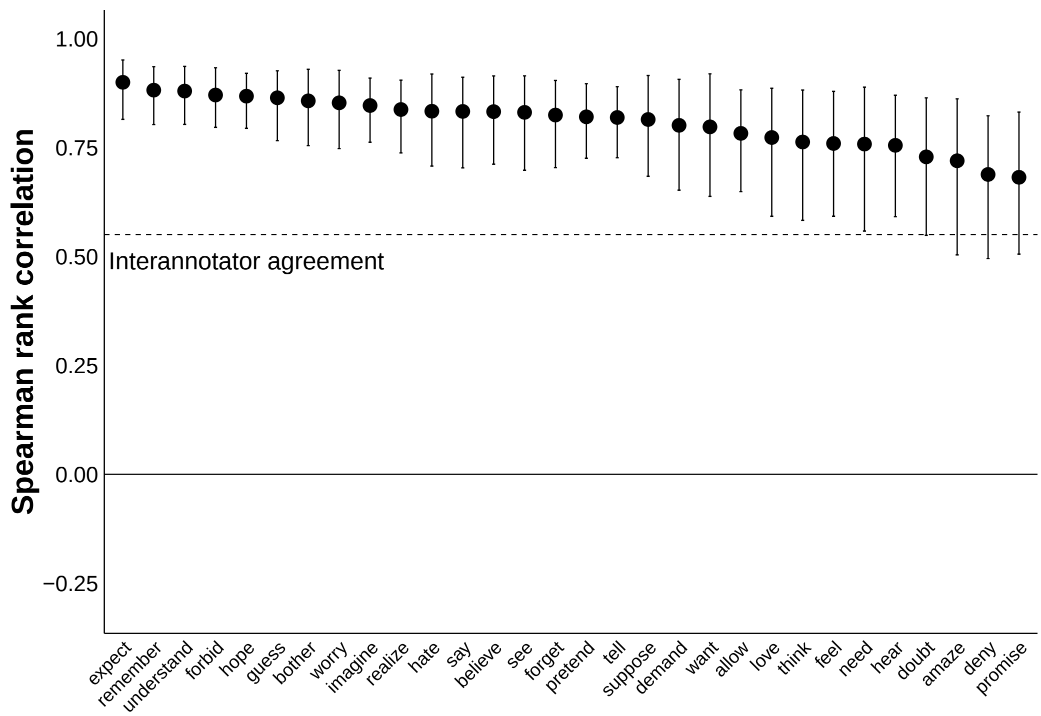

Figure 3: Correlation by verb between mean normalized verb-frame acceptability in MegaAcceptability and one-verb-per-list dataset. The dashed line shows mean interannotator agreement.

which is very close to the correlation found in our validation experiments (Section 3).

To compare the agreement between the normalized ratings from the MegaAcceptability dataset to those from this one-verb-per-list dataset, we applied the normalization used for the MegaAcceptability dataset (Appendix C) to these data and then computed the correlation by verb. Figure 3 shows this agreement which is extremely high across all verbs.

We take this as an indicator that testing one verb per list—at least in this set of frames—produces results that are just as valid as intermixing verbs. We thus tested the seven verbs above using this method. The resulting dataset is available on megaattitude.io.

References

Abadi, Martín, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu

Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu & Xiaoqiang Zheng. 2015. TensorFlow: Large-scale machine learning on heterogeneous systems. https://www.tensorflow.org/.

Agresti, Alan. 2012. *Categorical data analysis*. Hoboken, NJ: Wiley-Interscience 3rd edn.

Akaike, Hirotugu. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19(6). 716–723. https://doi.org/10.1109/tac.1974.1100705.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. https://doi.org/10.18637/jss.v067.i01.

Gelman, Andrew & Jennifer Hill. 2006. *Data analysis using regression and multilevel/hierarchical models*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511790942.

Schwarz, Gideon. 1978. Estimating the dimension of a model. *The Annals of Statistics* 6(2). 461–464. https://doi.org/10.1214/aos/1176344136.

White, Aaron Steven, Valentine Hacquard & Jeffrey Lidz. 2018. Semantic information and the syntax of propositional attitude verbs. *Cognitive Science* 42(2). 416–456. https://doi.org/10.1111/cogs.12512.

White, Aaron Steven & Kyle Rawlins. 2016. A computational model of S-selection. *Semantics and Linguistic Theory* 26. 641–663. https://doi.org/10.3765/salt.v26i0.3819.