

John E. Hayes^{1,2}, Alissa L. Allen^{1,2} and Samantha M. Bennett^{1,2,3}

¹ Sensory Evaluation Center, ² Department of Food Science, College of Agricultural Sciences, The Pennsylvania State University, University Park, PA. ³ Current address: Product Guidance and Insights, General Mills, Minneapolis, MN.

Background

- Category and Visual analog scales have been used for a century to generate quantitative data in perceptual, affective, or attitudinal research.
- Generalized scales have become extremely common for chemosensory research in the last decade, based on the premise they generate ratio data, and increase the validity of comparisons across individuals who cannot be randomly assigned to groups (Bartoshuk et al 2003; Snyder et al 2006). In particular, the general Labeled Magnitude Scale (gLMS) has been used hundreds of times in the chemosensory literature.
- In using the gLMS over the last decade, we have observed some participants treat the labels as categories, marking only at the adjectives, or on a pencil and paper ballots, even circling the words themselves.
- One solution to prevent clustering around the adjectives would be to strip the gLMS of internal adjectives (semantic labels), while keeping the labels at the ends of the scales (Dionne, et al., 2005; Snyder, et al., 2006). This version, the generalized Visual Analog Scale (gVAS), has only seen limited usage in the literature (Timpson et al 2007; Pickering et al 2010).
- To date, there are no reports comparing the unstructured gVAS to the gLMS in the same participants. In pilot analysis of existing data, we found the rank order of imagined sensations from the orientation was preserved, and the correlation across group means was very high. However, we also observed that there was systematic deviation in the raw values obtained from the two scales. As different participants were used in each study, it was unclear if the differences were due to the participants, the study context or the scale structure.
- Additionally, categorical behavior was clearly present in gLMS data, but this may have arisen from abbreviated instructions that did not explicitly stress participants should use the space in between the semantic labels.
- Here, we present 2 studies designed to answer the following questions:
 - do intensity ratings for orientation items and sampled stimuli vary across the gLMS and gVAS,
 - does changing the wording in the participant instructions alter clustering behavior on the gLMS,
 - which scale is easier to use, and
 - what is the relative resolving power between the two scales?

Methods

Participants – Reportedly healthy, non-smoking adults (aged 18-45 years) were recruited from the Penn State community. Procedures were IRB approved, informed consent was obtained, and participants were paid for their time.

Procedures – Orientation items and sampled stimuli were constant across all conditions. The orientation included 15 imagined or remembered sensations; oral and non-oral items were used to establish a generalized context. Stimuli (10mL) were presented in plastic soufflé cups at room temperature. Participants were instructed to take the entire sample into his or her mouth, swish for 3 seconds, and then spit prior to rating. Participants rinsed with 20C reverse osmosis water between every trial. Data were collected via Compusense *five* (Guelph ONT).

Study 1 Design – Participants (n=87) were randomized into one of three conditions (n=29 per condition).
 Condition 1 – using the gLMS with explicit instructions taken from Green (2002);
 Condition 2 – using the gLMS with instructions that did not specifically emphasize rating between the individual adjective labels (denoted as implicit here);
 Condition 3 – same instructions as Condition 2, but using the unstructured gVAS.

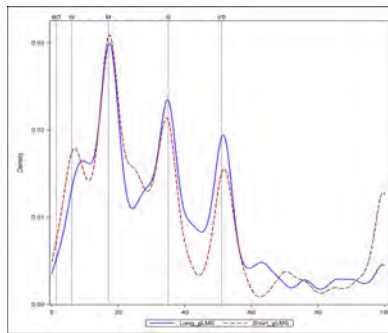
Study 1 Stimuli – Participants rated sweetness, sourness, burning/stinging, bitterness, umami/savory and saltiness for four samples: 0.5M sucrose, 0.41mM quinine hydrochloride, 0.56mM potassium chloride, and 25uM natural capsaicin.

Study 2 Design – Participants (n=58) completed a within-subjects crossover design. Participants attended two sessions one week apart: half used the gLMS first and half used the gVAS first. The implicit instructions from Experiment 1 were used for both scales.

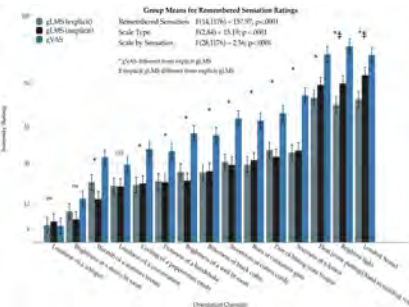
Study 2 Stimuli – Participants were asked to rate the overall intensity for five sucrose concentrations: 0.19M, 0.24M, 0.30M, 0.37M and 0.47M.

Experiment 1

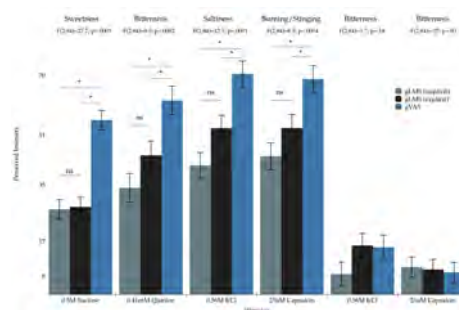
Finding 1: gLMS data exhibited clustering at the adjective labels, regardless of whether participants received explicit (blue) or implicit (red) instructions



Finding 2: For the orientation, raw values on the gVAS were higher than values obtained with the gLMS, and the effect was more noticeable for more intense sensations

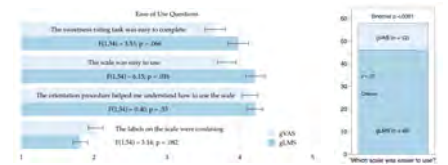


Finding 3: For sampled stimuli, gVAS ratings were higher than implicit gLMS or explicit gLMS ratings

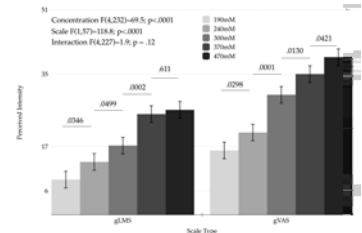


Experiment 2

Finding 4: Orientation was helpful across both scales, but users preferred the gLMS to the gVAS



Finding 5: Both scales revealed differences across sucrose stimuli, but there was some evidence for compression due to a false ceiling on the gLMS



Finding 6: Raw gVAS ratings were consistently higher than raw gLMS ratings, but the ratios between stimuli were preserved across both scales

Discussion

- Here, we find that the gVAS produces data similar to, but not identical with, data generated with the gLMS. Raw data obtained with the gVAS were consistently higher than ratings collected with the gLMS, and this was true for both imagined sensations during the orientation procedure and for sampled tastants and irritants.
- Notably, differences in ratings across scales disappeared when data were standardized to a cross-modal reference from the orientation (not shown). Also, Study 2 shows the *ratios* of intensity ratings were preserved across scales (eg, 300mM sucrose was ~1.9 times more intense than 190mM sucrose, regardless of which scale was used).
- Consistent with anecdotal reports, gLMS data were not normally distributed, as participants exhibited substantial clustering of responses near the verbal labels. Moreover, providing explicit written instructions to rate between the adjectives was *not* successful in reducing clustering. As would be expected given the lack of interior labels on the gVAS, we did not observe any evidence of clustering.
- Regarding resolving power, there was no clear advantage: both scales allowed participants to differentiate between sucrose samples. In leave-one-out resampling, the gLMS had a small advantage (not shown), but the gLMS also showed some evidence of a false ceiling at 'very strong'. That is, the inability to distinguish between the top two concentrations above are consistent with scale compression.
- Naive participants in a university setting clearly preferred the gLMS over the gVAS, reporting that the gLMS was easier to use. Whether this extends to other populations is unknown and needs to be tested. (This work was supported by an NIH Grant DC010904 to the lead author.)