Atten Percept Psychophys (2017) 79:1860–1869
DOI 10.3758/s13414-017-1334-5

Judgments of differences and ratios of subjective heaviness

Sergio Cesare Masin[1] · Andrea Brancaccio[1]

Published online: 17 May 2017

Abstract Experimental instructions to judge differences or ratios of subjective heaviness numerically are generally assumed to produce judgments linearly proportional to the respective heaviness differences or heaviness ratios. In this study, participants were instructed to numerically judge the difference or ratio of heaviness between two weights being lifted separately, either unimanually or bimanually. Weight values were combined factorially. Patterns of factorial curves revealed that unimanual lifting triggered linear judgments of heaviness differences, whereas bimanual lifting triggered nonlinear judgments of heaviness ratios. Lifting conditions produced these judgments independently of the instruction specifications to judge differences or ratios. These results suggest the interpretation that unimanual lifting triggers linear judgments of heaviness differences by default, whereas bimanual lifting triggers nonlinear judgments of heaviness ratios learned through experience. Implications for sensory measurement are noted.

Keywords Subjective heaviness · Difference judgment · Ratio judgment · Integration psychophysics · Sensory measurement

It is widely assumed that people instructed to judge sensory differences numerically produce judgments proportional to sensory differences and that people instructed to judge sensory ratios numerically produce judgments proportional to sensory ratios (Gescheider, 1997; Hutchings, 1999; Lawless, 2013; Marks, 1974; Meilgaard, Carr, & Civille, 2016; Stevens, 1975; Zwislocki, 2009). However, empirical data obtained from various sensory dimensions suggest that the majority of people produce judgments proportional to sensory differences when they execute the task to judge sensory ratios (Atkinson & Ward, 1972; Fagot & Stewart, 1969; Masin, 2014; Parker & Hickman, 1990; Ross & Di Lollo, 1971; Schneider, Parker, Farrell, & Kanow, 1976; Westermann, 1982). We investigated whether this could be the case for the sensory dimension of heaviness. This sensory dimension may be special, because, as was first broadly suggested by Fullerton and Cattell in 1892, it is plausible that people have learned in their past to conform their judgments of heaviness ratios to physical weight ratios through the numerical feedback received while they were weighing objects.

We tested this hypothesis using methods of integration psychophysics (Anderson, 1982, 1992, 1996, 2013; McBride, 1993; Weiss, 2006). Let Ψ_S and Ψ_V be the subjective heaviness of the first and second, respectively, of two objects lifted separately and unimanually. Let S and V, with $S < V$, be the weights that produce Ψ_S and Ψ_V, respectively. For each factorial combination of fixed values of S and V, participants were instructed to judge numerically the difference $\Psi_V - \Psi_S$ or the ratio Ψ_V/Ψ_S. Factorial curves were obtained by plotting judgments against V for each S, separately for difference or ratio instructions. The patterns of these factorial curves were used to assess the effects of the instructions.

Predictions

Instructions to judge heaviness differences produce judgments that may be modeled as

✉ Sergio Cesare Masin
scm@unipd.it

[1] Department of General Psychology, University of Padua, Via Venezia 8, 35131 Padova, Italy

$$J_D = c_1 \cdot (\Psi_V - \Psi_S) + c_2 \quad (1)$$

and instructions to judge heaviness ratios produce judgments that may be modeled as

$$J_R = c_3 \cdot \frac{\Psi_V}{\Psi_S} + c_4 \quad (2)$$

with c_1–c_4 parameters (Anderson, 1974). These models were tested as follows.

For the size and range of weights of stimuli used in the present study, magnitude estimation yields an exponent of the psychophysical power function for heaviness of 1.1 (Marks & Cain, 1972; Stevens & Rubin, 1970). This near-unity exponent indicates a nearly linear relation between heaviness and physical weight.[1] A fortiori, this relation may be considered linear in the present study because S and V varied in a narrow range.[2]

This practical linearity implies that Models 1 and 2 are equivalent to

$$J_D = c_5 \cdot (V - S) + c_2, \quad (3)$$

and

$$J_R = c_3 \cdot \frac{V}{S} + c_4, \quad (4)$$

respectively, with c_5 a parameter. Tests of Models 3 and 4 are indirect tests of Models 1 and 2, respectively.

Figure 1 serves to show the geometric implications of Models 3 and 4. These models were used to calculate J_D and J_R for each combination of values of S and V, with $c_3 = c_5 = 1$ and $c_2 = c_4 = 0$. The left diagram shows calculated J_D and the right diagram calculated J_R plotted against V for each S, respectively. Straight lines fit the calculated points.

Models 3 and 4 imply that the curves relating J_D or J_R to V are straight lines. Model 3 implies parallel, uniformly spaced straight lines. Model 4 implies divergent, hyperbolically spaced straight lines. The analysis of variance allows identifying which of these different geometric implications is empirically supported (Anderson, 1981, 1982; Weiss, 2006).

Models 3 and 4 predict a significant linear trend and non-significant quadratic trend for V (linear curves). Model 3 predicts a non-significant interaction (parallel curves) and a significant linear trend and non-significant quadratic trend for S (uniformly spaced curves). Model 4 predicts a significant interaction (diverging curves) and a significant quadratic trend for S (hyperbolically spaced curves).

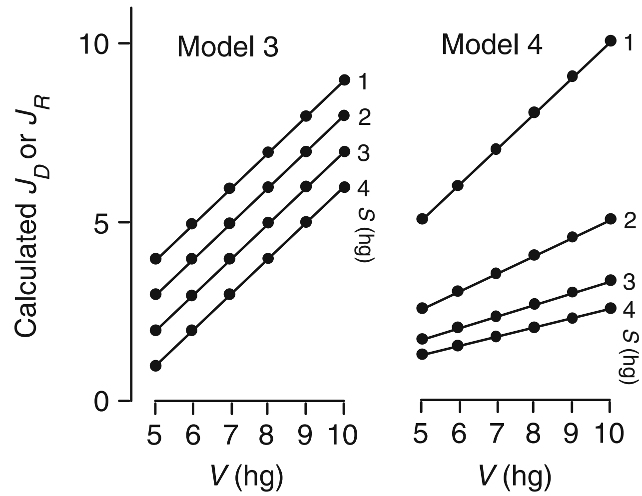

Fig. 1 Different patterns of straight lines fitting theoretical heaviness judgments calculated using Models 3 and 4

Experiment 1

Method

Participants

Twenty psychology students participated in the experiment. They were divided in two equally numerous groups: Groups 1 and 2.

Stimuli

Ten cylindrical smooth tin cans with 8.5-cm diameter and 10-cm height were used as stimuli. Four were called *standard* and had a weight S of 1, 2, 3, or 4 hg, and six were called *variable* and had a weight V of 5, 6, 7, 8, 9, or 10 hg. These weight values were fixed by filling stimuli with uniformly distributed pellets held steady by cotton wool.

The values of standards were combined factorially with the values of variables, making 24 stimulus pairs. These stimulus pairs were presented randomly two times consecutively. Each stimulus pair was presented as follows.

Participants saw no stimulus before or during the experiment. While sitting at a table, they faced a 135- × 62-cm opaque vertical screen with its base edge coinciding with the closest table edge. Two 17- × 9-cm holes were cut in the screen such that their base also coincided with the closest table edge, and they were spaced 18-cm apart symmetrically about

[1] It has been hypothesized that a power function with exponent m relates magnitude estimates of heaviness to heaviness and that another power function with exponent k relates heaviness to physical weight (Attneave, 1962; Rule & Curtis, 1982). It follows that a power function with exponent $n = m \cdot k$ relates magnitude estimates of heaviness to physical weight. Curtis, Attneave, and Harrington (1968) estimated m to be 1.14. This estimate, if reliable, and the finding that $n = 1.1$ imply that $k \approx 1$, that is, that heaviness is nearly linearly related to physical weight.

[2] A simple graphical analysis shows that, if the power function relating heaviness to physical weight should have an exponent k that differs largely from unity—for example, if k should be 0.5 or 1.5—this function would be practically indistinguishable from a linear function within the range of values of V from 5 to 10 hg used in the present study.

the screen's central axis. We refer to these holes as the left and right holes as viewed by the participant. On each trial, each stimulus was placed on the table on soft material approximately 10 cm from the left or right hole. Participants reached the stimuli through the holes.

Procedure

Participants in Groups 1 and 2 reached the stimuli through the left hole. Each trial consisted of the following orderly series of events: a standard was placed on the table; the participants inserted their preferred hand in the left hole, lifted the standard, and then retracted their hand from the left hole; a variable was placed on the table in place of the standard; the participants inserted their preferred hand in the left hole and lifted the variable.

To reduce the risk of mental calculations, the participants were never given any example of the numerical judgments they were instructed to produce during the experiment.

Group 1 received the following difference instructions: "On each trial, express the difference of the heaviness of the second can with respect to that of the first can as a number: the larger the difference the larger the number. Numbers may be integers or fractional numbers and must be 0 if the two cans have equal heaviness and greater than 0 if the second can is heavier. Please report if the first can feels heavier."

Group 2 received the following ratio instructions: "On each trial, express the ratio of the heaviness of the second can to the heaviness of the first can as a number. That is, judge how many times the second can is heavier than the first can. Numbers may be integers or fractional numbers and must be 1 if the two lifted cans have equal heaviness and greater than 1 if the second can feels heavier. Please report if the first can feels heavier."

The results showed that no participant in Group 2 and only one participant in Group 1 reported only once that the first can felt heavier. When this report occurred, the participant was asked to judge the difference of the heaviness of the first can with respect to that of the second can. The negative of the difference was taken as the participant's judgment.

Participants received four practice trials before the actual experiment.

Results

Figure 2 shows mean judgments obtained from unimanual lifting plotted against V separately for each S, for difference and ratio instructions.

Difference and ratio instructions produced patterns of roughly parallel factorial curves. Difference instructions produced a pattern of factorial curves statistically equivalent to a set of moderately non-equidistant straight lines barely converging rightward. Ratio instructions produced a pattern of

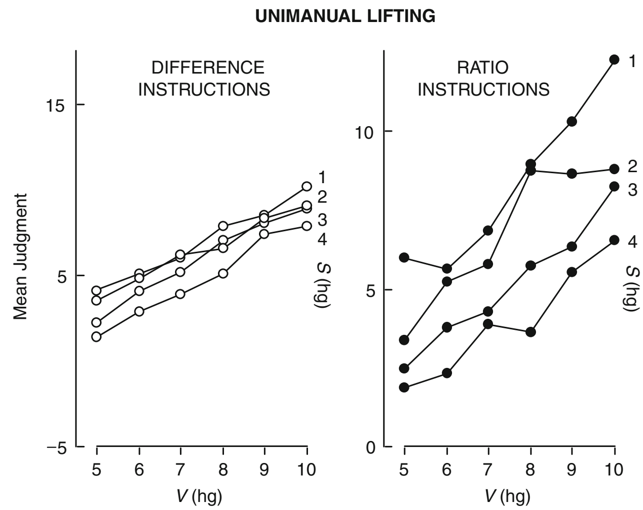

Fig. 2 Mean judgments produced under instructions to judge heaviness differences or to judge heaviness ratios between two objects lifted unimanually, with S the weight of the first object and V that of the second

factorial curves statistically equivalent to a set of parallel equidistant straight lines, replicating a prior finding (Masin, 2014). The finding that mean judgments were related linearly to V supports the above-mentioned indications that the power function relating heaviness to physical weight is practically indistinguishable from a linear function for values of V in the range of 5 to 10 hg.

Data analysis for difference instructions

The results were assessed using 4 (S) × 6 (V) analyses of variance. For three subgroups of five, four, and one participants, individual factorial curves were roughly parallel, converged rightward, and diverged rightward with nonsignificant, significant, and nonsignificant interactions, $F(15, 60) = 1.3$, $F(15, 45) = 2.1$ ($\eta_p^2 = 0.41$), $p < .05$, and $F(15, 15) = 2.2$, respectively. The participant with diverging factorial curves produced judgments in the range of 0–2000. He was excluded from further analyses.

The interaction for the group of the other nine participants was significant, $F(15, 120) = 1.74$ ($\eta_p^2 = 0.18$), $p = .05$, with significant linear and quadratic trends for S, $F(1, 8) = 11$ ($\eta_p^2 = 0.58$) and $F(1, 8) = 8$ ($\eta_p^2 = 0.50$), $p < .05$, and with a significant linear trend and not significant quadratic trend for V, $F(1, 8) = 20$ ($\eta_p^2 = 0.72$), $p < .005$, and $F(1, 8) < 1$, respectively.

Data analysis for ratio instructions

For three subgroups of five, three, and two participants, individual factorial curves were roughly parallel, converged rightward, and diverged rightward with no significant interactions, $F(15, 15) = 1.1$, $F(15, 30) = 1.7$, and $F(15, 60) < 1$, respectively.

For the entire group the interaction was not significant, $F(15, 135) = 1.3$, with significant linear trends, $F(1, 9) = 6.2$ ($\eta_p^2 = 0.41$), $p < .05$, and $F(1, 9) = 28$ ($\eta_p^2 = 0.76$), $p < .005$, and non-significant quadratic trends for both S and V, $F(1, 9) < 1$ and $F(1, 9) < 1$, respectively.

Discussion

It is widely assumed that people instructed to numerically judge heaviness differences produce judgments proportional to heaviness differences and that people instructed to numerically judge heaviness ratios produce judgments proportional to heaviness ratios (Cross & Rotkin, 1975; Donovan & Ross, 1969; Harper & Stevens, 1948; Ross & Di Lollo, 1970; Stevens & Rubin, 1970). Because the analyses of variance confirmed all the geometric implications of Model 3, the results suggest that most participants produced judgments linearly or nearly linearly proportional to heaviness differences both when they executed the instructions to judge heaviness differences and when they executed the instructions to judge heaviness ratios.

These results align with prior findings that people would actually respond to sensory differences when they execute the instructions to judge sensory ratios (Atkinson & Ward, 1972; Popper, Parker, & Galanter, 1986; Schneider, 1980; Ross & Di Lollo, 1971).

Under the instructions to judge heaviness ratios, the participants did not conform their judgments to physical weight ratios. It could be that this result was due to participants lifting the stimuli unimanually. Indeed, people interact with objects bimanually considerably more often than unimanually (Coté, 2014; Kotranza, Quarles, & Lok, 2006; Terrenghi, Kirk, Sellen, & Izadi, 2007). One may thus hypothesize that, by learned association, people could conform their numerical judgments of heaviness ratios to the respective physical weight ratios only or prevalently when they use two hands for lifting. The next experiment tested this hypothesis.

Experiment 2

Method

Participants

Twenty psychology students participated in the experiment. None had participated in Experiment 1. They were divided in two equally numerous groups: Groups 1 and 2.

Stimuli

The stimuli and the stimulus conditions were the same as those used for Experiment 1.

Procedure

For Groups 1 and 2 each trial consisted of the following orderly series of events: a standard and a variable were placed on the table; the participants inserted their left and right hands through the left and right holes, respectively; the participants simultaneously grasped and lifted the standard and the variable.

Groups 1 and 2 received the difference and ratio instructions used in Experiment 1 with "first can" and "second can" replaced by "left can" and "right can," respectively.

The results showed that in Groups 1 and 2 the number of participants who reported just 0, 1, 2, or 3 times and just 0, 1, 2, 3, or 5 times that the left can felt heavier was respectively 1, 2, 3, and 4 and 1, 3, 4, 1, and 1. In these cases, participants were asked to judge the difference or the ratio of the heaviness of the left can with respect to that of the right can. Negatives of judged differences and reciprocals of judged ratios were taken as the participant's judgments. Participants received four practice trials before the actual experiment.

Results

Figure 3 shows mean judgments obtained from bimanual lifting plotted against V separately for each S, for difference and ratio instructions. The patterns of factorial curves obtained from difference and ratio instructions were statistically equivalent to a set of diverging uniformly spaced straight lines.

Data analysis for difference instructions

The results were assessed by 4 (S) × 6 (V) analyses of variance using $p \leq 0.1$ as a criterion for marginal significance. For two subgroups of three and seven participants, individual factorial curves were roughly parallel and diverged rightward with non-significant and marginally significant interactions, $F(15, 30) < 1$ and $F(15, 90) = 1.67$ ($\eta_p^2 = 0.22$), $p = .07$, respectively. For the subgroup of 7 participants, the linear trend was marginally significant for both S, $F(1, 6) = 5.4$ ($\eta_p^2 = 0.48$), $p = .06$, and V, $F(1, 6) = 4.1$ ($\eta_p^2 = 0.41$), $p = .09$, whereas the quadratic trend was not significant for S, $F(1, 6) < 1$, and marginally significant for V, $F(1, 6) = 4.1$ ($\eta_p^2 = 0.41$), $p = .09$, respectively.

For the entire group the interaction was marginally significant, $F(15, 135) = 1.7$ ($\eta_p^2 = 0.16$), $p = .06$, with significant linear trends, $F(1, 9) = 6.4$ ($\eta_p^2 = 0.42$) and $F(1, 9) = 5.3$ ($\eta_p^2 = 0.37$), $p < .05$, and nonsignificant quadratic trends, $F(1, 9) < 1$ and $F(1, 9) = 2.5$, for both S and V, respectively.

Data analysis for ratio instructions

Individual factorial curves diverged rightward for each participant. For the entire group of participants the interaction was significant, $F(15, 135) = 4.1$ ($\eta_p^2 = 0.31$), $p < .001$, with a

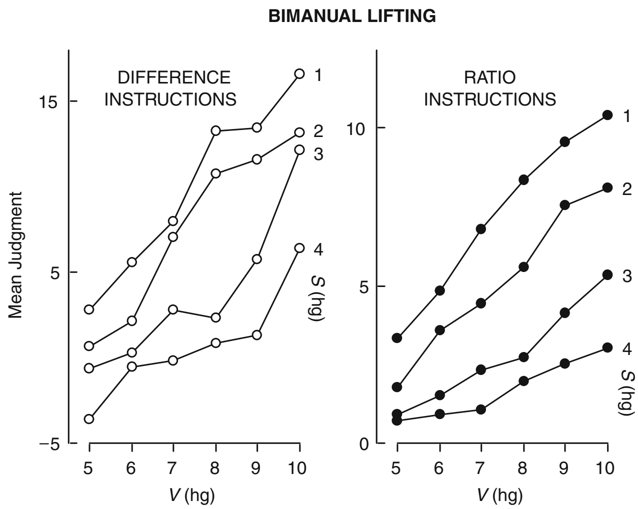

Fig. 3 Mean judgments produced under instructions to judge heaviness differences or to judge heaviness ratios between two objects lifted bimanually, with *S* the weight of the left object and *V* that of the right object

significant linear trend, $F(1, 9) = 26.5$ ($\eta_p^2 = 0.75$) and $F(1, 9) = 16.6$ ($\eta_p^2 = 0.65$), $p < .005$, and nonsignificant quadratic trend, $F(1, 9) = 2.7$ and $F(1, 9) < 1$, for both *S* and *V*, respectively.

Discussion

Nonlinear ratio judgments

Because the analyses of variance rejected the hypothesis of Model 4 that factorial curves should be spaced hyperbolically with significant quadratic trend for *S*, the results in Fig. 3 suggest that most participants produced judgments nonlinearly related to heaviness ratios.

What judgment operations produced nonlinear judgments of heaviness ratios causing the uniform spacing of factorial curves? Tentatively, we offer the following ideas.

We propose that people try to adjust their judgments of heaviness ratios to physical weight ratios through the use of abilities other than the ability to judge heaviness ratios. Let Ψ_S and Ψ_V be correspondingly the heaviness of the left and right of two stimuli lifted separately and bimanually. People could, for example, judge Ψ_V on a bounded scale and learn to vary the upper bound of this scale inversely with Ψ_S. These judgment operations may be modeled as

$$J_R^* = k_1 \cdot (k_2 - \Psi_S) \cdot \Psi_V + k_3, \tag{5}$$

with k_1–k_3 parameters and $k_2 > \Psi_S$.

For the reasons mentioned above this model is linearly equivalent to

$$J_R^* = k_4 \cdot (k_5 - S) \cdot V + k_3, \tag{6}$$

with k_4 and k_5 parameters and $k_5 > S$. Albeit incompletely, Model 6 mimics Model 4 because both models predict that factorial curves are diverging straight lines with slope varying inversely with *S*.

In Fig. 4 the diagram shows factorial curves calculated using Model 6 with $k_1 = 0.25$, $k_2 = 5$, and $k_3 = 0$. These curves are uniformly spaced in agreement with the empirical uniformly spaced diverging factorial curves in Fig. 3.

Lifting condition effects

The results showed that the lifting condition remarkably influenced participants' judgments. For most participants, unimanual lifting generated roughly parallel factorial curves whereas bimanual lifting generated diverging factorial curves. Parallelism and divergence of factorial curves occurred irrespective of the instruction specifications to judge ratios or differences.

These results were obtained without giving participants any example of the numerical judgments they were instructed to produce during the experiment. Prior experiments using factorial designs comparable to that used in the present study provided instructions to judge differences or ratios of heaviness that contained examples of the numerical judgments to be produced during the experiment (Birnbaum & Veit, 1974; Marks & Cain, 1972; Mellers, Davis, & Birnbaum, 1984; Rule, Curtis, & Mullin, 1981). Difference instructions were found to yield parallel factorial curves and ratio instructions to yield diverging factorial curves. Numerical examples alter ratio judgments considerably as shown by the change in exponent of the psychophysical power function produced by such examples (Baird, Kreindler, & Jones, 1971; Gibson & Tomko, 1972; Guirao, 1991; Robinson, 1976). Even more alarming is the possibility that numerical examples could cause mental calculations (Laming, 1997, p. 140). Indeed, some of the participants who received examples of how to produce numerical judgments during the experiment spontaneously reported that during the experiment they were ranking the stimuli and mentally calculated the difference or the ratio of ranks when they were expected to judge heaviness differences or ratios, respectively (Mellers et al., 1984). Other participants may have mentally calculated judgments without reporting it. Mental calculations may determine the pattern of factorial curves whether or not participants can judge differences or ratios.[3] In the present study, it seems plausible that the fact that participants were

[3] Empirical evidence also shows that simple feedback and even the mere verbal suggestion of a quantitative relationship between the subjective magnitudes of stimuli combined factorially may alter the pattern of factorial curves substantially. For example, factorial curves that are found to converge rightward or to be parallel before feedback or before verbal suggestion may diverge rightward after feedback or after verbal suggestion (Chasseigne, Lafon, & Mullet, 2002; Lafratta, 2007).

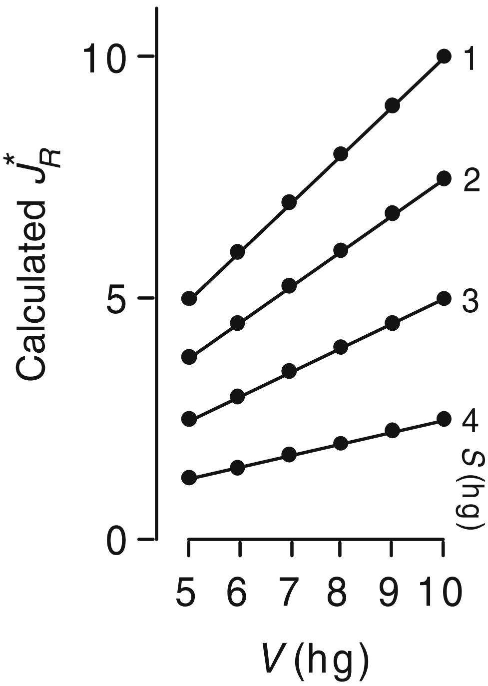

Fig. 4 Pattern of straight lines fitting theoretical heaviness judgments calculated using Model 6. Physically these lines are uniformly spaced while perceptually they look sparser at the bottom due to a geometric-optical illusion

given no example of numerical judgments minimized the risk of occurrence of mental calculations.

Why did the participants' judgments depend on the lifting condition and why did these judgments override the instruction specifications? We offer the following broad interpretation.

Previous empirical data suggest that people produce judgments of sensory differences by default when they execute the task to judge sensory ratios. For heaviness, people may try to convert their default judgments into the respective weight ratios through the numerical feedback received when they weigh objects. The present results suggest that converted ratio judgments are nonlinear approximations to heaviness ratios. People would make these judgments only or prevalently in association with using two hands for lifting because people interact with objects bimanually considerably more frequently than unimanually. It may be that unimanual lifting triggers judgments of heaviness differences by default and that bimanual lifting triggers approximate judgments of heaviness ratios by learned association.

Experiment 3

The heaviness of a lifted object could vary with that of a contextual lifted object (Dinnerstein, Curcio, & Chinsky, 1966; Williams, Ross, & Di Lollo, 1966). Models 1, 2, and 5 assume invariance of Ψ_S with V and of Ψ_V with S. But these assumptions were untested: S and V may have contextually influenced Ψ_V and Ψ_S, respectively, and these perceptual effects may have influenced the judgments of $\Psi_V - \Psi_S$ or Ψ_V/Ψ_S. The next experiment tested this hypothesis.

Method

Participants

Thirty-four psychology students participated in the experiment. None had participated in Experiments 1 or 2.

Stimuli

Stimuli and stimulus presentation conditions were the same as those used for Experiment 1. A plastic can of 0.2 hg and a tin can of 16 hg, equal in size to the stimuli, were used as anchors.

Procedure

On each trial, participants lifted only a standard with their left hand, only a variable with their right hand, or both a standard and a variable with their left and right hands, respectively. The anchors were presented once immediately before the first trial. Half the participants lifted the anchors simultaneously, the lighter anchor in the left hand and the heavier anchor in the right hand. The other participants lifted the anchors successively, first the heavier anchor with their right hand and then the lighter anchor with their left hand.

For each trial, participants were asked to rate the heaviness of a standard (Ψ_S) or of a variable (Ψ_V) on a 1–100 scale with 1 the lighter and 100 the heavier anchor heaviness.

The 4 standards and 6 variables lifted singularly and the 24 factorial combinations of standard and variable were presented twice consecutively, with stimuli and combinations of stimuli in random intermixed order. Only the standards were rated in one presentation of 24 combinations of standard and variable and only the variables in the other presentation.

Results

Let R_S and R_V be the ratings of Ψ_S and Ψ_V, respectively. Figure 5 shows mean $R_V - R_S$ plotted against V for each S, for unimanual or bimanual lifting.

Minor findings

In Fig. 5, in the left diagram the curvature at the right end of curves reveals a slight end-anchor effect. The quadratic trends for S and V were not significant and significant, $F(1, 33) = 2.9$, and $F(1, 33) = 20$ ($\eta_p^2 = 0.38$), $p < .001$, respectively. In the right diagram the spacing between curves slightly increased with S. The quadratic trends for S and V were significant and

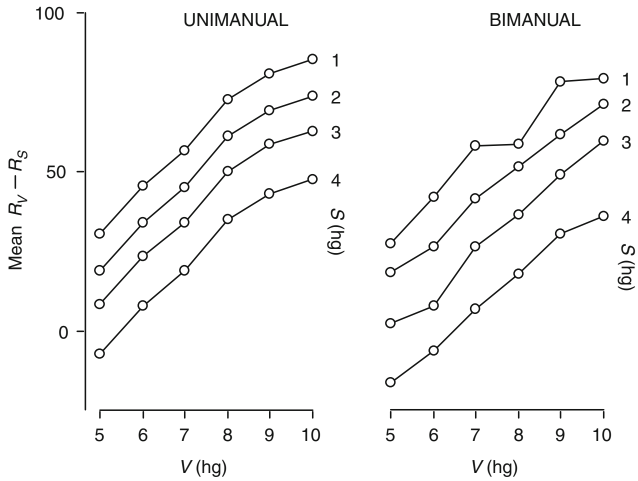

Fig. 5 Results derived from ratings of heaviness, R_V and R_S, of stimuli with weight V or S lifted by one hand while the other hand was holding no stimulus or a stimulus with weight S or V, respectively. For unimanual or bimanual lifting, each curve relates the mean calculated difference $R_V - R_S$ to V for a different S

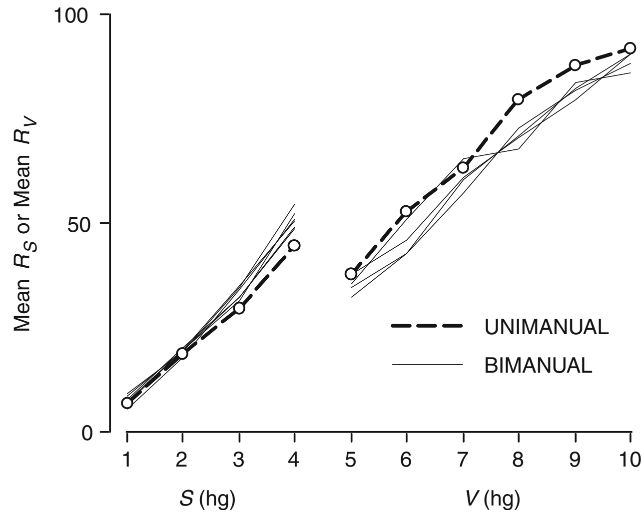

Fig. 6 The left curves show mean rated heaviness R_S plotted against stimulus weight S while the right curves show mean rated heaviness R_V plotted against stimulus weight V for unimanual (dashed curves) or bimanual lifting (solid curves). The left solid curves show the results separately for each V and the right solid curves show the results separately for each S

not significant, $F(1, 32) = 11$, $p < .005$, and $F(1, 32) = 2.8$, respectively.

Main finding

In Fig. 5, in the right diagram the near parallelism of curves shows that mean $R_V - R_S$ varied with $V - S$ at a rate independent of S. Thus, the contextual influences that S and V may have had on Ψ_V and Ψ_S did not differentially influence the judgments of $\Psi_V - \Psi_S$ or Ψ_V/Ψ_S.

A 2 (anchor presentation mode) × 4 (S) × 6 (V) analysis of variance supported this finding. The interaction of S with V, $F(15, 480) < 1$, the effect of anchor presentation mode, $F(1, 32) < 1$, and the interactions involving this factor, $F(3, 96) < 1$, $F(5, 160) < 1$, and $F(15, 480) = 1.5$, were not significant.

Supplementary findings

Figure 6 helps to discuss the following three supplementary findings. The two dashed curves show the results for unimanual lifting and the solid curves show those for bimanual lifting. The left dashed curve shows mean R_S plotted against S while the right dashed curve shows mean R_V plotted against V. The left solid curves show mean R_S plotted against S for each V and the right solid curves show mean R_V plotted against V for each S.

The results for unimanual lifting confirm previous findings (Shen, 1936). The dashed curves show that a standard lifted by the left hand would have been felt lighter if it had been lifted by the right hand.

Bimanual lifting generated no significant heaviness contrast effect. Two 4 (S) × 6 (V) analyses of variance were carried out on the data that yielded the solid curves in Fig. 6. For the left solid curves, the effect of V and the interaction were not significant, $F(1, 165) < 1$ and $F(15, 495) < 1$, respectively. For the right solid curves, the effect of S and the interaction were also not significant, $F(3, 99) < 1$ and $F(15, 495) = 1.3$, respectively.

Bimanual lifting generated a significant heaviness assimilation effect. The left dashed and solid curves show that Ψ_S was larger for standards lifted together with a variable than for standards lifted singularly. The right dashed and solid curves show that Ψ_V was smaller for variables lifted together with a standard than for variables lifted singularly.

As may be seen in Fig. 5, heaviness assimilation caused a moderate but significant expansion of mean $R_V - R_S$ in bimanual lifting. A 4 (S) × 7 (V, including V presented without S) analysis of variance was performed on the data that yielded the left dashed and solid curves in Fig. 6, whereas a 5 (S, including S presented without V) × 6 (V) analysis of variance was performed on the data that yielded the right dashed and solid curves in Fig. 6. For the left dashed and solid curves, the effect of V was significant and the interaction not significant, $F(6, 198) = 2.4$ ($\eta_p^2 = .07$), $p < .05$, and $F(18, 594) = 1.4$, respectively. For the right dashed and solid curves, the effect of S was significant and the interaction not significant, $F(4, 132) < 5.9$ ($\eta_p^2 = 0.15$), $p < .001$, and $F(20, 660) = 1.3$, respectively.

Discussion

Large evidence shows that ratings of sensory intensity are linearly related to sensory intensity when floor and ceiling effects are removed, typically by using two anchor stimuli: one larger than the largest experimental stimulus and one smaller that the smallest experimental stimulus (Anderson, 1996, pp. 94–96, 2013; Hofmans, Mairesse, & Theuns, 2007; Masin & Toffalini, 2009). That is, $R_V = a \cdot \Psi_V + b$ and $R_S = a \cdot \Psi_S + b$ yielding that $R_V - R_S = a \cdot (\Psi_V - \Psi_S)$ with a and b parameters. Thus, in Fig. 5 the results for bimanual lifting suggest that both mean $R_V - R_S$ and $\Psi_V - \Psi_S$ varied with $V - S$ at a rate constant with S. This rate of variation of $R_V - R_S$ was statistically constant notwithstanding the occurrence of heaviness assimilation.

Curve parallelism in both Figs. 2 and 5 supports the idea that curves in Fig. 2 were parallel due to the linear relation of judgments of $\Psi_V - \Psi_S$ with $\Psi_V - \Psi_S$ itself. Divergence and uniform spacing of curves in Fig. 3 and curve parallelism in Fig. 5 support the interpretation that divergence and uniform spacing of curves in Fig. 3 may have resulted from participants producing nonlinear approximate judgments of heaviness ratios.

General discussion

People's abilities to judge differences or ratios of heaviness were investigated using methods from integration psychophysics (Anderson, 1982, 1996, 2013; McBride, 1993; Weiss, 2006). Participants were instructed to lift separately two objects, either unimanually or bimanually, and to judge numerically the difference or ratio of heaviness between the objects for each factorial combination of weights of the objects. The patterns of factorial curves obtained when difference or ratio judgments of heaviness were used as the dependent variable yielded the following findings.

The hypothesis that judgments of heaviness differences are related linearly to heaviness differences predicts that factorial curves form a pattern of parallel straight lines. In unimanual lifting, obtained factorial curves formed a pattern statistically equivalent to a set of parallel or closely parallel straight lines both when participants judged differences and when they judged ratios of heaviness. These results reveal that judgments of heaviness differences were related linearly or nearly linearly to heaviness differences both when the instructions required judging heaviness differences and when the instructions required judging heaviness ratios.

The finding of linear or nearly linear judgments of heaviness differences was confirmed by the further finding that the judgments of heaviness differences (Fig. 2) and the heaviness differences themselves (Fig. 5) varied with the larger of the compared stimulus weights at a rate that was constant or nearly constant with the smaller stimulus weight. These results confirm previous findings, obtained for various sensory dimensions, that the majority of people judge sensory differences under the task of judging sensory ratios (Atkinson & Ward, 1972; Parker & Hickman, 1990; Ross & Di Lollo, 1971; Schneider et al., 1976; Westermann, 1982).

The hypothesis that judgments of heaviness ratios are related linearly to heaviness ratios predicts hyperbolically spaced factorial curves. The factorial curves obtained from bimanual lifting formed a pattern statistically equivalent to a fan of diverging uniformly spaced straight lines, both when heaviness differences were judged and when heaviness ratios were judged. Because the obtained factorial curves were uniformly spaced rather than being hyperbolically spaced, participants' judgments were nonlinear approximations to heaviness ratios.

This finding of nonlinear judgments of heaviness ratios agrees with the abovementioned empirical evidence suggesting that people may not be able to judge sensory ratios and with the hypothesis specific for the heaviness sensation that people learn to compensate for this lack of ability by using alternative judgment operations. Model 5 provides an example of alternative operations yielding judgments that nonlinearly approximate heaviness ratios, consistently with the present results.

In unimanual and bimanual lifting, the findings that judgments of heaviness differences and heaviness ratios may involve nonequivalent processes of weight information integration may be interpreted broadly as follows. In general, people appear to judge sensory differences by default, that is, irrespective of whether they are instructed to judge differences or ratios. It is plausible that people have in their lives often attempted to convert their default heaviness judgments into the corresponding physical weight ratios when they received feedback about these weight ratios while they were weighing objects. The present results suggest that these converted judgments were nonlinear approximations to heaviness ratios. People may learn to associate these converted judgments prevalently to the use of two hands, because they interact with objects bimanually considerably more frequently than unimanually. Judgments of heaviness differences would thus be linear in unimanual lifting by default, whereas judgments of heaviness ratios would be nonlinear in bimanual lifting by learned association.

The assumption that people can judge ratios of sensory intensities is used largely in psychology to derive sensory psychophysical functions (Baird, 1997; Florentine, Popper, & Fay, 2010; Gescheider, 1997; Laming, 2004; Marks, 1974; Zwislocki, 2009). These functions typically are intended to provide valid ratio-scale measures of sensory intensity (Stevens, 1975). Such ratio-scale measures are often also assumed to allow for the comparison across individuals of intensities of various psychological dimensions related to well-being such as, among many others, pain, perceived stress from life events, perceived exertion, and quality of food (Borg & Kaijser, 2006; Lawless & Heymann, 2010; Rahe, Veach, Tolles, & Murakami, 2000; Ruskin, Amaria, Warnock, & McGrath, 2011). However, comparisons across individuals based on judgments deriving from instructions to judge ratios may be valid only if people can accurately judge ratios. The present results support Models 1 and 5, suggesting that people lack the ability to judge ratios. That is, under the instructions to judge heaviness ratios, people's judgments appear to be either linear representations of heaviness differences or nonlinear representations of heaviness ratios. These results indicate that psychophysical functions based on ratio judgment provide inaccurate ratio-scale measures of sensory intensity. Instead, with Ψ_S constant, Model 1 implies that $J_D = c_1 \cdot \Psi_V + c_6$ and Model 5 that $J^*_R = k_6 \cdot \Psi_V + k_3$ with c_6 and k_6 parameters. That is, these models imply that judged heaviness differences and judged heaviness ratios are valid measures of heaviness on an interval scale.[4] The present results thus support the idea that sensory psychophysical functions based on instructions to judge ratios can only provide accurate or nearly accurate interval-scale measures of sensory intensity.

Acknowledgements We would like to thank Norman H. Anderson, Scott Parker, and an anonymous reviewer for constructive comments on an earlier draft of this paper.

[4] Using feedback, participants may be trained to produce measures of heaviness on a ratio scale. We first note that, using feedback, participants may be trained to produce specific judgments of sensory ratios (King & Lockhead, 1981; Marks, Galanter, & Baird, 1995; West, Ward, & Khosla, 2000). For example, for each of different values of Ψ_V, on each training trial one may present a fixed value of Ψ_S, the participant may be asked to judge Ψ_V/Ψ_S, and then a value J_F for this ratio calculated using the equation $J_F = m \cdot V^n$, with m and n arbitrary parameters, may be finally presented to the participant as feedback. For fixed arbitrary values of m and n, participants progressively learn to convert their judgments of Ψ_V/Ψ_S into the respective value of J_F. Given these results, participants may be trained to produce judgments of heaviness ratios as follows. For each factorial combination of S and V, participants may be asked to judge Ψ_V/Ψ_S and then a valid ratio calculated by Model 4 with fixed arbitrary parameters may be presented as feedback. The conversion of participant's judgments into valid ratio judgments would be complete when, in line with Model 4, the resulting factorial curves have become spaced hyperbolically.

References

Anderson, N. H. (1974). Algebraic models in perception. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. 2, pp. 215–298). New York: Academic Press.

Anderson, N. H. (1992). Integration psychophysics and cognition. In D. Algom (Ed.), *Psychophysical approaches to cognition* (pp. 14–113). Amsterdam: Elsevier.

Anderson, N. H. (1981). *Foundations of information integration theory*. New York: Academic Press.

Anderson, N. H. (1982). *Methods of information integration theory*. New York: Academic Press.

Anderson, N. H. (1996). *A functional theory of cognition*. Hillsdale, NJ: L. Erlbaum Associates.

Anderson, N. H. (2013). Unified psychology based on three laws of information integration. *Review of General Psychology, 17,* 125–132.

Atkinson, W. H., & Ward, L. M. (1972). On the internal consistency of magnitude estimations. *Perceptual and Motor Skills, 35,* 903–913.

Attneave, F. (1962). Perception and related areas. In S. Koch (Ed.), *Psychology: A study of a science* (Vol. 4, pp. 619–659). New York, NY: McGraw-Hill.

Baird, J. C. (1997). *Sensation and judgment: Complementarity theory of psychophysics*. Mahwah, NJ: Erlbaum.

Baird, J. C., Kreindler, M., & Jones, K. (1971). Generation of multiple ratio scales with a fixed stimulus attribute. *Perception & Psychophysics, 9,* 399–403.

Birnbaum, M. H., & Veit, C. T. (1974). Scale convergence as a criterion for rescaling: Information integration with difference, ratio, and averaging tasks. *Perception & Psychophysics, 15,* 7–15.

Borg, E., & Kaijser, L. (2006). A comparison between three rating scales for perceived exertion and two different work tests. *Scandinavian Journal of Medicine & Science in Sports, 16,* 57–69.

Chasseigne, G., Lafon, P., & Mullet, E. (2002). Aging and rule learning: The case of the multiplicative law. *American Journal of Psychology, 115,* 315–330.

Coté, C. A. (2014). Haptic exploration in elementary school age children. *OTJR: Occupation, Participation and Health, 34,* 4–11.

Cross, D. V., & Rotkin, L. (1975). The relation between size and apparent heaviness. *Perception & Psychophysics, 18,* 79–87.

Curtis, D. W., Attneave, F., & Harrington, T. L. (1968). A test of a two model for magnitude estimation. *Perception & Psychophysics, 3,* 25–31.

Dinnerstein, D., Curcio, F., & Chinsky, J. (1966). Contextual determination of apparent weight as demonstrated by the method of constant stimuli. *Psychonomic Science, 5,* 251–252.

Donovan, R. J., & Ross, J. (1969). The dimensionality of judgments of heaviness. *Perception & Psychophysics, 6,* 247–249.

Fagot, R. F., & Stewart, M. (1969). Tests of product and additive scaling axioms. *Perception & Psychophysics, 5,* 117–123.

Florentine, M., Popper, A. N., & Fay, R. R. (Eds.). (2010). *Loudness*. New York: Springer.

Fullerton, G. S., & Cattell, J. M. (1892). *On the perception of small differences with special reference to the extent, force and time of movement*. Philadelphia: University of Pennsylvania Press.

Gescheider, G. A. (1997). *Psychophysics: The fundamentals*. Mahwah, NJ: Erlbaum.

Gibson, R. H., & Tomko, D. L. (1972). The relation between category and magnitude estimates of tactile intensity. *Perception & Psychophysics, 12,* 135–138.

Guirao, M. (1991). A single scale based on ratio and partition estimates. In S. J. Bolanowski & G. A. Gescheider (Eds.), *Ratio scaling of* *psychological magnitude: In honor of the memory of S. S. Stevens* (pp. 59–78). Hillsdale: Erlbaum.

Harper, R. S., & Stevens, S. S. (1948). A psychological scale of weight and a formula for its derivation. *American Journal of Psychology, 61,* 343–351.

Hofmans, J., Mairesse, O., & Theuns, P. (2007). An alternative for prescribed integration rules in testing the linearity of a response measure. *Teorie & Modelli, 12*(1-2), 259–268.

Hutchings, J. B. (1999). *Food color and appearance*. Gaithersburg, MD: Aspen Publishers.

King, M. C., & Lockhead, G. R. (1981). Response scales and sequential effects in judgment. *Perception & Psychophysics, 30,* 599–603.

Kotranza, A., Quarles, J., & Lok, B. (2006). Mixed reality: Are two hands better than one? In *Proceedings of the ACM symposium on virtual reality software and technology* (pp. 31–34). New York: Association for Computing Machinery.

Lafratta, A. (2007). On the activation of information integration rules. *Teorie & Modelli, 12*(1–2), 53–62.

Laming, D. (1997). *The measurement of sensation*. New York: Oxford University Press.

Laming, D. (2004). *Human judgement: The eye of the beholder*. London: Thompson.

Lawless, H. T. (2013). *Quantitative sensory analysis: Psychophysics, models and intelligent design*. Hoboken, NJ: Wiley-Blackwell.

Lawless, H. T., & Heymann, H. (2010). *Sensory evaluation of food: Principles and practices*. New York: Springer.

Marks, L. E. (1974). *Sensory processes: The new psychophysics*. New York: Academic Press.

Marks, L. E., & Cain, W. S. (1972). Perception of intervals and magnitudes for three prothetic continua. *Journal of Experimental Psychology, 94,* 6–17.

Marks, L. E., Galanter, E., & Baird, J. C. (1995). Binaural summation after learning psychophysical functions for loudness. *Perception & Psychophysics, 57,* 1209–1216.

Masin, S. C. (2014). Test of the ratio judgment hypothesis. *Journal of General Psychology, 141,* 130–150.

Masin, S. C., & Toffalini, E. (2009). On the bisection operation. *Attention, Perception, and Psychophysics, 71,* 1566–1663.

McBride, R. L. (1993). Integration psychophysics: The use of functional measurement in the study of mixtures. *Chemical Senses, 18,* 83–92.

Meilgaard, M. C., Carr, B. T., & Civille, G. V. (2016). *Sensory evaluation techniques*. Boca Raton, FL: CRC Press.

Mellers, B. A., Davis, D. M., & Birnbaum, M. H. (1984). Weight of evidence supports one operation for "ratios" and "differences" of heaviness. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 216–230.

Parker, S., & Hickman, D. (1990). Heaviness ratio estimates follow weight ratios for some subjects. In F. Müller (Ed.), *Proceedings of the sixth annual meeting of the International Society for Psychophysics* (pp. 100–105). Würzburg, Germany: International Society for Psychophysics.

Popper, R., Parker, S., & Galanter, E. (1986). Dual loudness scales in individual subjects. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 61–69.

Rahe, R. H., Veach, T. L., Tolles, R. L., & Murakami, K. (2000). The stress and coping inventory: An educational and research instrument. *Stress Medicine, 16,* 199–208.

Robinson, G. H. (1976). Biasing power law exponents by magnitude estimation instructions. *Perception & Psychophysics, 19,* 80–84.

Ross, J., & Di Lollo, V. (1970). Differences in heaviness in relation to density and weight. *Perception & Psychophysics, 7,* 161–162.

Ross, J., & Di Lollo, V. (1971). Judgment and response in magnitude estimation. *Psychological Review, 78,* 515–527.

Rule, S. J., & Curtis, D. W. (1982). Levels of sensory and judgmental processing: Strategies for the evaluation of a model. In B. Wegener (Ed.), *Social attitudes and psychophysical measurement* (pp. 107–122). Hillsdale, NJ: Erlbaum.

Rule, S. J., Curtis, D. W., & Mullin, L. C. (1981). Subjective ratios and differences in perceived heaviness. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 459–466.

Ruskin, D., Amaria, K. A., Warnock, F. F., & McGrath, P. A. (2011). Assessment of pain in infants, children, and adolescents. In D. C. Turk & R. Melzack (Eds.), *Handbook of pain assessment* (pp. 213–241). New York: Guilford Press.

Schneider, B. (1980). Individual loudness functions determined from direct comparisons of loudness intervals. *Perception & Psychophysics, 28,* 493–503.

Schneider, B., Parker, S., Farrell, G., & Kanow, G. (1976). The perceptual basis of loudness ratio judgments. *Perception & Psychophysics, 19,* 309–320.

Shen, N. C. (1936). Handedness effects in simultaneous lifting of weights by both hands. *Journal of Experimental Psychology, 19,* 64–76.

Stevens, J. C., & Rubin, L. L. (1970). Psychophysical scales of apparent heaviness and the size-weight illusion. *Perception & Psychophysics, 8,* 225–230.

Stevens, S. S. (1975). *Psychophysics: Introduction to its perceptual, neural, and social prospects*. New York: Wiley.

Terrenghi, L., Kirk, D., Sellen, A., & Izadi, S. (2007). Affordances for manipulation of physical versus digital media on interactive surfaces. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1157–1166). New York: Association for Computing Machinery.

Weiss, D. J. (2006). *Analysis of variance and functional measurement: A practical guide*. New York: Oxford University Press.

West, R. L., Ward, L. M., & Khosla, R. (2000). Constrained scaling: The effect of learned psychophysical scales on idiosyncratic response bias. *Perception & Psychophysics, 62,* 137–151.

Westermann, R. (1982). Empirical test of scale type resulting from the power law for heaviness. *Perceptual and Motor Skills, 55,* 1167–1173.

Williams, M. Y., Ross, J., & Di Lollo, V. (1966). Sensory contrast effects in the judgment of lifted weights. *Psychonomic Science, 6,* 137–138.

Zwislocki, J. J. (2009). *Sensory neuroscience: Four laws of psychophysics*. New York: Springer.