Citation:
Saccone EJ, Goldsmith RM, Buckingham G, Chouinard PA (accepted on 4/8/2019) Container size exerts a stronger influence than liquid volume on the perceived weight of objects. Cognition. DOI: 10.1016/j.cognition.2019.104038.

School of Psychology and Public Health, La Trobe University, Melbourne, Victoria, Australia

Sport and Health Sciences, College of Life and Environmental Sciences, University of Exeter, Exeter, United Kingdom

Corresponding author:
Dr Elizabeth J. Saccone
School of Psychology and Public Health
La Trobe University
Edwards Road, Flora Hill
Victoria 3552 Australia
Ph: +61(0) 3 5444 7489
Email: e.saccone@latrobe.edu.au

Many features of an object can influence how we predict and perceive its weight. The current study evaluated the relative contributions of sensory and conceptual processing of object features on weight perception. We employed a novel paradigm to investigate how container size and the amount of liquid inside can influence the perceived weight of bottles and the forces deployed when lifting them. Stimulus pairs always had the same mass but could vary in liquid volume (full vs half-full bottle) or size (large vs small bottle; size-weight illusion (SWI)). In Experiment 1, participants lifted the stimuli via strings, which served to isolate the influence of visual from kinaesthetic information about the size of stimuli on perception and lifting behaviour. In Experiment 2, participants lifted the stimuli via handles that were attached directly to the objects. This lifting style is more likely to include deviations from true vertical lifting, which should theoretically provide more kinaesthetic information about the size of the stimuli. Experiment 1 did not produce any weight illusion. Experiment 2 produced a weight illusion but only when container size differed. Thus, liquid volume did not influence perceived weight when container size was held constant in either experiment. Curiously, additional control experiments revealed that participants could not discriminate between the different sized bottles solely from the kinaesthetic information received from a handle-based lift, suggesting that size might be processed differently when making explicit perceptual judgements about it than when influencing weight perception. Together, these findings suggest that weight illusions are driven more strongly by the kinaesthetic processing of stimulus features than predictions arising from conceptual weight cues.

1. Introduction

The ability to predict an object's weight is key to our skilled interactions with the physical world. When acting on an unfamiliar object, weight is predicted based on past experience with similar objects or objects with similar features and an appropriate motor response is selected (Gordon, Westling, Cole, & Johansson, 1993). These predictive processes allow us to interact with the objects around us with little delay between planning and carrying out actions. Research has shown that an object's size is particularly influential in predicting its weight (Gordon, Forssberg, Johansson, & Westling, 1991).

Research examining the forces deployed when lifting an object demonstrates that size provides a strong basis for sensorimotor predictions. Researchers can measure both the horizontal force applied by the fingers to grasp an object (grip force) and the vertical force applied by the muscles of the arm during lift-off (load force). The initial degree of force deployed is controlled by feedforward models that operate on predictions of heaviness (Chouinard, Leonard, & Paus, 2005; Johansson & Westling, 1984) and reflect the expectation of how much force is required to lift the object. In particular, forces deployed the first time an individual lifts an unfamiliar object reflects *sensorimotor predictions* of the object's weight. For example, when lifting two novel objects of different sizes, greater force is typically applied to the larger one because larger objects frequently weigh more (Buckingham & Goodale, 2010b; Buckingham & Goodale, 2010c; Flanagan & Beltzner, 2000; Gordon et al., 1991; Grandy & Westwood, 2006).

Predictions of weight based on an object's size can also influence how heavy it feels when lifted. To illustrate, in a size-weight illusion (SWI) experiment, two or more objects appear identical except for their size, but have been engineered to have the same mass (Charpentier, 1886, 1891). In this unusual context, the expectation that the larger object is heavier is violated and instead, the smaller of the two SWI objects feels heavier (Buckingham & Goodale, 2013; Charpentier, 1886, 1891; Chouinard, Large, Chang, & Goodale, 2009; Flanagan & Beltzner, 2000). One explanation for the illusion relates to our conceptual understanding of the relationship between size and weight (Buckingham & Goodale, 2013; Buckingham & MacDonald, 2016). We learn early in life that the larger of two objects usually weighs more and this association is continually reinforced during our lifetime (Buckingham & Goodale, 2010c, 2013). When sensory input contradicts expectations, the opposite perceptual experience occurs in that the smaller object is felt as heavier. Examining the forces deployed when first lifting SWI objects tends to reflect this expectation of weight on the basis of size (Buckingham & Goodale, 2010c; Buckingham, Ranger, & Goodale, 2011b; Flanagan & Beltzner, 2000; Grandy & Westwood, 2006). During the initial lifts, which represent sensorimotor predictions of weight, greater force is typically deployed for the larger object than is needed, indicating that the object was lighter than expected. The reverse is typically true for the smaller object in that too little force is applied. A landmark study by Flanagan and Beltzner (2000) demonstrated that after a small number of lifts, the forces tend to change to reflect the objects' true (equal) weight (see Buckingham & Goodale, 2010b; Buckingham et al., 2011b; Grandy & Westwood, 2006 for replications).

The particular mechanism by which weight expectations influence perception is unknown (e.g., Buckingham & MacDonald, 2016). It may be that expectations contribute to perception in a top-down manner (Buckingham, 2014; Ellis & Lederman, 1998) whereby some conceptual understanding of the object's features must be computed prior to lifting, providing context to influence the subsequent perceptual experience of weight based on predicted weight. However, recently Plaisier, Kuling, Brenner, and Smeets (2019) provided evidence against this idea. They demonstrated that participants still experienced the SWI even when they did not become aware of the objects' size until after the lift had started (i.e., up to 400 ms after lift-off).

Regardless, converging evidence from other weight illusions demonstrates that predicted weight can influence perceived weight (see Saccone & Chouinard, 2019b for a review). In these other weight illusions, a stimulus feature other than size is varied and it is this other feature that provides a *conceptual cue* to the object's weight – that is, a cue that allows us to predict weight based on the learned association between that feature and weight. These conceptual cues to weight may depend primarily on the formation of semantic associations, which are operationally defined as knowledge acquired during our lifetime (Tulving, 1972). Weight illusion paradigms have exploited conceptual cues such as objects' apparent material, density and unique identity – cues that normally predict weight differences but fail to do so in a contrived, weight illusion context when mass is held constant across the stimuli. When these typical feature-weight associations are violated, the opposite perceptual experience results, in a consistent manner to the SWI.

To provide an example, in the material-weight illusion (Buckingham, Cant, & Goodale, 2009; Ellis & Lederman, 1999; Seashore, 1899; Wolfe, 1898), objects have the same size and mass but vary in their apparent material. For example, participants might be presented with two, equally-sized cubes that appear to be comprised of either Styrofoam or brass, but have in fact been modified to have the same mass. Styrofoam objects are normally lighter than brass objects but in the context of the illusion, where the objects weigh the same, the former is felt heavier. Thus, these other weight illusions provide converging evidence that predictions based on learned, conceptual feature-weight associations typically produce the opposite perceptual experience when those predictions are not met.

Although size can serve as a conceptual weight cue, Saccone and Chouinard (2019b) have reviewed evidence that size may also influence weight perception in a different manner. In brief, one reason that size may be different is that it can have a remarkably stronger and more consistent effect on illusory weight perception than other conceptual cues (e.g., Buckingham, Bieńkiewicz, Rohrbach, & Hermsdörfer, 2015; Buckingham & Goodale, 2013; Buckingham, Goodale, White, & Westwood, 2016; Saccone, Landry, & Chouinard, 2019). Another reason is that the effect of size on perceived weight seems to vary with sensory input modality. To explain, in the SWI, information about size is gained through some combination of visual, haptic, and kinaesthetic input. For example, when an object is viewed and held in the hand, its size is processed visually. Haptic information about its size is also received through the pressure exerted on the touch receptors in the skin. Kinaesthetic information is also obtained from the hand and arm muscles providing feedback signals to the brain about the object's size and weight as the object is stabilised in the air. If object features influence perceived weight via semantic associations, then perception should not vary with the sensory modality providing the same feature information – in the way that patients with semantic dementia are equally impaired in processing semantic content of stimuli regardless of sensory input modality (Marinkovic et al., 2003; Patterson, Nestor, & Rogers, 2007).

However, research demonstrates that a lift including haptic and somatosensory information about size produces a stronger SWI than visual information. Specifically, researchers have isolated the influence of visual information about size from somatosensory information by having participants lift stimuli via a string or pulley system (Buckingham, Milne, Byrne, & Goodale, 2015; Ellis & Lederman, 1993; Masin & Crestoni, 1988; Wolf, Bergmann Tiest, & Drewing, 2018). These paradigms still elicit the

SWI; however, this visual-based version of the SWI tends to be weaker than when size is processed via somatosensation (Saccone et al., 2019; Wolf et al., 2018). Thus, size may not influence perception in the same manner as other conceptual weight cues that are based on semantic associations. The latter may depend more on low-level sensory mechanisms, with somatosensory channels exerting a stronger influence than visual channels (Saccone et al., 2019).

Of interest, Plaisier and Smeets (2015) compared the effect of size on the SWI with a related conceptual weight cue: volume of material. They varied overall size but not volume of material by employing “spacer” objects. These objects each consisted of two small, shallow blocks (60 mm x 60 mm x 18 mm) that were connected by a cylindrical bar, akin to a miniature dumbbell. They varied the length of the connecting bar such that the two blocks were set different distances from one another. Thus, the overall size (i.e., outer edges) of the objects varied while volume of material was held constant, therefore providing roughly the same conceptual (i.e., volume) information about weight. Using these objects, the authors reported patterns of weight perception that mirrored the SWI – the “smaller” objects (i.e., blocks set closer together) were perceived as heavier than the “larger” ones (i.e., blocks set further apart). This study supports the view that the overall size of an object has a particularly strong influence on weight perception, more so than other features that provide conceptual information about weight.

However, Plaisier and Smeets’ (2015) stimuli were rather unusual. It is uncertain how well participants could predict the weight of these unfamiliar dumbbell objects. Predictions were conceivably more difficult for these stimuli than for objects we typically handle, which are solid rather than separated. We are not accustomed to predicting the weight of objects for which overall size varies but the volume of material does not. On the contrary, we are much more accustomed to handling objects that have a constant size and changing content, for example, mugs, bottles, bags, suitcases. Given that predicted weight can influence perceived weight, the perceivers' familiarity of the objects is a relevant consideration. Although the authors' logic is sound – that volume of material should and does provide a conceptual cue to weight – perhaps when such unfamiliar or unusual objects are used, predictions relating to size simply override other weight cues such as volume of material. However, the idea that size cues exert a stronger influence on perception for unfamiliar objects would have to be tested with more common, ecologically valid objects, such as those used in the present investigation. Furthermore, Plaisier and Smeets (2015) did not measure forces applied during lifting and therefore did not examine how the properties of their task objects might influence *sensorimotor* predictions, which is arguably another important consideration for understanding weight perception (Dijker, 2014). In short, it is unclear as to why size influenced perception so strongly above volume of material in Plaisier and Smeets' (2015) study. Is it because size represents a more familiar, well established weight cue in the context of such unusual objects, because size influences perception in a different way to other conceptual cues (Saccone & Chouinard, 2019b), or because size exerts different sensorimotor predictions than volume?

To answer these questions, the present study examined predicted and perceived weight based on size cues compared with a highly familiar, conceptual cue to weight: liquid volume content. Liquid volume is a visual weight cue that is encountered frequently in daily life. Experience tells us that a full bottle of milk should be considerably heavier and require more force to lift than a bottle that only contains the last few mouthfuls. In this vein, Nowak and Hermsdörfer (2003) demonstrated that participants' altered their maximum force and rates of force applied when lifting objects containing different levels of visible liquid volume. The containers had a constant size and therefore the liquid content was the feature providing the cue to weight. Although liquid volume is a highly familiar, conceptual cue to weight, it has not been examined in an illusory weight context before.

Thus, the current study employed a novel paradigm to compare the effects of container size and liquid volume on perceived weight. Sensorimotor predictions of weight were also examined by measuring forces deployed during lifting. Stimuli were a pair of large, identical semi-transparent bottles; one appeared full of liquid and the other appeared half-full, but they were manipulated to have the same mass (see bottles A and B in Figure 1). The results from this stimulus pair were compared to a control experiment, in which different participants lifted a pair of bottles that met criteria for a SWI; namely the large, full bottle described above (bottle A in Figure 1), as well as a small, full bottle (bottle C in Figure 1) that had the same mass as the large one. Importantly, the small, full bottle was intended to have the same apparent liquid content as the half-full bottle. Thus, the higher-order, conceptual cue of liquid volume content was held constant across the two stimulus pairs, whereas overall size, which may influence weight perception in other ways above and beyond providing a conceptual cue, was varied only in the SWI (large and small) pair.

First, we hypothesised that liquid volume would influence both sensorimotor predictions (i.e., reflected in forces deployed during the initial lifts) and weight estimates throughout the entire testing session. Specifically, liquid content may provide a cue to weight such that greater lifting forces are deployed in initial trials for the full than half-full bottle. Furthermore, in line with the perceptual experience of the SWI and other weight illusions, the half-full bottle may be perceived as heavier than the full bottle.

Second, we hypothesised that container size would influence sensorimotor predictions (i.e., forces deployed during initial lifts) and exert an even stronger influence on perceived weight. Specifically, we predicted the typical findings with respect to sensorimotor predictions based on size in that greater lifting forces would be deployed in initial trials for the large compared to small bottle. We also predicted a stronger effect on perception for the large and small pair, meaning that there would be a greater difference in perceived heaviness between the large and small bottles than between the full and half-full bottles. In line with Flanagan and Beltzner's (2000) landmark study that measured both forces and perception over the course of the experiment, we further predicted the difference in lifting forces across the stimuli to adapt over time, converging to reflect the true (i.e., equal) weights of the stimuli.

This study comprises the following experiments. The first experiment used a paradigm in which participants lifted the bottle pairs via strings. This was the first SWI experiment to measure lifting forces using a strings-based lift. The second experiment was identical to the first except that the stimuli were lifted via a handle that was attached directly to the lids of the bottles. This lifting method is in line with many studies that have measured forces, including Flanagan and Beltzner (2000). After conducting these two experiments, we carried out additional control experiments to further understand the results obtained in the two main experiments.

2. Experiment 1

Experiment 1a determined if the conceptual weight cue of liquid volume would a) produce an illusory weight experience (i.e., the half-full bottles would feel heavier than the full bottle of equal mass) and b) demonstrate the typical pattern of sensorimotor predictions demonstrated in earlier studies (i.e., more forces would initially be applied to the full bottle than the half-full one and that differences in forces between the two bottles would disappear as the experiment progressed). In Experiment 1a, the conceptual cue of liquid volume varied while the size of the bottles was held constant (full vs half-full bottles). These results were then compared to Experiment 1b, in which both liquid content and bottle size varied (SWI; large vs small bottles). Based on previous studies, we hypothesised that participants would perceive the SWI and that more forces would initially be applied to the larger object than the smaller one, but that force differences would dissipate as the experiment progressed.

If both container size and liquid volume content influence predicted and perceived weight via conceptual cues, then visual information should suffice in producing a perceptual weight illusion for both stimulus pairs. Thus, in Experiments 1a and 1b, participants lifted the bottles via strings, in line with other SWI studies that have isolated visual information about size using this lifting style (Buckingham, Milne, et al., 2015; Ellis & Lederman, 1993; Masin & Crestoni, 1988; Wolf et al., 2018). Previous studies using string-based lifts have never measured fingertip forces. Instead, forces have been measured in paradigms where participants lifted stimuli via a force transducer handle attached directly to the objects (Buckingham, Bieńkiewicz, et al., 2015; Buckingham & Goodale, 2010a, 2010c, 2013; Buckingham, Goodale, et al., 2016; Buckingham et al., 2011b; Buckingham, Ranger, & Goodale, 2012; Flanagan & Beltzner, 2000; Grandy & Westwood, 2006; Mon-Williams & Murray, 2000). In these cases, there could still be a degree of kinaesthetic feedback about stimulus size and distribution of mass from the torques applied during the lift. Conversely, lifting the same objects with strings should, in theory, eliminate this information completely as it only allows a truly vertical lift. The current study is the first to investigate forces applied to lift SWI objects when only the visual modality provides size information.

2.1 Method

2.1.1 Experiment 1a: Full and half-full bottles lifted with strings.

2.1.1.1 Participants. Fourteen right-handed individuals (8 females, 6 males; age: M = 24.5 years, SD = 7.4 years) from the La Trobe University community were recruited for the experiment. This sample size is similar to other experiments reporting illusory weight perception that also measured force deployment (e.g., Baugh, Kao, Johansson, & Flanagan, 2012; Buckingham et al., 2011b, experiment 1). Handedness was measured with the 10-item Edinburgh Handedness Inventory (Oldfield, 1971). Participants were classed as right-handed if they reported using their right hand for at least 7 of the 10 tasks (e.g., writing, throwing, using a toothbrush) on the inventory. Participants had normal or corrected-to-normal vision and gave written informed consent to participate. These criteria apply to all experiments reported in this manuscript. All study procedures were approved by the La Trobe University Human Research Ethics Committee.

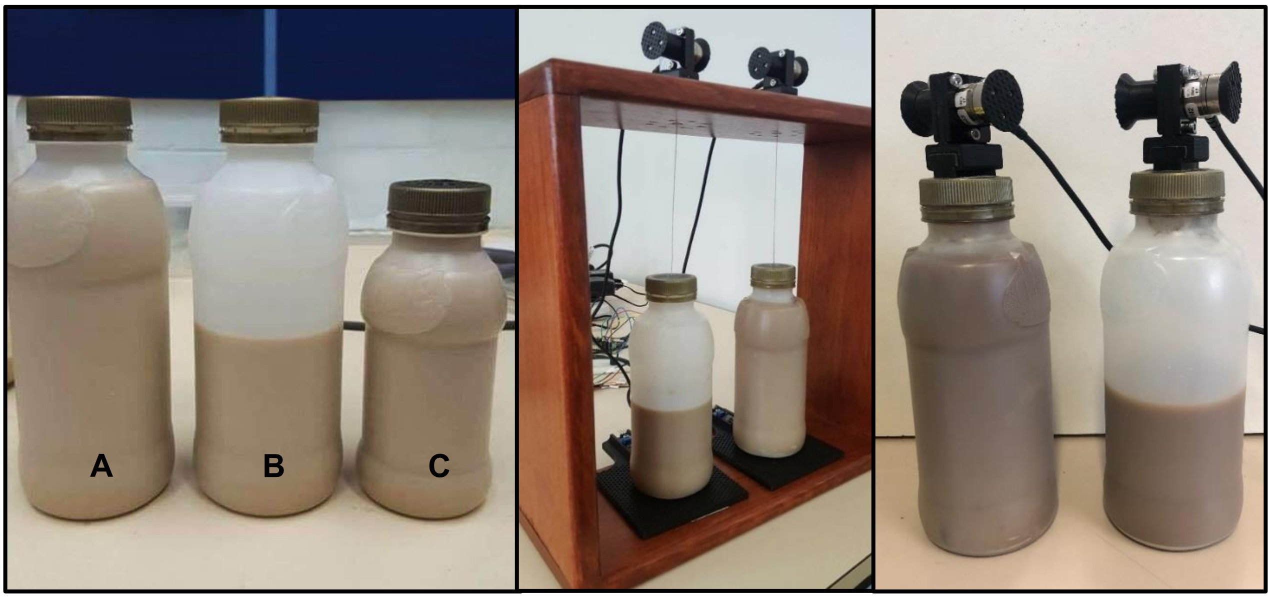

Figure 1. Left panel: the experimental stimuli. All bottles weighed 409 g and contained fake coffee-flavoured milk. Experiment 1a used bottles A (full bottle) and B (half-full bottle). Experiment 1b used bottles A (large bottle) and C (small bottle). Middle panel:

the experimental apparatus from Experiment 1a. This example image depicts bottles A and B. Bottles rested on light sensor pads, which recorded when bottles were raised from the platform. Bottles were attached via strings to the force transducers, which rested on the top of the wooden frame. Right panel: Bottles A and B with transducer handles attached to the lids, which were employed in Experiment 2.

2.1.1.2 Stimuli. Stimuli were two plastic bottles with 500ml capacity (height: 178 mm; minimum diameter (bottle neck): 385 mm; maximum diameter: 720 mm). One bottle was full of opaque liquid that mimicked coffee-flavoured milk and the other approximately half-full of the same liquid (see Figure 1). The stimuli simulated a brand of cold coffee-flavoured milk that is familiar to Australians. The liquid in the bottles was made from PVA glue diluted with water and coloured with coffee so that it mimicked this drink in both colour and viscosity. Both bottles weighed 409 g. The half-full bottle contained a hidden tube of lead ballast to increase its weight, whereas an empty tube within the full bottle displaced enough liquid to reduce its weight to 409 g. Care was taken to ensure the weights were secured centrally inside the bottles and attached to the bottom to prevent an uneven distribution of mass, which can influence perceived heaviness (Amazeen & Turvey, 1996; Plaisier & Smeets, 2015).

2.1.1.3 Apparatus. The main experimental apparatus consisted of a wooden frame (340 mm high, 340 mm wide), which rested on the table at which participants sat during the experiment. The bottles were placed inside the wooden frame (see Figure 1). The bottles' lids were attached to non-stretch Dacron braided fishing line, which had been fed through holes in the top panel of the wooden frame. Strings were 110 mm in length. Connected to the other end of the strings were 3D printed handles equipped with six-axis force transducers (Nano17 F/T; ATI Industrial Automation, Garner, North Carolina, USA), which participants used to lift the bottles. The transducers recorded grip (horizontal) and load (vertical) forces during the lift for the index finger only (thus, both the grip and load forces reported in this paper are presumed to be half of what was actually applied during lifting). The wooden frame acted as a support for the transducers, holding them level before the lift began. The bottles rested on light sensor pads, which recorded when the bottles were raised from the platform.

The experiment was controlled via a Dell Precision T1700 computer using a program designed in-house in MATLAB (The Mathworks, Inc., 2016, Natick, MA, USA). The program played two beeps to signal when the participant should lift and lower the bottles, and recorded the data from the force transducers and light sensors.

2.1.1.4 Procedure. Participants sat at a table in front of the wooden frame so that they had a clear view of the stimuli throughout the experiment. Prior to the lifting task, participants were asked to estimate how heavy each object would be. Weight estimates were made using an absolute magnitude estimation procedure, whereby participants provided heaviness estimates using any scale they preferred, without specifying an anchor or lower and upper limits (Buckingham & Goodale, 2013; Buckingham et al., 2011b; Flanagan, Bittner, & Johansson, 2008). It was made clear that a larger score corresponded to a heavier weight.

Participants began the lifting trials after providing preliminary heaviness estimates. The experimenter began each trial by sounding a beep, which signalled participants to lift the stimulus. Participants remained seated and lifted each bottle with their right hand, gripping the force transducer with their index finger and thumb. They were asked to look at the bottle and to raise it approximately 5 cm from the platform, and to lift in a smooth, confident manner. They maintained their view of the bottle as they held it aloft for three seconds, after which a second beep would sound, prompting them to place the bottle down. After each lift, the participants were asked to report a weight judgement, using the same scale they used for the preliminary weight estimate. Participants were explicitly instructed to concentrate on their experience of the objects and report their perception for each lift. They were told that there was no right or wrong answer, and that they could keep their response the same throughout the experiment or change it, as long as it reflected their experience.

The lifting procedure was repeated 20 times per bottle in alternation, for a total of 40 lifts. The starting bottle was counterbalanced across participants, as well as the side of the frame (right or left) in which the bottles appeared. After the experiment had finished, participants were thanked for their time, debriefed appropriately, and given a compensatory gift voucher.

2.1.2 Experiment 1b: Large and small (SWI) bottles lifted with strings.

2.1.2.1 Participants. Fourteen different participants (10 females, 4 males; age: *M* = 28.8 years, *SD* = 13.74 years) were recruited for Experiment 1b.

2.1.2.2 Apparatus, stimuli and procedure. Apparatus, stimuli and procedure were identical to Experiment 1a with the following exceptions. Participants lifted one large bottle of 500 ml capacity (identical to Experiment 1a) and one small bottle of 250 ml capacity (height: 144 mm; minimum diameter (bottle neck): 385 mm; maximum diameter: 627 mm; see bottle C in Figure 1). Both bottles appeared full of the same opaque, coffee-coloured liquid. The large bottle was the same object employed in Experiment 1a as the "full bottle". The small bottle's weight was adjusted in the same manner as the half-full bottle in Experiment 1a, in that it contained an identical, hidden tube of lead ballast and weighed 409 g. Note that the small bottle in this experiment had approximately the same apparent amount of liquid as the half-full bottle. Owing to the difference in size for the two bottles in Experiment 1b, the string attached to the small bottle was 140 mm.

2.1.3 Data analysis. The dependent measures (in *italics*) comprised perceptual ratings of heaviness as well as force and load phase data. *Preliminary heaviness estimates* and post-lift *perceptual heaviness ratings* comprised of magnitude estimates, which participants provided on an unconstrained numerical scale of their choosing. It is plausible that participants may alter how they scale their estimates after lifting objects for the first time compared to before. Thus, for the purposes of analyses, the *preliminary heaviness estimates* were not standardised whereas the post-lift perceptual ratings were standardised into *Z* scores, allowing for more meaningful comparisons across participants who used different scales. The *Z* scores were calculated by taking the participant's overall mean score, subtracting it from each rating observation, then dividing this difference by the standard deviation of the mean. Force data were recorded at a sample rate of 400 Hz along the X, Y, and Z axes. The force data were smoothed using a fourth-order, zero-phase lag, low-pass Butterworth filter, with a cut-off frequency of 14 Hz prior to analysis (Buckingham, Goodale, et al., 2016; Chouinard et al., 2009; Flanagan & Beltzner, 2000). These data yielded measures of grip force, defined as the force applied normal to the surface of the transducer handle (i.e., force applied horizontally), and load force, defined as the resultant vector of force tangential to the handle surface (i.e., force applied vertically).

For each trial, the grip and load force signals, in Newtons (N), were plotted over time (seconds). The first peak in these signals after lift-off was selected by a rater (R.M.G.) as the *maximum grip and load forces* for that trial (Figure 2A). This method of identifying peak values is common in weight illusion studies that analyse force variables (Baugh et al., 2012; Buckingham et al., 2009; Buckingham et al., 2011b; Flanagan & Beltzner, 2000; Grandy & Westwood, 2006). To establish the rate at which these forces were applied, each force signal was differentiated using a three-point central difference equation to determine, at each time point, the rate of force applied in N/s (Chouinard et al., 2009; Flanagan & Beltzner, 2000). Here, peak values accounting for the rise in force at lift-off were taken as the *maximum grip and load force rates* (Figure 2B). Once again, the peak values were identified on a trial-by-trial basis by the same rater. In establishing these measurements, the rater selected on the plots where the peaks occurred and our program selected the true maximum point nearby, reducing error in our measures. *Load phase duration* was taken as the time between the moment the participant began applying a load force greater than 0.2 N to the point of object lift-off, as measured by the light sensor. Unlike the other measures, *load phase durations* were determined in a completely automated manner. A second rater (P.A.C.) independently processed (i.e., selected peaks in the force data) for five participants chosen at random. The inter-rater reliability between the two raters was excellent (peak grip force: $r(198) = 0.99$, $p < .001$; peak load force $r(198) = 0.98$, $p < .001$; peak rate in grip force: r(198) = 0.93, $p < .001$; peak rate in load force: $r(198) = 0.99$, $p < .001$; load phase: $r(198) = 1$, $p < .001$).

In the manuscript, we present analyses that were conducted on the following dependent variables. To measure the effects of liquid content and container size on expected and perceived weight perception, we analysed the unstandardised, pre-lift weight estimates and the mean standardised post-lift perceptual heaviness ratings. To measure the effects of liquid content and container size on sensorimotor prediction and adaptation, we present the analysis of two of the force variables: mean peak grip and load force rates (Buckingham, Bieńkiewicz, et al., 2015; Buckingham & Goodale, 2010c; Buckingham, Goodale, et al., 2016). Arguably, these variables are more likely than the other force variables to reflect sensorimotor prediction because the peaks in the rates at which forces are applied occur during the earlier stages of the lift, and are therefore less influenced by sensorimotor feedback acquired during the lift (Johansson & Westling, 1988). However, the analyses for the other force variables are presented in the Supplementary Material. Paired sample *t*-tests compared the pre-lift estimates across stimuli as unstandardised values. To echo what was said earlier, this was done in case participants used a different scale after lifting the objects for the first time, which would confound all perceptual measures if they were all standardised. All other variables were analysed with a 2 (Bottle; full/large, half-full/small) x 5 (Trial: 1, 5, 10, 15, 20) repeated measures ANOVA. Pairwise comparisons are presented with a family-wise Bonferroni correction applied. Greenhouse-Geisser corrections were applied whenever sphericity could not be assumed as determined by a Mauchly's test. For a small number of trials, force data were missing owing to participant lifting errors (e.g., fumbling, lifting before the beep indicated the start of the trial). For the purpose of analysis, those data points were replaced with the mean of the immediately preceding and subsequent trials. Data analysis was identical for both experiments.

In the interest of transparent statistical reporting, particularly as this is the first study to measure force and load phase variables when strings are used to lift the weight illusion stimuli, descriptive statistics from *all* force and load phase variables, and analyses on the variables not presented in the manuscript are reported in the Supplementary Material. The data are also publicly available at https://osf.io/n9aeu/.

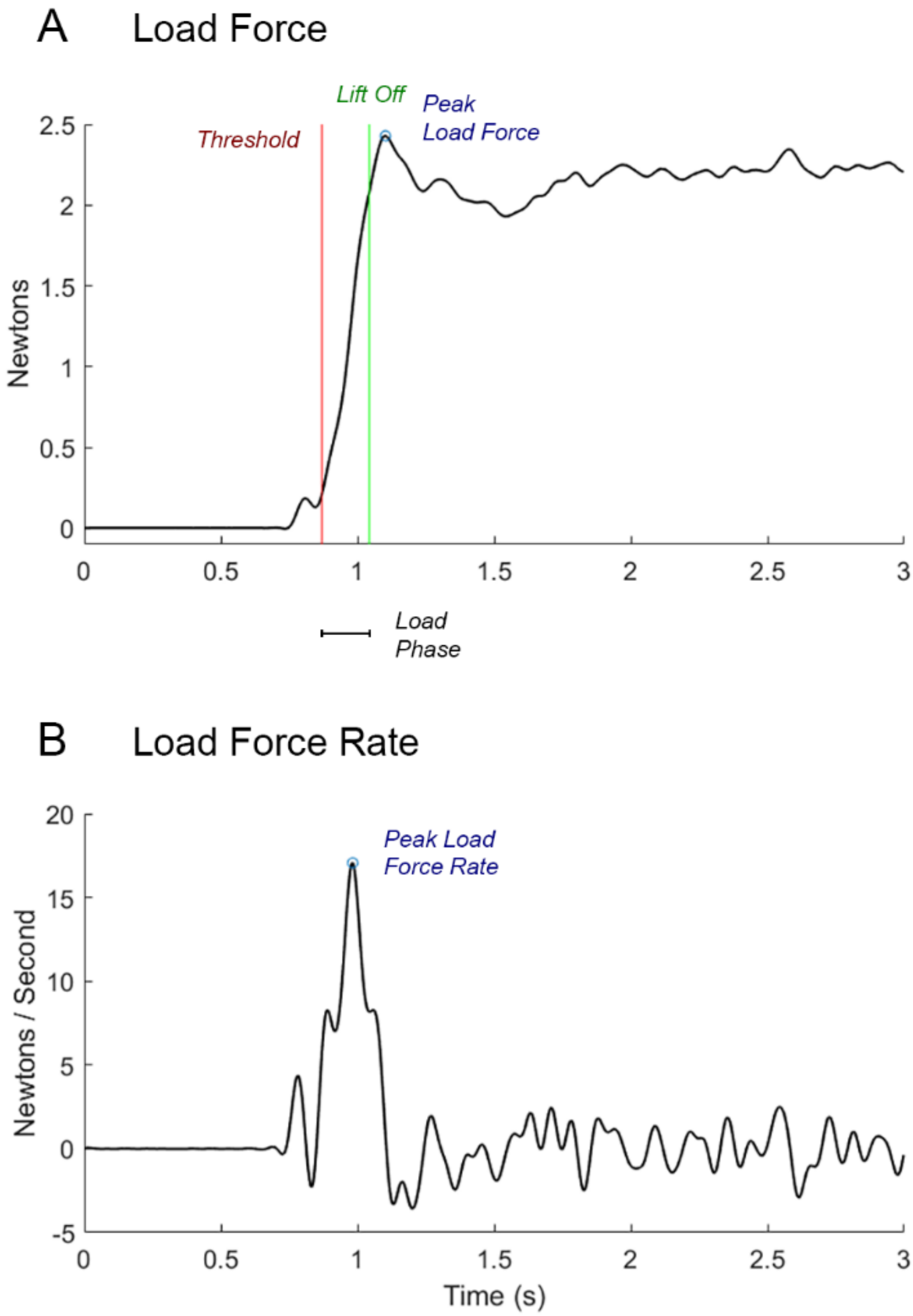

Figure 2. Illustrative example of the force variables. This figure shows the load force (A) and load force rate (B) in a representative participant on a given trial. The blue circles within the figure demonstrate how the peak values of grip and load force and their rates were selected on a trial-by-trial basis. In our experiments, the force transducers were calibrated to measure forces in Newtons. Recordings were taken from the index finger only. A light sensor recorded the time when lift off occurred (green line). The load phase was defined as the time in milliseconds from when the load force first reached 0.2 N (red line) to the time with lift off occurred. The load force signal was differentiated to

determine the load force rate at each time point in Newtons per second. For every trial, the peak force and peak force rates (blue circles) were determined using a semi-automated procedure whereby the experimenter (R.M.G.) selected on the plots where the peak signals occurred and computer program selected the true maximum point nearby.

2.2 Results

2.2.1 Experiment 1a: Full and half-full bottles lifted with strings. In summary, contrary to what we had hypothesised, participants reported that the full bottle felt heavier than the half-full one. Furthermore, the force data did not reflect our hypothesis that more force would be applied to the full compared to the half-full bottle at the start of the experiment and that the forces applied to the two objects would become more similar as the experiment progressed.

2.2.1.1 Preliminary heaviness estimates prior to lifting. Unstandardised magnitude estimates made prior to lifting indicated a trend towards participants expecting the full bottle (M = 84.43, SD = 153.34) to be heavier than the half-full bottle (M = 44.18, SD = 78.79), $t(13)$ = 2.01, p = .065, Cohen's d = 0.33. Not quite reaching significance in this case could be explained by noise arising from the use of different scales across participants (e.g., scores ranged from 3-500).

2.2.1.2 Perceptual heaviness ratings during lifting. Mean standardised perceptual ratings for the full and half-full bottles are displayed in Figure 3A. The 2 x 5 repeated measures ANOVA revealed a main effect of Bottle, $F(1, 13)$ = 9.56, p = .009, η_p^2 = .42. Contrary to our hypothesis, the full bottle was perceived as heavier than the half-full bottle. There was no main effect of Trial, $F(2, 24)$ = 2.24, p = .132, η_p^2 = .15, indicating no evidence of significant change in the perceptual ratings across the experiment. The interaction between these two factors was also not significant, $F(2, 25)$ = 0.13, p = .868, η_p^2 = .01, indicating that this illusory weight difference between the bottles was perceived similarly throughout the experiment.

2.2.1.3 Peak grip force rate. Mean peak grip force rates for the full and half-full bottles are displayed in Figure 3B. The 2 x 5 repeated measures ANOVA revealed neither a main effect of Bottle, $F(1, 13) = 0.19$, $p = .670$, $\eta_p^2 = .01$, nor Trial, $F(4, 52) = 0.71$, $p = .591$, $\eta_p^2 = .05$. The Bottle x Trial interaction was also not significant, $F(4, 52) = 1.57$, $p = .197$, $\eta_p^2 = .11$.

2.2.1.4 Peak load force rate. Mean peak load force rates for the full and half-full bottles are displayed in Figure 3C. The 2 x 5 repeated measures ANOVA revealed no main effect of Bottle, $F(1, 13) = 0.29$, $p = .598$, $\eta_p^2 = .02$. There was a significant main effect of Trial, $F(4, 52) = 4.60$, $p = .003$, $\eta_p^2 = .26$. Pairwise comparisons indicated significantly greater mean peak load forces in trial 15 compared to trial 1 ($p = .006$). None of the other comparisons were significant (all ps > .144). There was no significant interaction, $F(4, 52) = 1.47$, $p = .223$, $\eta_p^2 = .10$.

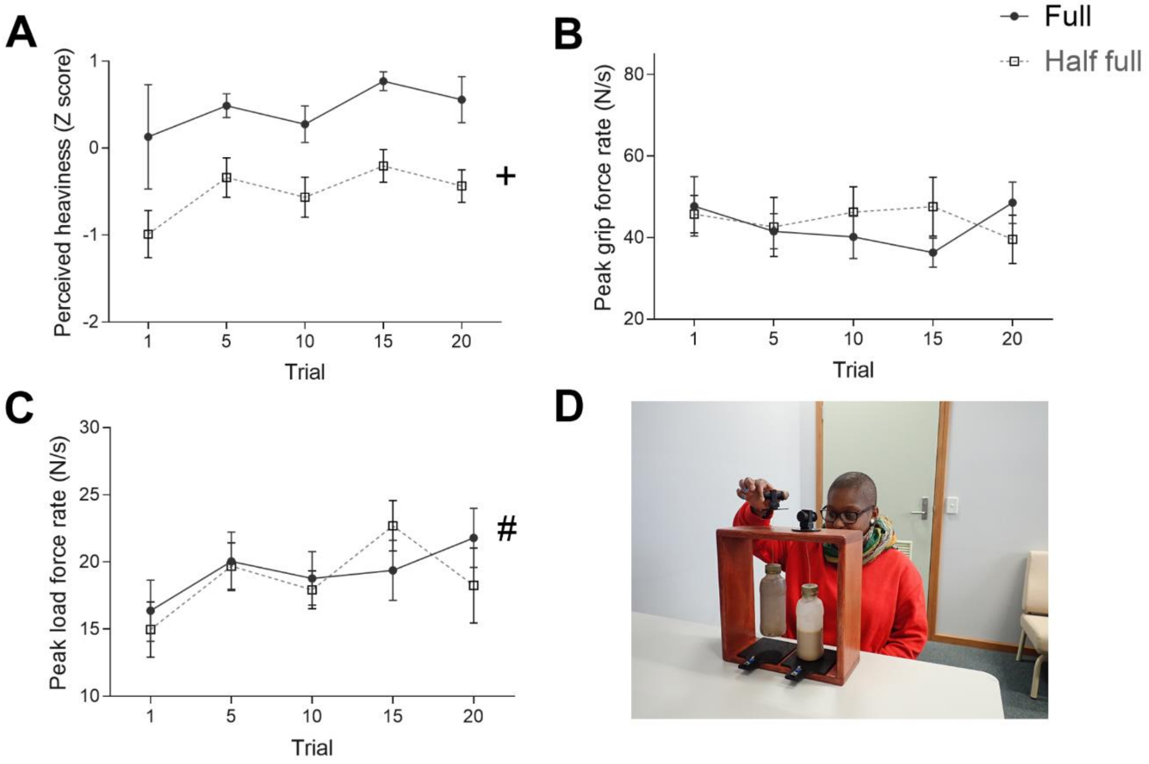

Figure 3. Mean standardised perceptual ratings (A), peak grip force rates (B) and peak load force rates (C) for the full (Bottle A in Figure 1) and half-full (Bottle B in Figure 1) bottles and image depicting a lifting trial (D) in Experiment 1a. The lifter reported a clear view of the stimuli at all times. Error bars denote standard errors around the mean. Cross (+) denotes a significant (p <.05) main effect of Bottle. Pound (#) denotes a significant (p <.05) main effect of Trial.

2.2.2 Experiment 1b: Large and small (SWI) bottles lifted with strings. In summary, perceptual ratings did not reflect a SWI. In contrast to a typical SWI, participants reported that the large bottle felt heavier than the small one, but only for the first trial. For the remainder of the experiment, participants did not perceive a difference in weight between the bottles. The force data also did not reflect our hypotheses. Namely, the forces deployed in initial trials were not greater for the larger compared to the smaller bottle, nor did they become more similar as the experiment progressed.

2.2.2.1 Preliminary heaviness estimates prior to lifting. Unstandardised magnitude estimates made prior to lifting indicated that participants expected the large bottle (M = 411.86, SD = 226.45) to be heavier than the small bottle (M = 216.71, SD = 126.54), $t(13) = 6.32$, $p < .001$, Cohen's d = 1.06.

2.2.2.2 Perceptual heaviness ratings during lifting. Mean standardised perceptual ratings for the large and small bottles are displayed in Figure 4A. There was neither a main effect of Bottle, $F(1, 13) = 0.97$, $p = .343$, $\eta_p^2 = .07$, nor Trial, $F(1, 19) = 3.72$, $p = .056$, $\eta_p^2 = .22$. However, there was a significant Bottle x Trial interaction, $F(2, 27) = 8.59$, $p = .001$, $\eta_p^2 = .40$. Contrary to most past SWI experiments, and our hypothesis, pairwise comparisons revealed significantly heavier perceptual estimates for the large bottle than the small bottle in trial 1 ($p = .006$). There were no differences in perceived heaviness across the bottles for the other trials (all ps > .284).

2.2.2.3 Peak grip force rate. Mean peak grip force rates for the large and small bottles across all trials are displayed in Figure 4B. There was neither a significant main effect of Bottle, $F(1, 13) = 4.00$, $p = .067$, $\eta_p^2 = .24$, nor Trial, $F(2, 24) = 2.46$, $p = .109$, $\eta_p^2 = .16$. The interaction was also not significant, $F(2, 23) = 2.31$, $p = .128$, $\eta_p^2 = .15$.

2.2.2.4 Peak load force rate. Mean peak load force rates for the large and small bottles across all trials are displayed in Figure 4C. There was neither a significant main effect of Bottle, $F(1, 13) = 0.17$, $p = .688$, $\eta_p^2 = .01$, nor Trial, $F(2, 24) = 0.58$, $p = .551$, $\eta_p^2 = .04$, and no significant interaction between the two factors, $F(2, 32) = 1.52$, $p = .233$, $\eta_p^2 = .10$.

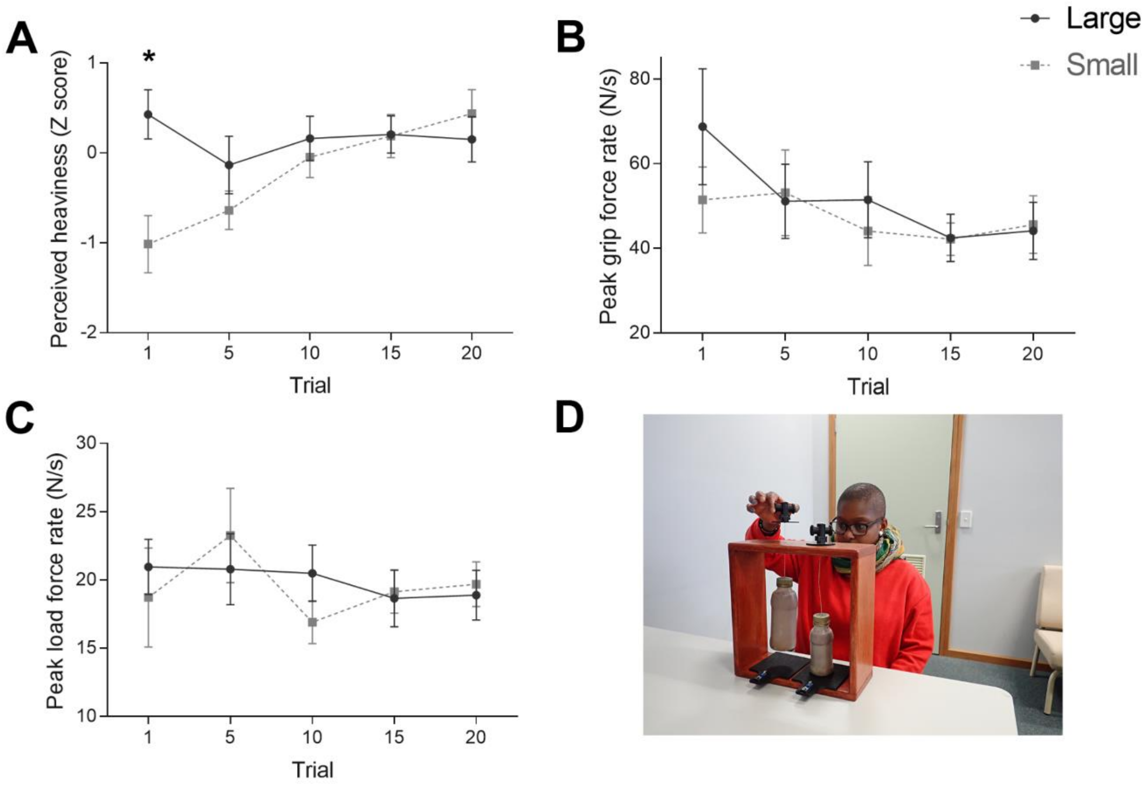

Figure 4. Mean standardised perceptual ratings (A), peak grip force rates (B) and peak load force rates (C) for the large (Bottle A in Figure 1) and small (Bottle C in Figure 1) bottles and image depicting a lifting trial (D) in Experiment 1b. The lifter reported a clear view of the stimuli at all times. Error bars denote standard errors around the mean. Asterisk (*) denotes a significant (*p* <.05, Bonferroni family-wise corrected) pairwise comparison between the bottles within a trial. There were no significant (*p* <.05) main effects of Bottle or Trial.

2.3 Discussion

Results from Experiment 1 did not support our hypotheses. Instead, participants in Experiment 1a reported the *opposite* perceptual pattern to that predicted by weight illusions in that they perceived the full bottle as heavier than the half-full bottle. However, this same unusual, reverse weight illusion pattern was also evident in the first trial of the SWI experiment (Experiment 1b), in that participants initially reported the large bottle as heavier than the small one. One potential explanation for these findings is that the participants may have deduced (incorrectly) that the experimenter expected them to state that the full (or large) bottle felt heavier than the half-full (or small) bottle, in line with what one would expect under natural conditions. Within the experimental context, participants may have assumed a trick regarding the weight of the bottles and as such simply reported what they thought the experimenter wanted to hear or what they expected to be true. Interestingly, although participants reported this unusual pattern with the SWI bottles in the first trial of Experiment 1b, there appears to be a trend for a change in perceptual reports over the course of the experiment. Perhaps participants in Experiment 1b initially felt the same pressure to respond in a certain way as in Experiment 1a, but that the SWI, which is typically a remarkably strong effect (Saccone et al., 2019), began to override this tendency. Regardless, it is clear from Experiment 1a that liquid volume did not influence perceived weight in the manner predicted by other weight illusions.

The force data in Experiment 1 did not demonstrate results similar to those in other studies where participants misapplied forces in line with predicted weight differences on initial trials but then learned to apply more veridical forces after repeated lifts (Buckingham & Goodale, 2010b; Buckingham & Goodale, 2010c; Buckingham et al., 2011b; Buckingham et al., 2012; Flanagan & Beltzner, 2000; Grandy & Westwood, 2006). One issue to consider is that there are inconsistencies across these past studies with respect to which particular force variables are presented as evidence for sensorimotor prediction (i.e., differences in forces between objects) during initial trials and sensorimotor adaptation (i.e., the convergence of forces between objects) in subsequent trials. We will discuss this matter further in the General Discussion. Regardless, Experiment 1 did not produce what is generally considered to be the typical pattern for any of the force variables in a weight illusion experiment. Instead, the force data suggest a cautious approach to lifting the stimuli. For example, in Experiment 1b there was a pattern of overestimation in grip force rates for early compared to later trials but this did not differ statistically across the two bottles. The deployment of these forces may relate to the use of strings. It is possible that this unusual, artificial lifting style influenced participants' strategy for applying forces compared to previous studies in which transducer handles were attached directly to objects.

Overall, contrary to our hypotheses, neither stimulus pair produced the expected perceptual weight illusions in Experiment 1. These findings suggest that the conceptual cue of liquid volume does not elicit a typical illusory weight experience. However, given that the paradigm also did not produce findings that are typical of the (extremely robust) SWI when container size varied, it is difficult to draw definite conclusions about how the conceptual cue of liquid volume might influence predicted and perceived weight when vision is the only modality with which size information is provided.

In contrast to the current experiment, past research using string-based lifts has demonstrated that visual information alone can elicit a SWI (Anderson, 1970; Buckingham, Milne, et al., 2015; Ellis & Lederman, 1993; Kawai, Henigman, MacKenzie, Kuang, & Faust, 2007; Masin & Crestoni, 1988; Werber & King, 1962; Wolf et al., 2018). However, in light of the current experiment's failure to elicit a SWI, it may be that some additional somatosensory information about the stimuli is particularly important for the illusion. In this vein, evidence described earlier suggests that the illusion is stronger when test objects are held directly in the hands, and hence the lift includes haptic and kinaesthetic feedback about stimulus size (Ellis & Lederman, 1993; Plaisier & Smeets, 2015; Wolf et al., 2018). Furthermore, as mentioned earlier, SWI studies that measure forces do so by having participants lift stimuli via an attached transducer handle. A recent meta-analysis by Saccone et al. (2019) demonstrated a weaker SWI when objects are lifted a) with strings than when they are lifted b) by a handle or hefted directly in the hands. Therefore, it seems that there is something about the somatosensory information received about the objects when they are handled - either directly or via an attached handle - that produces a stronger SWI than when they are lifted with a string.

Although there is no direct tactile feedback of size during a handle-based lift, there might be more kinaesthetic feedback about size from the distribution of mass that is gained from the torques applied when lifting an object via a handle compared with a string. This is because the latter ensures that the stimuli are lifted vertically whereas the former may deviate from true vertical lifting even in the most coordinated participants. Any subtle rotational forces applied where the string is gripped is unlikely to give the lifter any additional information about the object, which is dangling on the other end of the string. In contrast, during a handle-based lift, the resistance of the object to rotational forces could still be detected kinaesthetically, which may provide subtle enough cues to some of its properties, such as size or distribution of mass. If kinaesthetic feedback is in fact important in driving the SWI, then this could indicate that size does not simply represent a conceptual cue to predicted and perceived weight. This possibility was investigated in Experiment 2.

3. Experiment 2

Experiment 2 was identical to Experiment 1 except that participants lifted stimuli via a force transducer that was attached directly to the lids of the bottles (see Figure 1, right panel). If object features influence perceived weight via a conceptual cue, one that relies on semantic associations, then perception should not vary according to which sensory modality provides information about that feature. Therefore, if kinaesthetic information about differently sized objects produces a stronger SWI than visual information, then this would suggest that size informs predicted and perceived weight via a different mechanism than conceptual cues such as liquid volume.

With respect to the perceptual reports of weight, Experiment 2 had two competing hypotheses. First, if both liquid volume and container size influence perceived weight via a conceptual weight cue, then Experiment 2 was predicted to produce perceptual weight illusions for both stimulus pairs. That is, the half-full bottle may be perceived as heavier than the full bottle, and likewise the small bottle may be perceived as heavier than the large bottle. However, if size influences weight perception via additional mechanisms that are more sensory driven, in which kinaesthetic information exerts a particularly strong effect, then Experiment 2 should produce a stronger weight illusion when container size differs (i.e., large vs small).

It is possible that the unusual lifting style in Experiment 1 produced the unexpected findings in the force data. Experiment 2 employed a lifting style that is commonly used in weight illusion studies that measure fingertip forces (e.g., Buckingham, Goodale, et al., 2016; Buckingham et al., 2011b; Buckingham et al., 2012; Flanagan & Beltzner, 2000; Grandy & Westwood, 2006). Therefore, Experiment 2 was deemed more likely to produce force data that reflects sensorimotor prediction in initial trials and sensorimotor adaptation in subsequent trials.

3.1 Method

3.1.1 Experiment 2a: Full and half-full bottles lifted with handles.

3.1.1.1 Participants. Fourteen right-handers (11 females, 3 males; age: *M* = 21.79 years, *SD* = 2.99 years) from the La Trobe University community participated in the experiment.

3.1.1.2 Apparatus, stimuli and procedure. Apparatus, stimuli and procedures were identical to Experiment 1a except that participants lifted the bottles off of a table via force transducers attached directly to the bottle lids. Accordingly, the wooden frame and strings were not used.

3.1.2 Experiment 2b: Large and small (SWI) bottles lifted with handles.

3.1.2.1 Participants. Fourteen different right-handers (11 females, 3 males; age: *M* = 29.21 years, *SD* = 12.90 years) participated.

3.1.2.2 Apparatus, stimuli and procedure. Apparatus and procedure were identical to Experiment 2a. Stimuli were identical to those employed in Experiment 1b.

3.1.3 Data analysis. Data analysis procedures were identical to Experiment 1.

3.2 Results

3.2.1 Experiment 2a: Full and half-full bottles lifted with handles. In Experiment 2a, participants did not perceive any difference in weight between the full and half-full bottles. Peak load force rates were greater for the full bottle than the half-full bottle. However, there was no evidence suggesting that these force differences dissipated as the experiment progressed.

3.2.1.1 Preliminary heaviness estimates prior to lifting. Unstandardised magnitude estimates made prior to lifting indicated that participants expected the full bottle (*M* = 275.29, *SD* = 363.69) to be heavier than the half-full bottle (*M* = 137.71, *SD* = 181.79), $t(13) = 2.83$, $p = .014$, Cohen's $d = 0.48$.

3.2.1.2 Perceptual heaviness ratings during lifting. Mean standardised perceptual ratings for the full and half-full bottles are displayed in Figure 5A. There was neither a main effect of Bottle, $F(1, 13) = 0.13$, $p = .722$, $\eta_p^2 = .00$, nor Trial, $F(4, 52) = 0.04$,

p = .987, η_p^2 = .00. The interaction between the two factors was also not significant, F(2, 30) = 2.36, p = .105, η_p^2 = .15.

3.2.1.3 Peak grip force rate. Mean peak grip force rates for the full and half-full bottles are displayed in Figure 5B. There was neither a significant main effect of Bottle, F(1, 13) = 0.35, p = .565, η_p^2 = .03, nor Trial, F(2, 29) = 1.63, p = .213, η_p^2 = .11, and the interaction between the two factors was also not significant, F(2, 24) = 0.64, p = .524, η_p^2 = .05.

3.2.1.4 Peak load force rate. Mean peak load force rates for the full and half-full bottles are displayed in Figure 5C. There was a significant main effect of Bottle, F(1, 13) = 6.56, p = .024, η_p^2 = .34, reflecting higher peak load force rates for the full bottle than the half-full bottle. There was no significant main effect of Trial, F(3, 34) = 0.92, p = .428, η_p^2 = .07, and the interaction between the two factors was also not significant, F(4, 52) = 1.02, p = .408, η_p^2 = .07.

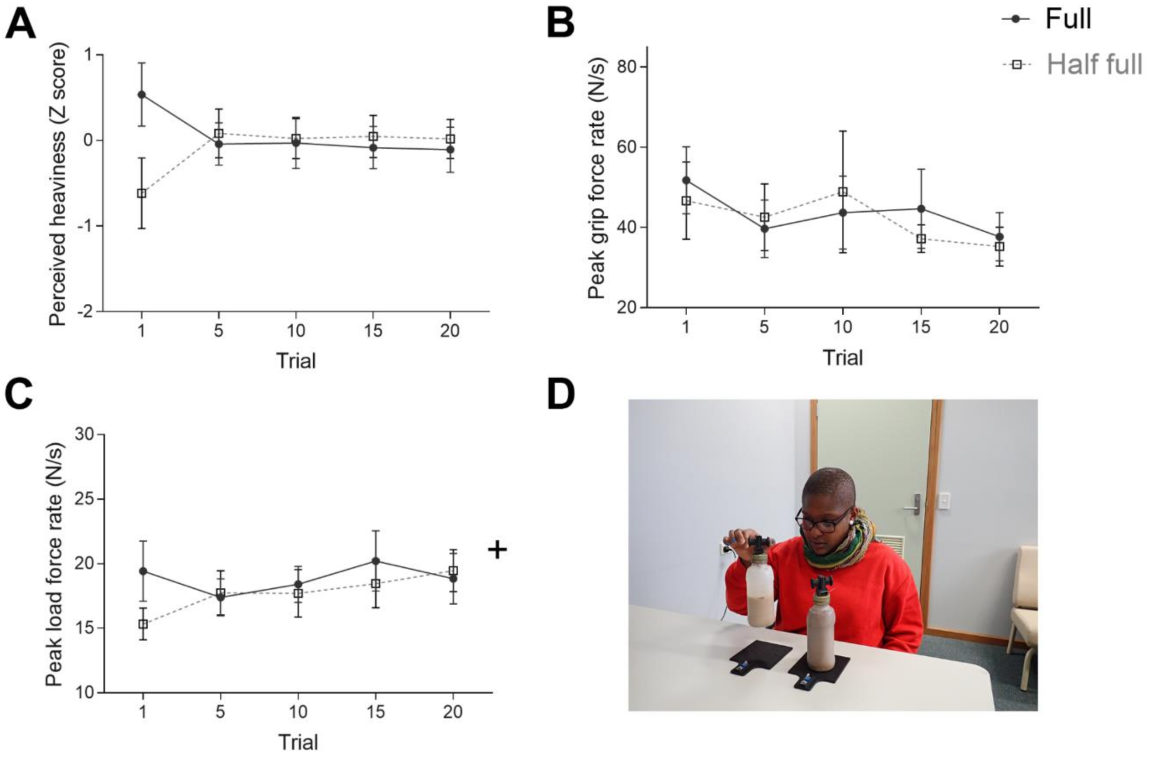

Figure 5. Mean standardised perceptual ratings (A), peak grip force rates (B) and peak load force rates (C) for the full (Bottle A in Figure 1) and half-full (Bottle B in Figure 1) bottles and image depicting a lifting trial (D) in Experiment 2a. Error bars denote standard errors around the mean. Cross (+) denotes a significant (p <.05) main effect of Bottle.

3.2.2 Experiment 2b: Large and small (SWI) bottles lifted with handles. In summary, participants initially reported the larger bottle as heavier, which is consistent with Experiment 1b. However, as the experiment progressed, the pattern reversed and participants reported a SWI from trial 10 onwards. Peak grip force rates demonstrated faster rate of force deployed for the large than small bottle. There was no evidence suggesting that this difference dissipated over time.

3.2.2.1 Preliminary heaviness estimates prior to lifting. Unstandardised magnitude estimates made prior to lifting indicated that participants expected the large bottle (*M* = 221. 17, *SD* = 235.30) to be heavier than the small bottle (*M* = 124.39, *SD* = 139.47), *t*(13) = 3.51, *p* = .004, Cohen's *d* = 0.50.

3.2.2.2 Perceptual heaviness ratings during lifting. Mean standardised perceptual ratings for the large and small bottles are displayed in Figure 6A. There was no significant main effect of Bottle, *F*(1, 13) = 2.39, *p* = .146, η_p^2 = .16. There was a significant main effect of Trial, *F*(2, 24) = 4.63, *p* = .022, η_p^2 = .26, although none of the pairwise comparisons survived the Bonferroni correction (all *p*s > .083). There was a strong, significant interaction between the two factors, *F*(2, 22) = 16.90, *p* < .001, η_p^2 = .57. Pairwise comparisons revealed that for trial 1, the large bottle was perceived as heavier than the small bottle (*p* = .032), which is consistent with Experiment 1b. There was no difference in perceived heaviness across the bottles for trial 5 (*p* = .090), however, the small bottle was perceived as heavier in trials 10 (*p* = .006), 15 (*p* = .003) and 20 (*p* = .007). Thus, aside from in the early trials, participants perceived a robust SWI.

3.2.2.3 Peak grip force rate. Mean peak grip force rates for the large and small bottles are displayed in Figure 6B. The ANOVA revealed a significant main effect of Bottle, *F*(1, 13) = 16.75, *p* = .001, η_p^2 = .56, reflecting higher grip force rates for the large bottle. There was also a main effect of Trial, *F*(2, 25) = 3.98, *p* = .034, η_p^2 = .23, although none of the pairwise comparisons survived the Bonferroni correction (all *p*s > .252). There was no interaction between the two factors, *F*(3, 33) = 0.84, *p* = .466, η_p^2 = .06.

3.2.2.4 Peak load force rate. Mean peak grip force rates for the large and small bottles are displayed in Figure 6C. There was neither a main effect of Bottle, *F*(1, 13) = 1.19, *p* = .296, η_p^2 = .08, nor Trial, *F*(4, 52) = 0.32, *p* = .862, η_p^2 = .02, and no significant interaction, *F*(2, 28) = 0.33, *p* = .739, η_p^2 = .03.

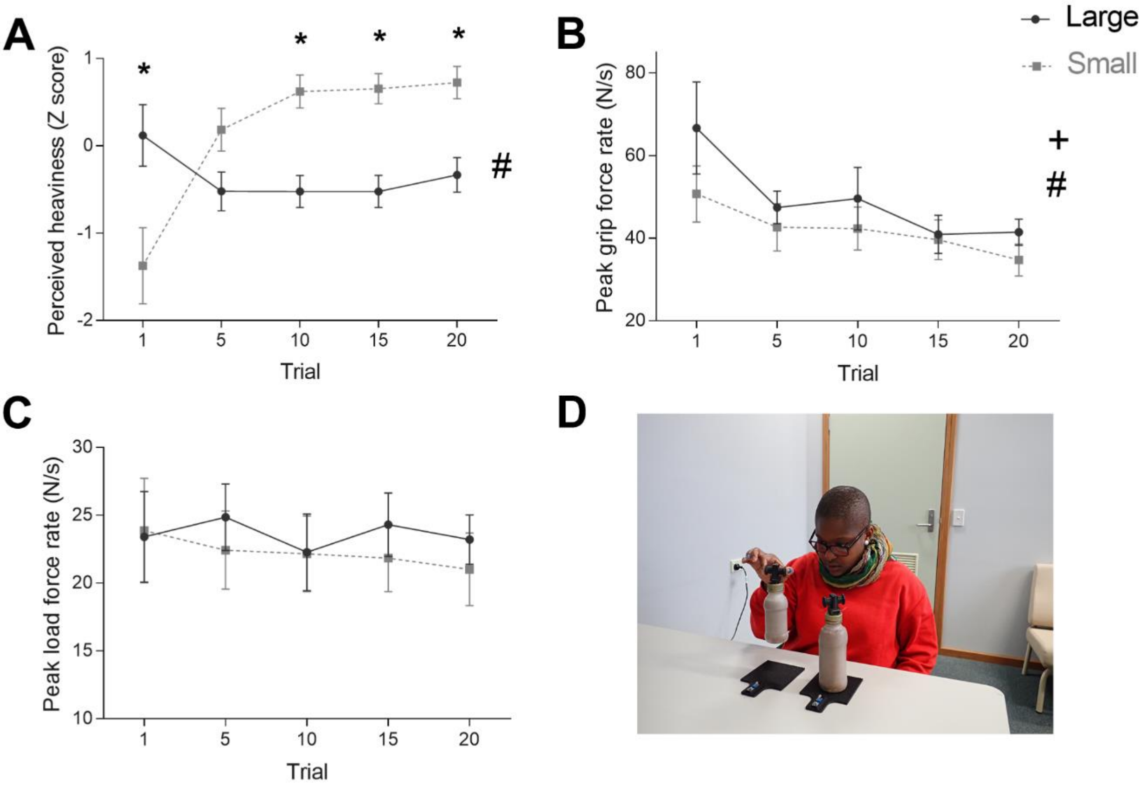

Figure 6. Mean standardised perceptual ratings (A), peak grip force rates (B) and peak load force rates (C) for the large (Bottle A in Figure 1) and small (Bottle C in Figure 1) bottles and image depicting a lifting trial (D) in Experiment 2b. Error bars denote standard errors around the mean. Cross (+) denotes a significant (*p* <.05) main effect of Bottle. Pound (#) denotes a significant (*p* <.05) main effect of Trial. Asterisk (*) denotes a significant (*p* <.05, Bonferroni family-wise corrected) pairwise comparison between the bottles within a trial.

3.3 Discussion

Participants lifting objects with strings in Experiment 1 did not report a SWI, whereas participants lifting with handles directly mounted on the stimuli in Experiment 2 experienced a SWI. Together, these findings suggest that some additional somatosensory processing of the differently sized objects is important for the illusion. Of note, participants initially reported the reverse pattern that the larger bottle was heavier, which is consistent with the early trials of Experiment 1a. Perhaps participants initially felt the same pressure to give a particular response – one consistent with what one would expect according to typical size or liquid volume cues to weight – but that the power of the illusion quickly prevailed. Thus, the paradigm in Experiment 2 elicited a SWI – yet participants did not report a weight difference between the full and half-full bottles. This finding suggests that when container size is held constant, liquid volume does not produce an illusory weight experience in the manner of other conceptual weight cues, such as material or identity. Taken together, Experiment 2 indicates that size may influence perceived weight via a different mechanism to other conceptual cues that elicit illusory weight experiences.

Force data were relatively more typical in Experiment 2. Container size influenced the degree of force applied when lifting the bottles in Experiment 2b. Namely, higher grip force rates were deployed for the large than small bottle. Liquid volume exerted some influence on the force applied in Experiment 2a, in that peak load force rates were higher for the full bottle compared with the half-full bottle. These findings provide some evidence of the sensorimotor predictions that the full bottle would be heavier than the half-full bottle, and likewise for the large bottle compared with the small one. In other respects, the force data did not show the expected pattern, particularly regarding sensorimotor adaptation in later trials that is considered typical for weight illusion paradigms. That is, comparatively higher force rates evident for the full (Experiment 2a) and large bottles (Experiment 2b) did not diminish over the course of the experiment.

The findings from Experiments 1 and 2 demonstrate that lifting objects by an attached handle is more likely to produce a SWI than a strings-based lift. However, it is unknown which particular information is received by the somatosensory system during this type of lift that accounts for a stronger or more reliable SWI. Although there is no haptic feedback about the dimensions of the object during a handle-based lift, because the object is not gripped directly in the hand, we have argued that during this type of lift there could be some information about size that can be detected via kinaesthetic feedback, perhaps via the torques applied during the lift. However, we note that there is no empirical evidence demonstrating that this is the case. This is surprising, given that so many contemporary SWI experiments have employed this lifting technique (Buckingham, Bieńkiewicz, et al., 2015; Buckingham & Goodale, 2010a, 2010c, 2013; Buckingham, Goodale, et al., 2016; Buckingham et al., 2011b; Buckingham et al., 2012; Flanagan & Beltzner, 2000; Grandy & Westwood, 2006; Mon-Williams & Murray, 2000). Accordingly, we conducted an additional control experiment to determine if size differences can be detected solely through the kinaesthetic information gained during a handle-based lift.

We conducted further control experiments to address two other remaining issues. First, there is a potential alternative explanation for why participants reported a weight difference for the large/small bottles but not the full/half-full pair. Although we engineered the small and half-full bottles to have the same apparent liquid volume content (i.e., 250 ml in each), which ensured that the apparent liquid content was matched across the two stimulus pairs, it is unknown whether or not participants *perceived* comparable volume differences for the two pairs. There is evidence that perceived volume is influenced by container shape and size (e.g., Raghubir & Krishna, 1999), which is relevant because the small and half-full bottles had different dimensions in our study. If participants perceived a greater difference in liquid content for the large/small bottles than for the full/half-full bottles, then this could explain why participants only reported a difference in perception for the former. Thus, perceived volume content of the stimuli was also tested in an additional control experiment.

Second, there is a consideration regarding the paradigm used in the first two experiments. In some weight illusion studies, all stimuli are hidden from view except for the one object that is lifted in that particular trial (e.g., Buckingham & Goodale, 2010c; Buckingham et al., 2011b; Plaisier & Smeets, 2015). In our experiments, the stimulus pair was visible throughout all of the lifting trials. Perhaps having the two stimuli constantly visible influenced participants' interpretation of the experiment. This is particularly relevant in light of the response bias that we suggest explains some curious results in Experiments 1 and 2. On this issue, Vicovaro and Burigana (2014) demonstrated a SWI of comparable strength when participants viewed and lifted stimulus pairs a) simultaneously, with one object in each hand, or b) consecutively, using the same hand to lift one object after the other. Both stimuli in the pair were constantly visible for both lifting methods, though. Accordingly, we aimed to confirm our perceptual findings in a new sample of participants who viewed and lifted one stimulus at a time.

4. Additional Control Experiments

The first control experiment aimed to determine if size differences between objects can be detected solely from the kinaesthetic information obtained during a handle-based lift. Participants' vision was obscured as they lifted the bottle stimuli via a handle. For the sake of completeness, we also had them perform the same task lifting the stimuli via strings. Note that both lifting styles include lifting an object via the same sized handle – either attached directly to the bottle lid or to the other end of a string. However, we will continue to refer to the two lifting styles as handle- or strings-based. Participants were asked to judge the relative size of the objects that were attached to the handle, objects they could neither see nor touch.

Second, we tested the perceived liquid volume content of the two bottle pairs from Experiments 1 and 2. Specifically, we investigated if the perceived difference in liquid content for the large/small pair was greater than for the full/half-full pair. Participants were presented with the two pairs and asked to report perceived volume content of each stimulus based only on visual information.

Third, we aimed to confirm the weight perception findings from Experiments 1 and 2 using a paradigm in which participants viewed and lifted only one stimulus at a time. In the interest of replication, we had participants lift the stimuli using both handles and strings.

4.1 Method

4.1.1 Participants. Sixteen right-handed individuals (8 females, 8 males; age: *M* = 21.56 years, *SD* = 1.27 years) from the La Trobe University community participated in all tasks.

4.1.2 Stimuli. We employed the three bottle stimuli from Experiments 1 (strings attached) and 2 (handles attached). We also used an additional set of stimuli during practice trials for the size perception task. Because we anticipated that reporting magnitude estimates of size for objects they could neither see nor touch would be an unusual and difficult task for participants, we also had them perform this task using a set of familiar objects that varied in size: six nested Russian wooden dolls (see Figure 7A). We did not attach handles to the dolls in order to test whether or not size differences could be discriminated for objects that were lifted directly. The dolls ranged in both height as well as diameter, meaning that participants' grip aperture during the lift varied with the size of the doll (e.g., smallest doll: height = 35 mm, diameter at grip point = approximately 13 mm; largest doll: height = 146 mm, diameter at grip point: approximately 55 mm).

4.1.3 Apparatus. The wooden frame from Experiment 1 was used for the strings-based lifts. Because we were not measuring fingertip forces, we used 3D-printed replicas of the real transducer handles for the handle-based lifts. Additional apparatus for the experiment included vision-obscuring glasses. The lenses of the glasses were filled with black plasticine and blinders were attached to their sides to prevent peripheral visual access. The glasses were worn at all times during the size perception task and also in between trials of the other tasks, so that participants did not observe the experimenter handling the objects. Last, three transparent, plastic food storage containers were employed during the volume perception task. One was a large container (capacity: 1.8 L; approximately 850 mm x 850 mm x 2600 mm) and two were smaller (capacity: 0.9 L; approximately 900 mm x 900 mm x 1200 mm).

4.1.4 Procedure. Participants performed the experiments in the same order (each detailed in subsections below). First, participants performed the size perception task, in which they lifted the nested dolls and the bottles one at a time, without visual access. This task was completed first so that size judgements were made before participants had ever seen or touched the objects. The volume perception task was performed second. Participants viewed but did not lift or touch the bottles, to ensure that volume content judgements were based on visual information only. Participants reported perceived volume in two ways, as detailed below, prior to judging the objects' weight.

4.1.4.1 Size perception without vision. Participants sat at a table wearing the vision-obscuring glasses. They were asked to lift various objects, one at a time, between the thumb and forefinger of their right hand and to provide magnitude estimates of each object's size. Instructions for the magnitude estimates were comparable to Experiments 1 and 2, with higher values denoting a larger size. The experimenter helped to guide the participants' hand to touch the object to begin each lift. Participants first lifted the nested dolls. They grasped the top of the dolls directly.

Participants then lifted each bottle stimulus from Experiments 1 and 2 in the same manner. In this case, participants were told they would be lifting different objects, each attached to the same-sized handle. They were asked to provide a magnitude estimate of size of the object that was attached to the handle, which they could not see. Each bottle was lifted twice with each lifting style (strings and handle) and lifting style was blocked in a counterbalanced order (i.e., participants completed all the strings-based or handle-based trials first). Stimulus presentation order was randomised.

4.1.4.2 Volume perception with vision. For trials in this task, the bottles were presented in the pairs from Experiments 1 and 2 (i.e., full/half-full and large/small). Each trial began with one of the two pairs placed side-by-side on the table in front of the participant. Participants then removed the vision-obscuring glasses and were asked not to touch the bottles. The two small, empty, food storage containers were placed in front of the bottles. Participants were given the large, transparent container, which was full of water, and asked to pour an amount of water into each of the two empty containers that would represent the amount of liquid contained in the two bottle stimuli. Participants were allowed to make adjustments to the amount of water in the two small containers until they were satisfied with the representative quantities. Then, the experimenter weighed each container, recorded the amount of water in each one (in grams, which converts to millilitres of water with parity), and then replaced the water from both small containers back into the large container for the next trial.

As a second measure of perceived volume, participants were also asked to provide a magnitude estimate of the relative volume of liquid in each bottle stimulus, with larger numbers representing greater volumes. These two approaches for measuring perceived volume are known to correlate well (Saccone & Chouinard, 2019a). After the participants provided magnitude estimates, their vision was obscured and the experimenter swapped the left/right positions of the two bottles. The participant performed a second trial for that bottle pair. This procedure was then repeated for the second bottle pair. The order that each pair was presented, as well as the starting left/right position of each bottle, was counterbalanced across participants.

4.1.4.3 Weight perception with vision. Participants viewed and lifted each bottle, one at a time, and provided magnitude estimates of weight as in Experiments 1 and 2. Because we suspected a response bias in the earlier experiments (i.e., participants responded in a manner they thought they should, which may not necessarily reflect what they perceived), we stressed more explicitly to participants in this experiment that we were interested in the weight they were experiencing, regardless of what they were expecting or what they believed they should be experiencing. The experimenter repeated this instruction several times with the aim of reducing the possibility of response bias. Participants lifted each bottle twice using each lifting style, for a total of 12 trials. Again, lifting style was blocked in a counterbalanced order. Stimulus presentation order was randomised.

4.1.5 Data analysis. Magnitude estimates of size, volume, and weight were all transformed into *Z* scores in the same manner as in Experiments 1 and 2. Means of these standardised perceptual ratings of bottle *size* and *weight* were analysed with a 3 (Bottle; full, half-full, small) x 2 (Lifting style; strings, handle) repeated measures ANOVA. Mean standardised ratings of size for the nested dolls were analysed with a one-way repeated measures ANOVA, with Doll as a factor with 6 levels (dolls 1 (smallest) to 6 (largest)). To test the perceived volume difference across the two bottle pairs, perceptual estimates of volume were analysed with a 2 (Pair; full/half-full, large/small) x 2 (Apparent content; 250 ml, 500 ml) repeated measures ANOVA. The same ANOVA model was also performed on the mean amount of water (ml) poured to represent the volume content of each bottle. Pairwise comparisons are presented with a family-wise Bonferroni correction applied.

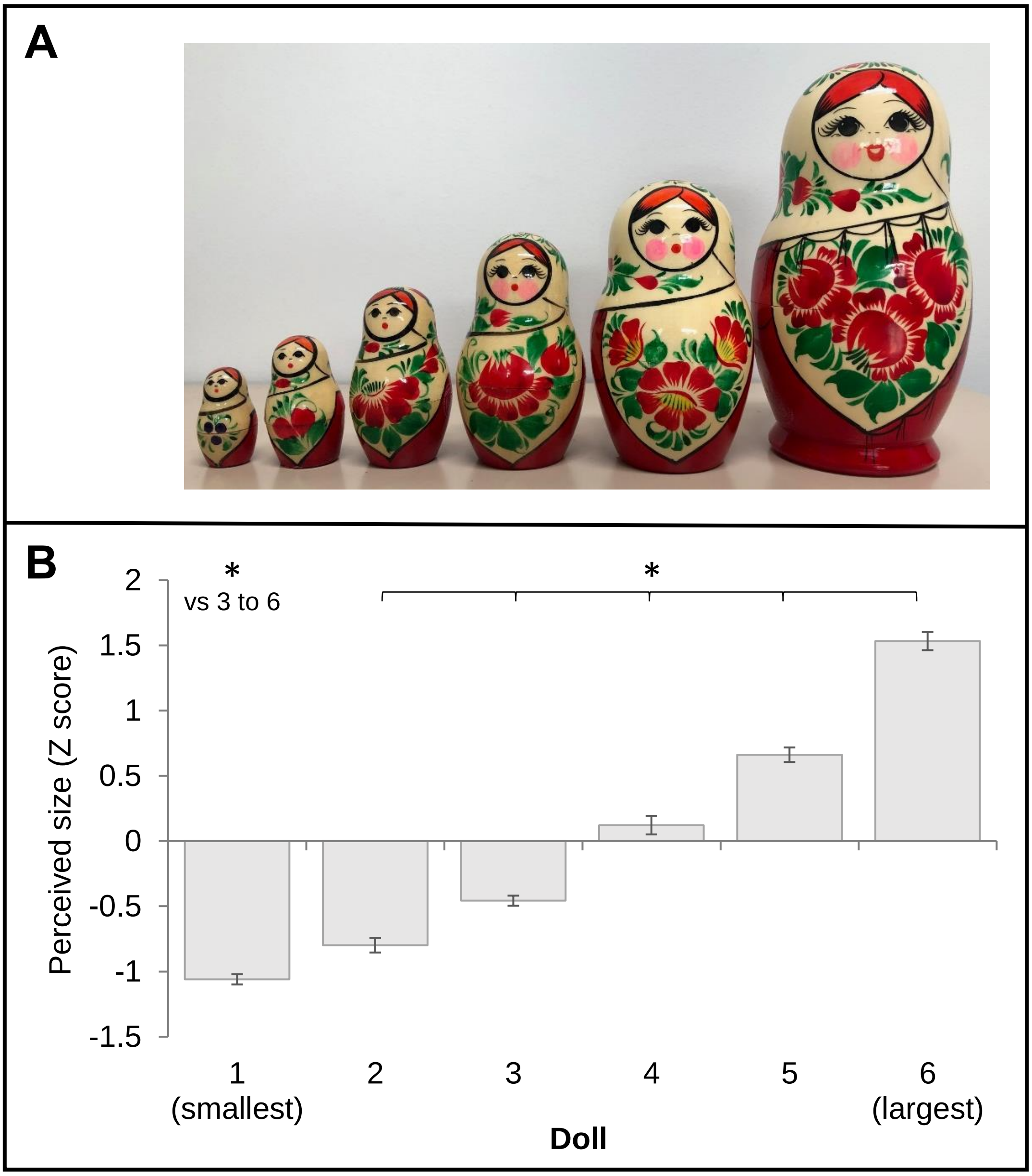

Figure 7. A: Photograph of the six nested Russian dolls used in the size perception task. The dolls ranged from smallest (height = 35 mm, diameter at grip point = approximately 13 mm) to largest (height = 146 mm, diameter at grip point: approximately 55 mm). B: Mean perceptual size estimates for the six dolls. All family-wise Bonferroni-corrected pairwise comparisons were significant (all *ps* < .004), expect for the comparison between dolls 1 and 2 (*p* = .065).

4.2 Results

In summary, the results demonstrated that participants could not detect size differences between the bottles when lifted via an attached handle or string. In contrast, they could discriminate size between the nested dolls. The perceptual ratings suggested a comparable difference in perceived volume across both bottle pairs, whereas the water pouring method revealed a greater difference in perceived volume for the full/half-full pair than the large/small pair. Regarding perceived weight, participants rated the small bottle as heavier than the large bottle when lifting via a handle but not when lifting via a string. There was no perceived weight difference between the full and half-full bottles for either lifting style.

4.2.1 Size perception without vision.

4.2.1.1 Dolls. Mean perceptual size ratings of the dolls are displayed in Figure 7B. There was a main effect of Doll, $F(2.94, 75) = 249.98$, $p < .001$, $\eta_p^2 = .94$. Pairwise comparisons revealed significant differences between all possible pairs (all ps <.004) except for the comparison between the two smallest dolls ($p = .065$). These findings indicate that participants could use somatosensory information to estimate the size of the different dolls. Note that participants used a different grip aperture for each doll.

4.2.1.2 Bottles. Mean perceptual size ratings of the bottles are displayed in Figure 8A. There was neither a main effect of Bottle, $F(2, 28) = 3.05$, $p = .063$, $\eta_p^2 = .18$, nor Lifting style, $F(1, 14) = 0.21$, $p = .651$, $\eta_p^2 = .02$, and no interaction between the two factors, $F(1, 20) = 1.20$, $p = .307$, $\eta_p^2 = .08$. These findings contrast those obtained for the dolls. Somatosensory information did not allow participants to discriminate the differently sized bottles when they lifted the objects with a handle (with or without a string).

4.2.2 Volume perception with vision. Mean amounts of water poured to represent the content of the two bottle pairs are displayed in the upper panel of Figure

8B. There was no main effect of Pair, *F*(1, 15) = 0.25, $p = .623$, $\eta_p^2 = .02$, demonstrating that overall the perceived volume of the full/half-full pair was comparable to the large/small pair. There was a main effect of Apparent content, *F*(1, 15) = 115.44, $p < .001$, $\eta_p^2 = .885$, indicating greater perceived volume reported for the 500 ml bottle (full or large) than the bottles with 250 ml of apparent liquid (small and half-full bottles). There was also an interaction between Pair and Apparent content, *F*(1, 15) = 5.12, $p = .039$, $\eta_p^2 = .255$. Pairwise comparisons indicated significant differences between both the full and half-full bottles ($p <.001$) and the large and small bottles ($p <.001$). Of note, there was no difference between the small and half-full bottles ($p = .74$) or the large and full bottles (which were in fact the same bottle; $p = .245$). Paired sample t-tests explored the interaction further. Results suggest that the interaction is driven by a greater difference for the full/half-full pair, *t*(15) = 11.71, $p <.001$, Cohen's $d = 2.08$, than the large/small pair, *t*(15) = 9.30, $p <.001$, Cohen's $d = 1.71$.

Mean perceptual volume ratings are displayed in the lower panel of Figure 8B. There was no main effect of Pair, *F*(1, 15) = 2.27, $p = .153$, $\eta_p^2 = .13$, but there was a main effect of Apparent content, *F*(1, 15) = 982.66, $p < .001$, $\eta_p^2 = .99$, with higher perceptual estimates for the full (large) bottle than the half-full or small bottles. There was no interaction, *F*(1, 15) = 0.42, $p = .527$, $\eta_p^2 = .03$.

4.2.3 Weight perception with vision. Perceptual heaviness ratings are displayed in Figure 8C. There was a main effect of Bottle, *F*(2, 30) = 9.00, $p = .001$, $\eta_p^2 = .38$. Pairwise comparisons revealed that heaviness ratings were higher for the small bottle compared to the large (full) bottle ($p = .026$) as well as the half-full bottle ($p = .002$). There was no difference between the large (full) and half-full bottles ($p > .999$). There was also a main effect of Lifting style, *F*(1, 15) = 7.66, $p = .014$, $\eta_p^2 = .34$, reflecting higher estimates when bottles were lifted with strings compared to handles. There was a significant interaction between the factors, $F(1, 20) = 5.24$, $p = .024$, $\eta_p^2 = .259$. Pairwise comparisons revealed that when participants lifted bottles via a string, they rated the small bottle as heavier than the half-full bottle ($p = .025$) but there was no difference between the other bottles (ps > .917). When the bottles were lifted via a handle, the small bottle was perceived as heavier than both the large (full) bottle ($p = .001$) and the half-full bottle ($p = .001$), but there was no difference between the large (full) and half-full bottles ($p = .351$).

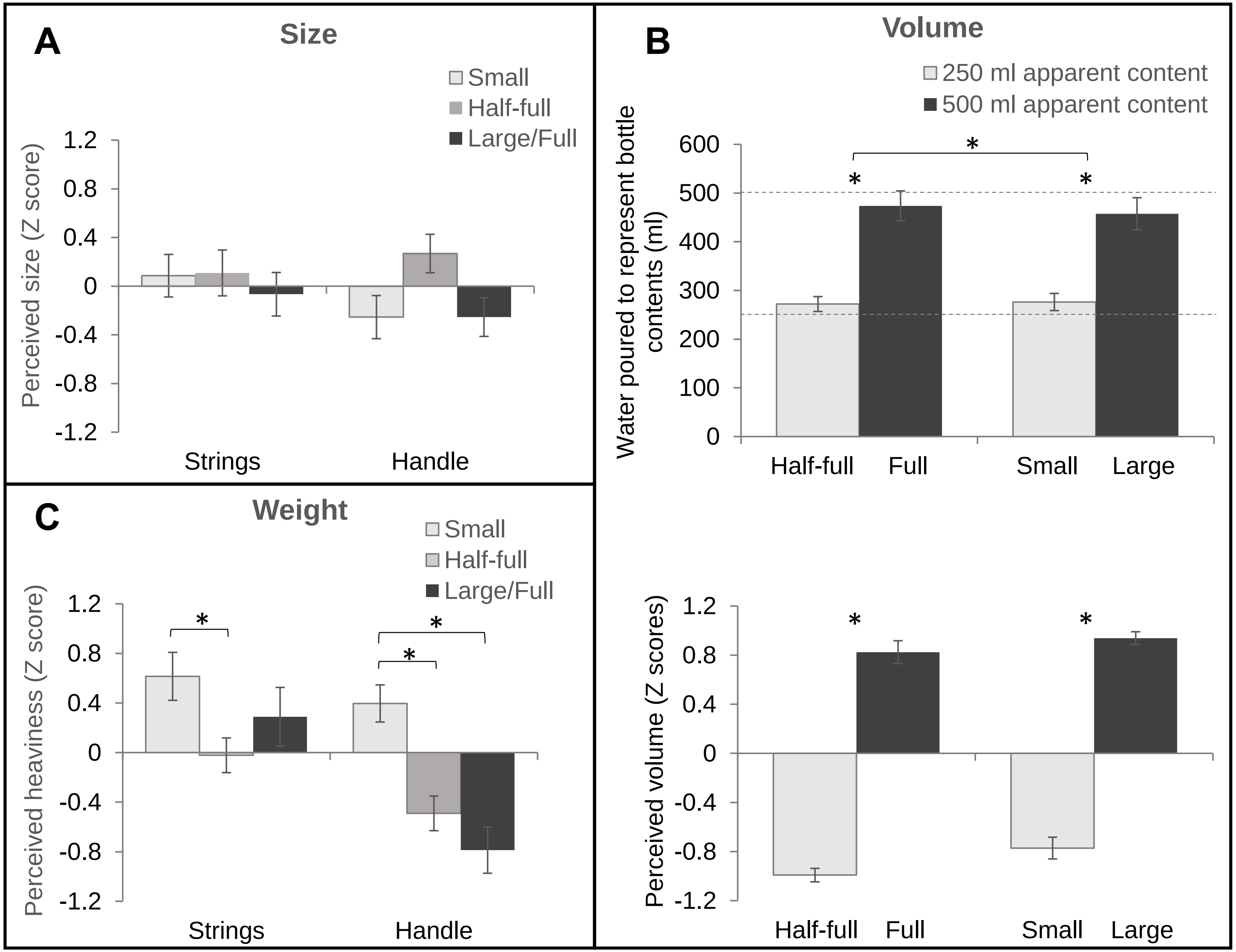
A
Size
Small
Half-full
Large/Full
Perceived size (Z score)
1.2
0.8
0.4
0
-0.4
-0.8
-1.2
Strings
Handle
B
Volume
250 ml apparent content
500 ml apparent content
Water poured to represent bottle contents (ml)
600
500
400
300
200
100
0
*
*
*
Half-full
Full
Small
Large
Perceived volume (Z scores)
1.2
0.8
0.4
0
-0.4
-0.8
-1.2
*
*
Half-full
Full
Small
Large
C
Weight
Small
Half-full
Large/Full
Perceived heaviness (Z score)
1.2
0.8
0.4
0
-0.4
-0.8
-1.2
*
*
*
Strings
Handle

Figure 8. A: Mean perceptual size estimates of the small, half-full and large (full) bottles when lifted via strings and an attached handle. B, upper: Mean amounts of water poured to represent the content of each stimulus for the full/half-full pair and the large/small pair. The dashed lines represent the actual apparent liquid content of the stimuli (i.e., 250 ml (half-full, small) or 500 ml (full, large). B, lower: and mean perceptual volume estimates for the full/half-full pair and the large/small pair. C: Mean perceptual heaviness estimates of the small, half-full and large (full) bottles when lifted via strings and an attached handle. Asterisk (*) denotes a significant ($p < .05$, Bonferroni family-wise correction where appropriate) pairwise comparison between stimuli/conditions.

4.3 Discussion

A number of conclusions can be drawn from these additional experiments. First and foremost, the weight perception task replicated the main findings from Experiments 1 and 2. Namely, the small bottle felt heavier than the large bottle when they were lifted via a handle but not when lifted with strings. There was no perceived weight difference between the full and half-full bottles for either lifting style. We did not replicate the unexpected reverse-weight illusion pattern found in Experiment 1a, which further supports the idea that this earlier finding is more likely to reflect a response bias than a true perceptual effect. Of note, participants in the control experiment also reported the small bottle as heavier than the half-full bottle when participants lifted with the handles. These two bottles were never compared in Experiments 1 and 2. This new information underscores that container size exerts a much stronger influence on perceived weight relative to the apparent amount of liquid content inside. Curiously, this effect was also present during the strings-based lifts. This latter finding should be interpreted with caution as it is unclear as to why lifting the full (large) bottle did not exert similar effects in this control experiment as well as in the Experiments 1 and 2. Future work could look into this matter further.

Interestingly, our findings suggest that these differences in perceived weight across the lifting styles were not due to detected size differences in the stimuli, at least in terms of a reported, conscious percept of size. Contrary to our hypothesis, participants could not discriminate between the differently sized bottles solely from the kinaesthetic information received from a handle-based lift. These findings are particularly interesting in light of the strong evidence our study provides that lifting style influences the strength or reliability of the SWI. Furthermore, although there were no perceived size differences for the bottles, participants could reliably detect differences in the nested dolls. Because participants lifted the dolls directly, rather than via a handle, we can reason that they received haptic information about the dolls' varied sizes, as well as proprioceptive feedback about the varying grip aperture of the thumb and forefinger. Thus, participants were indeed capable of detecting size differences during this task, but not when lifting the objects indirectly via a handle (without or without a string).

Last, these findings rule out the possibility that differences in perceived volume content across the bottle pairs account for differences in perceived weight across the pairs. It was important to investigate if there was a greater difference in perceived volume for the large/small pair. Perceived volume as measured by the water pouring task indicated that there was in fact a greater difference in perceived volume for the full/half-full pair. If anything, this finding suggests that there should be a greater expected weight difference for this pair than the large/small pair, according to the reasoning that predicted weight influences perceived weight. Participants' perceptual ratings of volume indicated instead that differences in volume were comparable across the two pairs. We cannot conclude that there was no difference in perceived volume for the small bottle and the half-full bottle because these two were never directly compared. Regardless, these data demonstrate overall that differences in perceived volume for the two pairs do not explain our findings that container size influenced perceived weight in a paradigm where liquid volume content did not.

5. General Discussion

The present study examined the effects of size and liquid volume content on the perceived weight of bottles and the forces applied when lifting them. This was the first study to examine the familiar weight cue of liquid volume in a weight illusion paradigm. We also varied lifting style to further evaluate the relative contributions of sensory versus conceptual processing on weight perception. In Experiment 1, participants lifted stimuli via strings, which served to isolate the influence of visual information about the stimuli on perception and lifting behaviour. This was the first study to examine force profiles when illusory objects were lifted with strings. Experiment 2 was identical to Experiment 1, except that the objects were lifted via handles that were attached directly to them. This lifting style may, in theory, result in more kinaesthetic information about the size of stimuli than when the stimuli are lifted with strings. This is because a strings-based lift is more likely to ensure that the stimuli are lifted vertically. The results demonstrated that liquid volume does not influence perceived weight in an illusory context when container size is held constant. These findings were replicated in our additional control experiments, which also ruled out of the possibility that differences in perceived liquid content of the bottle pairs explained these differences in perceived weight. Additionally, a control experiment revealed that the extra kinaesthetic information obtained when objects are lifted by an attached handle rather than a string does not translate to a conscious percept of size. These findings from the size perception task add to the novel contribution of our study in examining these two lifting styles, which are commonly used in weight illusion paradigms. Overall, our results provide strong evidence that weight perception in the SWI is more strongly driven by kinaesthetic processing mechanisms than predictions arising from conceptual weight cues.

In the current study, liquid volume did not produce a perceptual weight illusion in the same manner as other conceptual cues in other studies (for a review, see Saccone & Chouinard, 2019b). This finding is surprising given that a) liquid volume is a highly familiar object feature that predicts weight (Nowak & Hermsdörfer, 2003) and b) there is strong evidence that other conceptual weight cues produce an illusory weight percept, such as material or identity (Baugh et al., 2012; Buckingham et al., 2009; Buckingham, Ranger, & Goodale, 2011a; Ellis & Lederman, 1998, 1999; Seashore, 1899; Wolfe, 1898). One possibility is that size cues overrode other information that should predict weight (i.e., liquid volume), in the same manner as other studies that varied size as well as another weight-predicting feature (Buckingham & Goodale, 2013; Buckingham, Goodale, et al., 2016; Plaisier & Smeets, 2015). Note that our findings are consistent with Plaisier and Smeets' (2015) in this respect. However, this reasoning does not hold for other weight illusions, for example, the material-weight illusion. In this case, apparent material produces an illusory weight percept even though stimuli have the same physical size. It could be that conceptual-based effects are smaller and could not be detected in the present investigation. Regardless of why liquid volume did not produce a perceptual experience consistent with other weight illusions, the current findings demonstrate that container size influenced perceived weight in a paradigm where liquid volume content did not. These findings, coupled with evidence from previous literature (Buckingham, Bieńkiewicz, et al., 2015; Buckingham & Goodale, 2013; Buckingham, Goodale, et al., 2016), suggests that size influences perception in a different way to other features that provide a conceptual cue to weight.

The current findings also point to a strong influence of kinaesthetic processing in the SWI. Visual information was not sufficient to produce a SWI in Experiment 1, but the illusion was elicited in Experiment 2 when additional information about the stimuli was gained through kinaesthetic channels via a handle-based lift. This finding contrasts numerous studies that have reported a SWI using strings (Anderson, 1970; Ellis & Lederman, 1993; Kawai et al., 2007; Masin & Crestoni, 1988; Werber & King, 1962; Wolf et al., 2018), including in as few as four participants (i.e., Buckingham, Milne, et al., 2015's sighted control group). We did find partial evidence of a strings-based SWI in the control experiment, in that the small bottle was rated as heavier than the larger, half-full bottle; however, there was no difference in perceived weight between the small and large (full) bottles. The reason the current paradigm did not produce a reliable effect when stimuli were lifted with strings is unknown. Nonetheless, our findings are consistent with existing literature suggesting haptic and/or kinaesthetic information about the stimuli drives the perceptual experience of the illusion more strongly than vision alone (Ellis & Lederman, 1993; Saccone et al., 2019; Wolf et al., 2018). Furthermore, there are documented failures to elicit the SWI when vision is obscured and stimuli are lifted with strings (e.g., experiment 2 in Ellis & Lederman, 1993; experiment 1 in Wolf et al., 2018), whereas when objects are hefted in the hands or lifted via a handle, there is a strong SWI even without visual access (see Saccone et al., 2019's meta-analysis for several examples).

The notion that kinaesthetic feedback exerts a strong influence in the SWI is also supported by the neuropsychological literature. Although the SWI is demonstrably robust in spite of significant cerebellar (Rabe et al., 2009) and cortical damage (Buckingham, Bieńkiewicz, et al., 2015; Li, Randerath, Goldenberg, & Hermsdörfer, 2011; but see Halstead, 1945), one of the few documented cases of an absent SWI is provided by a patient with deafferentation. Buckingham, Michelakakis, and Cole (2016) report the case study of I.W., for whom tactile or proprioceptive feedback was lost after having acquired deafferentation 30 years prior. Accordingly, I.W. does not experience weight or heaviness via somatosensation, but rather is thought to infer mass from visual feedback of his lifting behaviour. Although he can discriminate between real weight differences to a comparable degree to a control group, I.W. does not experience a SWI.

On balance, the current findings suggest that size does not influence perception in the same manner as other conceptual weight cues that rely on acquired, semantic associations. This study also suggests that information obtained through kinaesthetic channels is particularly important in driving the perceptual experience of the SWI. However, an important and novel finding from the current study is that the information provided during this process does not translate to a conscious perception of size. The processing of size can be fulfilled by different neural mechanisms depending on what it is used for (e.g., Goodale, Milner, Jakobson, & Carey, 1991; for a review, see Sperandio & Chouinard, 2015). Thus, it could be the case that processing size for the purposes of perceiving it may depend on a different set of mechanisms than those used to influence weight perception, which relies heavily on somatosensory feedback. The question remains as to precisely why a handle-based lift is more likely to produce a SWI than a strings-based lift. What is the particular variable(s) detected by the somatosensory system during a handle-based lift that influences perceived weight so strongly in the SWI? In considering this question, we must consider alternative accounts of the SWI than those emphasising feature-weight associations. These alternative accounts are often referred to as *bottom-up* accounts (see Buckingham, 2014, and Saccone et al., 2019 for reviews; also see the General Discussion of Plaisier et al., 2019).

The current findings support one bottom-up account of the SWI and weight perception more generally provided by Amazeen and Turvey (1996). They demonstrated that it was the distribution of mass within an object, rather than simply mass, that was a stronger predictor of perceived weight (but see Zhu, Shockley, Riley, Tolston, & Bingham, 2013 for contrasting findings). Amazeen and Turvey provided evidence that kinaesthetic resistance to the rotational forces of the object is translated into a percept of weight. According to their account, rotational inertia is the critical variable accounting for the perceptual experience of the SWI. This explanation fits well with the findings in our study that indicate a driving role of kinaesthetic processing of objects that differ in size but not weight – kinaesthetic processing that does not translate to a conscious percept of size.

Another account that could speak to an important role of kinaesthetic processing is provided by Zhu and Bingham (2011). Rather than proposing that a physical object feature like density influences weight perception, their account implicates an action-relevant feature, drawing from Gibson's (1979) ecological view of object perception. Zhu and Bingham argued that size has been a significant factor for humans in estimating the throwability (and therefore heaviness) of objects since the hunter-gatherer days. They proposed that the effect of size on weight perception relates to a readiness to judge the "throwability" of objects in order to acquire objects that can be thrown the furthest. It is possible that this evolutionary reliance on the physical size of stimuli means that it is prioritised over other weight cues when making heaviness judgements. Our findings could speak to their account in that kinaesthetically-derived information may be more closely related to judging throwability.

There is other evidence in the weight illusion literature that supports these theories underscoring the influence of kinaesthetic, sensory processing rather than conceptual processing in the SWI. Consider that information about size and distribution of mass can be detected kinaesthetically whereas other features such as apparent material or identity cannot. To illustrate, in the material-weight illusion and some other weight illusions (e.g., Ellis & Lederman, 1998), stimuli have the same size, as well as mass, and therefore do not include size differences across stimuli that can be detected kinaesthetically. A number of studies suggest that the SWI is a considerably stronger and/or more reliable illusion than the material-weight illusion (Buckingham, Bieńkiewicz, et al., 2015; Buckingham & Goodale, 2013; Saccone et al., 2019; Vicovaro & Burigana, 2017). Additionally, SWI research demonstrates that size can have a remarkably consistent influence on weight perception, even when other weight-predicting features are varied (Buckingham & Goodale, 2013; Buckingham, Goodale, et al., 2016). Thus, the unique influence of kinaesthetic feedback from objects that differ in size could explain the stronger and more consistent effect of size on illusory weight perception in the SWI compared to other features in other weight illusions.

With respect to our force data, the current study was the first to investigate force deployment in an illusory weight context where liquid content could inform sensorimotor prediction. In general, these data did not show the expected pattern that greater force would be applied for the full bottle during initial lifts. These findings are particularly notable in the case of Experiment 2, which employed the typical lifting technique for a weight illusion paradigm measuring fingertip forces. However, we note that our SWI experiment also produced force profiles that were not entirely consistent with previous studies. These findings speak to an important criticism raised in a review paper by Dijker (2014) on earlier studies recording force data.

Dijker (2014) highlights the degree of inconsistency among published studies in terms of which particular force variables demonstrate sensorimotor adaptation to illusory stimuli and how quickly. He provided an example from Buckingham and Goodale

(2010b). In this SWI study, peak grip force and peak grip/load force rates showed the expected pattern, both in terms of initial sensorimotor predictions and adaptation. That is, greater forces were applied initially for the larger stimuli, and then this difference attenuated over a small number of trials as participants learned the true (i.e., equal) weights of the stimuli. However, peak load force showed poor adaptation over the course of the experiment; there was consistently greater peak load forces applied for the largest than the smallest object. This pattern is consistent with our Experiment 2b, in that peak grip force rates were consistently higher for the large bottle. In fact, we did not find evidence of statistically significant sensorimotor adaptation at all, with respect to statistical interaction effects between the factors of Bottle and Trial. We also note inconsistencies in terms of which particular force variables are reported in published studies. Some present peak *rates* of grip and load forces only (Buckingham, Bieńkiewicz, et al., 2015; Buckingham & Goodale, 2010c; Buckingham, Goodale, et al., 2016), whereas other studies present force rates as well as load phase durations (Buckingham et al., 2011b). Others focus on either grip forces (Buckingham et al., 2012; Flanagan & Beltzner, 2000) or load forces (Baugh, Yak, Johansson, & Flanagan, 2016; Flanagan et al., 2008), with or without load phase durations.

In contrast, Mon-Williams and Murray (2000) reported all five variables as we have done in the current study (see Supplementary Material for analyses not reported in the manuscript). We feel this is more transparent. Choosing which variables to report over others in an inconsistent manner masks rather than eliminates Type 1 error. The risk for making this type of error remains the same regardless of the number of variables one chooses to report. Ultimately, selecting to report only a subset of them is counterproductive in the long run as it does now allow the validity of force data reported across different weight illusion studies to be fully evaluated. Further validity on the recording of forces is warranted. Nonetheless, one should consider that we did not correct for acquiring and reporting multiple force measurements. Thus, one should deem our force results with a certain degree of caution until they are replicated.

Aside from addressing the primary aim of comparing physical size and conceptual weight cues, this study also highlights an additional, important issue for the weight illusion literature. It is possible that demand effects or participants' interpretation of the experimental context can influence results. The present study has demonstrated evidence of a response bias that is unlikely to reflect perception. Namely, the same, opposite pattern to what is typically seen in weight illusions was clearly evident in Experiment 1a, as well as the first trials of Experiments 1b, 2a and 2b. We can only speculate that this pattern relates to participants' incorrect interpretation of the experiment and/or that they felt they should respond in a particular way. This reasoning is supported by the findings from our additional control experiment, in which participants were instructed with great emphasis to report their perceptual experience, regardless of their interpretation of the experiment. We note also that this additional weight perception experiment differed from Experiments 1 and 2 in that stimuli were viewed and lifted one at a time, which might have influenced participants' interpretation and/or responses. The additional control experiment replicated the SWI found with the large and small bottles, whereas there were no differences in perceived weight between the full and half-full bottles. This pattern in the data was unexpected but highlight that these paradigms are subject to influence. It seems the SWI is a strong enough phenomenon to counteract this, whereas other weight illusions may not be.

Acknowledgments

This work was supported by the Australian Research Council (DP170103189). We have no conflicts of interest to declare. We thank Bailey Evans for his help with data collection for the additional control experiments.

References

Amazeen, E. L., & Turvey, M. T. (1996). Weight perception and the haptic size-weight illusion are functions of the inertia tensor. *Journal of Experimental Psychology: Human Perception and Performance, 22*(1), 213-232. doi:10.1037//0096-1523.22.1.213

Anderson, N. H. (1970). Averaging model applied to the size-weight illusion. *Perception & Psychophysics, 8*(1), 1-4. doi:10.3758/BF03208919

Baugh, L. A., Kao, M., Johansson, R. S., & Flanagan, J. R. (2012). Material evidence: Interaction of well-learned priors and sensorimotor memory when lifting objects. *Journal of Neurophysiology, 108*(5), 1262-1269. doi:10.1152/jn.00263.2012

Baugh, L. A., Yak, A., Johansson, R. S., & Flanagan, J. R. (2016). Representing multiple object weights: Competing priors and sensorimotor memories. *Journal of Neurophysiology, 116*(4), 1615-1625. doi:10.1152/jn.00282.2016

Buckingham, G. (2014). Getting a grip on heaviness perception: A review of weight illusions and their probable causes. *Experimental Brain Research, 232*(6), 1623-1629. doi:10.1007/s00221-014-3926-9

Buckingham, G., Bieńkiewicz, M., Rohrbach, N., & Hermsdörfer, J. (2015). The impact of unilateral brain damage on weight perception, sensorimotor anticipation, and fingertip force adaptation. *Vision Research, 115*, 231-237. doi:10.1016/j.visres.2015.02.005

Buckingham, G., Cant, J. S., & Goodale, M. A. (2009). Living in a material world: how visual cues to material properties affect the way that we lift objects and perceive their weight. *Journal of Neurophysiology, 102*(6), 3111-3118. doi:10.1152/jn.00515.2009

Buckingham, G., & Goodale, M. A. (2010a). The influence of competing perceptual and motor priors in the context of the size-weight illusion. *Experimental Brain Research, 205*(2), 283-288. doi:10.1007/s00221-010-2353-9

Buckingham, G., & Goodale, M. A. (2010b). The influence of competing perceptual and motor priors in the context of the size–weight illusion. *Experimental Brain Research, 205*(2), 283-288. doi:10.1007/s00221-010-2353-9

Buckingham, G., & Goodale, M. A. (2010c). Lifting without seeing: The role of vision in perceiving and acting upon the size weight Illusion. *PLoS One, 5*(3). doi:10.1371/journal.pone.0009709

Buckingham, G., & Goodale, M. A. (2013). Size Matters: A Single Representation Underlies Our Perceptions of Heaviness in the Size-Weight Illusion. *PLoS One, 8*(1). doi:10.1371/journal.pone.0054709

Buckingham, G., Goodale, M. A., White, J. A., & Westwood, D. A. (2016). Equal-magnitude size-weight illusions experienced within and between object categories. *Journal of Vision, 16*(3), 1-9. doi:10.1167/16.3.25

Buckingham, G., & MacDonald, A. (2016). The weight of expectation: Implicit, rather than explicit, prior expectations drive the size–weight illusion. *The Quarterly Journal of Experimental Psychology, 69*(9), 1831-1841. doi:10.1080/17470218.2015.1100642

Buckingham, G., Michelakakis, E. E., & Cole, J. (2016). Perceiving and acting upon weight illusions in the absence of somatosensory information. *Journal of Neurophysiology, 115*(4), 1946-1953. doi:10.1152/jn.00318.2015

Buckingham, G., Milne, J. L., Byrne, C. M., & Goodale, M. A. (2015). The size-weight illusion induced through human echolocation. *Psychological Science, 26*(2), 237-242. doi:10.1177/0956797614561267

Buckingham, G., Ranger, N. S., & Goodale, M. A. (2011a). The material–weight illusion induced by expectations alone. *Attention, Perception, & Psychophysics, 73*(1), 36-41. doi:10.3758/s13414-010-0007-4

Buckingham, G., Ranger, N. S., & Goodale, M. A. (2011b). The role of vision in detecting and correcting fingertip force errors during object lifting. *Journal of Vision, 11*(1), 4-4. doi:10.1167/11.1.4

Buckingham, G., Ranger, N. S., & Goodale, M. A. (2012). Handedness, laterality and the size-weight illusion. *Cortex, 48*(10), 1342-1350. doi:10.1016/j.cortex.2011.09.007

Charpentier, A. (1886). Sur les sensations de poids (note présentée par M. D'Arsonval le 3 avril)[On sensations of weight (note presented by Mr. D'Arsonval on the 3rd of April)]. *Comptes Rendus Hebdomadaires des Séances et Mémoires de la Société de Biologie, 38*, 169-170.

Charpentier, A. (1891). Analyse experimentale de quelgues elements de la sensation de poids. *Archives de Physiologie Normale et Pathologique, 3*, 122-135.

Chouinard, P. A., Large, M. E., Chang, E. C., & Goodale, M. A. (2009). Dissociable neural mechanisms for determining the perceived heaviness of objects and the predicted weight of objects during lifting: An fMRI investigation of the size-weight illusion. *NeuroImage, 44*(1), 200-212. doi:10.1016/j.neuroimage.2008.08.023

Chouinard, P. A., Leonard, G., & Paus, T. (2005). Role of the primary motor and dorsal premotor cortices in the anticipation of forces during object lifting. *Journal of Neuroscience, 25*(9), 2277-2284. doi:10.1523/JNEUROSCI.4649-04.2005

Dijker, A. J. M. (2014). The role of expectancies in the size-weight illusion: A review of theoretical and empirical arguments and a new explanation. *Psychonomic Bulletin & Review, 21*(6), 1404-1414. doi:10.3758/s13423-014-0634-1

Ellis, R. R., & Lederman, S. J. (1993). The role of haptic versus visual volume cues in the size-weight illusion. *Perception & Psychophysics, 53*(3), 315-324. doi:10.3758/BF03205186

Ellis, R. R., & Lederman, S. J. (1998). The golf-ball illusion: evidence for top-down processing in weight perception. *Perception, 27*(2), 193-201. doi:10.1068/p270193

Ellis, R. R., & Lederman, S. J. (1999). The material-weight illusion revisited. *Perception & Psychophysics, 61*(8), 1564-1576. doi:10.3758/bf03213118

Flanagan, J. R., & Beltzner, M. A. (2000). Independence of perceptual and sensorimotor predictions in the size- weight illusion. *Nature Neuroscience, 3*(7), 737-741. doi:10.1038/76701

Flanagan, J. R., Bittner, J. P., & Johansson, R. S. (2008). Experience can change distinct size-weight priors engaged in lifting objects and judging their weights. *Current Biology, 18*(22), 1742-1747. doi:10.1016/j.cub.2008.09.042

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Goodale, M. A., Milner, A. D., Jakobson, L., & Carey, D. (1991). A neurological dissociation between perceiving objects and grasping them. *Nature, 349*(6305), 154. doi:10.1038/349154a0

Gordon, A. M., Forssberg, H., Johansson, R. S., & Westling, G. (1991). Visual size cues in the programming of manipulative forces during precision grip. *Experimental Brain Research, 83*(3), 477-482. doi:10.1007/BF00229824

Gordon, A. M., Westling, G., Cole, K. J., & Johansson, R. S. (1993). Memory representations underlying motor commands used during manipulation of common and novel objects. *Journal of Neurophysiology, 69*(6), 1789-1796. doi:10.1152/jn.1993.69.6.1789

Grandy, M. S., & Westwood, D. A. (2006). Opposite perceptual and sensorimotor responses to a size-weight illusion. *Journal of Neurophysiology, 95*(6), 3887-3892. doi:10.1152/jn.00851.2005

Halstead, W. C. (1945). Brain injuries and the higher levels of consciousness. *Research Publications, Association for Research in Nervous & Mental Disease, 24*, 480-506.

Johansson, R. S., & Westling, G. (1984). Roles of glabrous skin receptors and sensorimotor memory in automatic control of precision grip when lifting rougher or more slippery objects. *Experimental Brain Research, 56*(3), 550-564. doi:10.1007/bf00237997

Johansson, R. S., & Westling, G. (1988). Coordinated isometric muscle commands adequately and erroneously programmed for the weight during lifting task with precision grip. *Experimental Brain Research, 71*(1), 59-71. doi:10.1007/bf00247522

Kawai, S., Henigman, F., MacKenzie, C. L., Kuang, A. B., & Faust, P. H. (2007). A reexamination of the size-weight illusion induced by visual size cues. *Experimental Brain Research, 179*(3), 443-456. doi:10.1007/s00221-006-0803-1

Li, Y., Randerath, J., Goldenberg, G., & Hermsdörfer, J. (2011). Size-weight illusion and anticipatory grip force scaling following unilateral cortical brain lesion. *Neuropsychologia, 49*(5), 914-923. doi:10.1016/j.neuropsychologia.2011.02.018

Marinkovic, K., Dhond, R. P., Dale, A. M., Glessner, M., Carr, V., & Halgren, E. (2003). Spatiotemporal dynamics of modality-specific and supramodal word processing. *Neuron, 38*(3), 487-497. doi:10.1016/S0896-6273(03)00197-1

Masin, S. C., & Crestoni, L. (1988). Experimental demonstration of the sensory basis of the size-weight illusion. *Perception & Psychophysics, 44*(4), 309-312. doi:10.3758/BF03210411

Mon-Williams, M., & Murray, A. H. (2000). The size of the visual size cue used for programming manipulative forces during precision grip. *Experimental Brain Research, 135*(3), 405-410. doi:10.1007/s002210000538

Nowak, D. A., & Hermsdörfer, J. (2003). Sensorimotor memory and grip force control: does grip force anticipate a self-produced weight change when drinking with a straw from a cup? *European Journal of Neuroscience, 18*(10), 2883-2892. doi:10.1111/j.1460-9568.2003.03011.x

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97-113. doi:10.1016/0028-3932(71)90067-4

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience, 8*, 976. doi:10.1038/nrn2277

Plaisier, M. A., Kuling, I. A., Brenner, E., & Smeets, J. B. J. (2019). When does one decide how heavy an object feels while picking it up? *Psychological Science, 30*(6), 822-829. doi:10.1177/0956797619837981

Plaisier, M. A., & Smeets, J. B. J. (2015). Object size can influence perceived weight independent of visual estimates of the volume of material. *Scientific Reports, 5*. doi:10.1038/srep17719

Rabe, K., Brandauer, B., Li, Y., Gizewski, E. R., Timmann, D., & Hermsdörfer, J. (2009). Size-weight illusion, anticipation, and adaptation of fingertip forces in patients with cerebellar degeneration. *Journal of Neurophysiology, 101*(2), 569-579. doi:10.1152/jn.91068.2008

Raghubir, P., & Krishna, A. (1999). Vital dimensions in volume perception: can the eye fool the stomach? *Journal of Marketing Research, 36*(3), 313-326. doi:10.1177/002224379903600302

Saccone, E. J., & Chouinard, P. A. (2019a). Barbie-cueing weight perception. *i-Perception*. doi:10.1177/2041669519850590

Saccone, E. J., & Chouinard, P. A. (2019b). The influence of size in weight illusions is unique relative to other object features. *Psychonomic Bulletin & Review, 26*(1), 77-89. doi:10.3758/s13423-018-1519-5

Saccone, E. J., Landry, O., & Chouinard, P. A. (2019). A meta-analysis of the size-weight and material-weight illusions. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-019-01604-x

Seashore, C. E. (1899). Some psychological statistics II. The material weight illusion. *University of Iowa Studies in Psychology, 2*, 36-46.

Sperandio, I., & Chouinard, P. A. (2015). The mechanisms of size constancy. *Multisensory Research, 28*(3-4), 253-283. doi:10.1163/22134808-00002483

Tulving, E. (1972). Episodic and semantic memory. From E. Tulving & W. Donaldson,(Eds.) Organization of memory (pp. 381-403): New York. Academic Press.

Vicovaro, M., & Burigana, L. (2014). Properties of the size-weight illusion as shown by lines of subjective equality. *Acta Psychologica, 149*, 52-59. doi:10.1016/j.actpsy.2014.03.001

Vicovaro, M., & Burigana, L. (2017). Contribution of surface material and size to the expected versus the perceived weight of objects. *Attention, Perception, & Psychophysics, 79*(1), 306-319. doi:10.3758/s13414-016-1212-6

Werber, M., & King, D. J. (1962). An investigation of the response mechanism of the size-weight illusions. *The Journal of General Psychology, 66*, 85-100. doi:10.1080/00221309.1962.9920498

Wolf, C., Bergmann Tiest, W. M., & Drewing, K. (2018). A mass-density model can account for the size-weight illusion. *PLoS One, 13*(2), e0190624. doi:10.1371/journal.pone.0190624

Wolfe, H. K. (1898). Some effects of size on judgments of weight. *Psychological Review, 5*(1), 25-54. doi:10.1037/h0073342

Zhu, Q., & Bingham, G. P. (2011). Human readiness to throw: The size-weight illusion is not an illusion when picking the best objects to throw. *Evolution and Human Behavior, 32*(4), 288-293. doi:10.1016/j.evolhumbehav.2010.11.005

Zhu, Q., Shockley, K., Riley, M. A., Tolston, M. T., & Bingham, G. P. (2013). Felt heaviness is used to perceive the affordance for throwing but rotational inertia does not affect either. *Experimental Brain Research, 224*(2), 221-231. doi:10.1007/s00221-012-3301-7

Supplementary Material

All variables were analysed with a 2 (Bottle; full (or large), half-full (or small) x 5 (Trial: 1, 5, 10, 15, 20) repeated measures ANOVA. Pairwise comparisons are presented with a family-wise Bonferroni correction applied. Greenhouse-Geisser corrections were applied whenever sphericity could not be assumed as determined by a Mauchly's test.

Results

Experiment 1a: Full and half-full bottles lifted with strings.

Peak grip force. Mean peak grip force for the full and half-full bottles across trials are displayed in Figure S1A. The 2 x 5 repeated measures ANOVA revealed no significant main effect of Bottle, $F(1, 13) = 0.33$, $p = .573$, $\eta_p^2 = .03$. There was a significant main effect of Trial, $F(4, 52) = 3.64$, $p = .011$, $\eta_p^2 = .22$. Pairwise comparisons indicated significantly greater mean peak grip forces in trial 1 compared to trial 10 ($p = .045$). None of the other comparisons were significant (all *p*s > .153). The Bottle x Trial interaction was not significant, $F(2, 23) = 0.98$, $p = .382$, $\eta_p^2 = .07$.

Peak load force. Mean peak load force for the full and half-full bottles across all trials are displayed in Figure S1B. The 2 x 5 repeated measures ANOVA revealed a main effect of Bottle, $F(1, 13) = 4.88$, $p = .046$, $\eta_p^2 = .27$, reflecting greater mean peak load forces for the full than half-full bottle. There was no significant main effect of Trial, $F(4, 52) = 1.45$, $p = .232$, $\eta_p^2 = .10$, and no significant interaction, $F(4, 52) = 0.08$, $p = .989$, $\eta_p^2 = .01$.

Load phase duration. Mean load phase duration for the full and half-full bottles are displayed in Figure S1C. There was no main effect of Bottle, $F(1, 13) = 1.17$, $p = .299$, $\eta_p^2 = .08$. The main effect of Trial was significant, $F(2, 28) = 8.21$, $p = .001$, $\eta_p^2 = .39$, reflecting longer mean load phase duration in trial 1 compared to trials 10 ($p = .009$) and

15 (p = .039). There were no differences across the other trials (all ps > .070). The interaction was not significant, F(4, 52) = 2.55, p = .050, η_p^2 = .16.

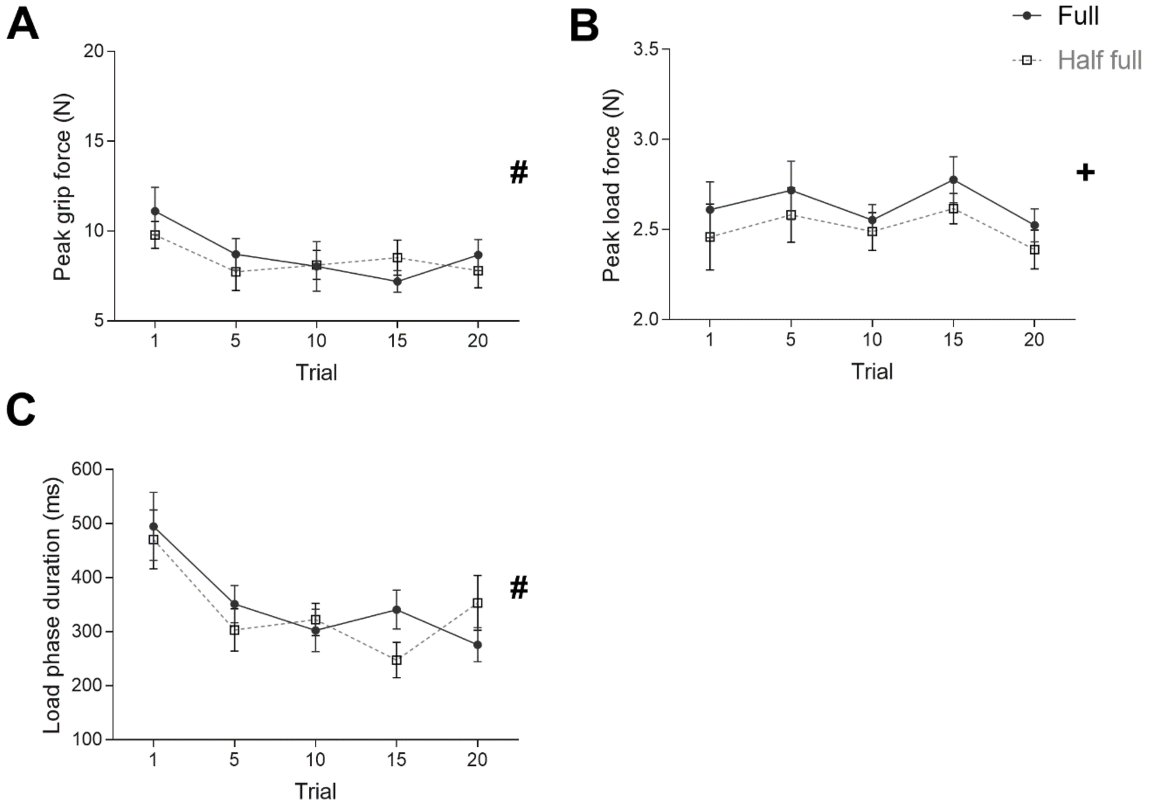

Figure S1. Mean peak grip forces (A), peak load forces (B) and load phase durations (C) for the full and half-full bottles in Experiment 1a. Error bars denote standard errors of the means. Cross (+) denotes a significant (p <.05) main effect of Bottle. Pound (#) denotes a significant (p <.05) main effect of Trial.

Experiment 1b: Large and small (SWI) bottles lifted with strings.

Peak grip force. Mean peak grip force for the large and small bottles across all trials are displayed in Figure S2A. The 2 x 5 repeated measures ANOVA revealed no main effect of Bottle, F(1, 13) = 0.33, p = .575, η_p^2 = .03. There was a significant main effect of

Trial, *F*(3, 34) = 7.27, p = .001, η_p^2 = .36. Pairwise comparisons indicated significantly greater mean peak grip forces in trial 1 compared to trials 15 (p = .003) and 20 (p = .003). None of the other comparisons were significant (all *p*s > .061). The interaction was also not significant, *F*(2, 23) = 0.83, p = .434, η_p^2 = .06.

Peak load force. Mean peak load force for the large and small bottles across all trials are displayed in Figure S2B. There was neither a significant main effect of Bottle, *F*(1, 13) = 3.409, p = .088, η_p^2 = .21, nor Trial, *F*(4, 52) = 1.07, p = .382, η_p^2 = .08. The interaction was also not significant, *F*(2, 26) = 0.75, p = .485, η_p^2 = .05.

Load phase duration. Mean load phase duration for the large and small bottles are displayed in Figure S2C. There was no significant main effect of Bottle, *F*(1, 13) = 0.06, p = .814, η_p^2 = .00. There was a significant main effect of Trial, *F*(4, 52) = 5.35, p = .001, η_p^2 = .29. Pairwise comparisons indicated significantly longer load phase durations in trial 1 compared to trials 10 (p = .036) and 15 (p = .034). None of the other comparisons were significant (all *p*s > .058). The interaction was also significant, *F*(4, 52) = 2.75, p = .038, η_p^2 = .18, however, none of the comparisons survived the Bonferroni correction (all *p*s > .116).

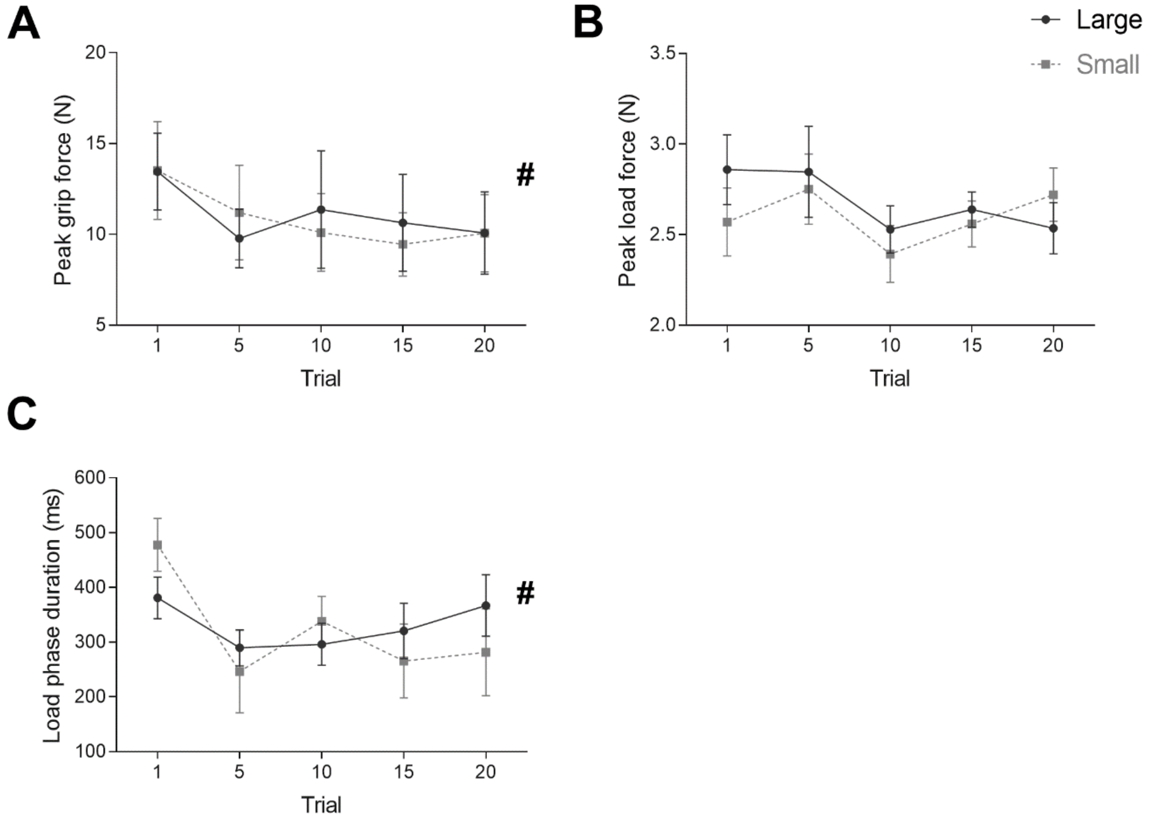

Figure S2. Mean peak grip forces (A), peak load forces (B) and load phase durations (C) for the large and small bottles in Experiment 1b. Error bars denote standard errors of the means. Pound (#) denotes a significant (*p* <.05) main effect of Trial.

Experiment 2a: Full and half-full bottles lifted with handles.

Peak grip force. Mean peak grip force for the full and half-full bottles across all trials are displayed in Figure S3A. There were no significant main effects of Bottle, *F*(1, 13) = 0.10, *p* = .752, η_p^2 = .01, nor Trial, *F*(2, 30) = 3.14, *p* = .052, η_p^2 = .19, and the interaction between the two factors was also not significant, *F*(2, 29) = 0.62, *p* = .564, η_p^2 = .05.

Peak load force. Mean peak load force for the full and half-full bottles across all trials are displayed in Figure S3B. There were no significant main effects of Bottle, *F*(1,

13) = 2.01, $p = .180$, $\eta_p^2 = .13$, or Trial, $F(4, 52) = 0.28$, $p = .893$, $\eta_p^2 = .02$, and the interaction between the two factors was also not significant $F(4, 52) = 0.60$, $p = .666$, $\eta_p^2 = .04$.

Load phase duration. Mean load phase durations for the full and half-full bottles are displayed in Figure S3C. There was no significant main effect of Bottle, $F(1, 13) = 3.48$, $p = .085$, $\eta_p^2 = .21$, but there was a significant main effect of Trial, $F(4, 52) = 6.36$ $p < .001$, $\eta_p^2 = .33$. However, none of the comparisons survived the Bonferroni correction (trials 1 vs 10, $p = .050$, all other $ps > .132$). The Bottle x Trial interaction was not significant, $F(4, 52) = 2.41$, $p = .061$, $\eta_p^2 = .16$.

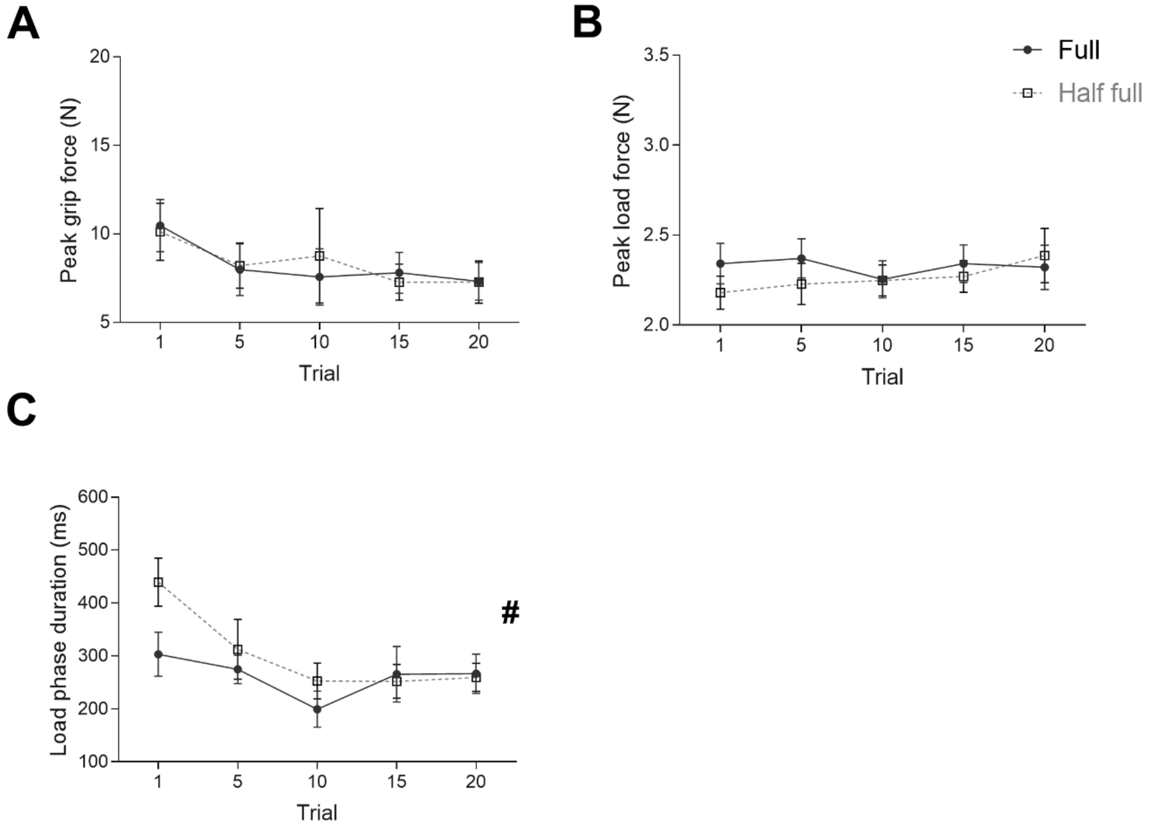

Figure S3. Mean peak grip forces (A), peak load forces (B) and load phase durations (C) for the full and half-full bottles in Experiment 2a. Error bars denote standard errors of the means. Pound (#) denotes a significant ($p < .05$) main effect of Trial.

Experiment 2b: Large and small (SWI) bottles lifted with handles.

Peak grip force. Mean peak grip force values for the large and small bottles are displayed in Figure S4A. There was a strong, significant main effect of Bottle, $F(1, 13) = 21.85$, $p < .001$, $\eta_p^2 = .63$, reflecting higher peak grip forces for the large bottle. There was also a significant main effect of Trial, $F(2, 21) = 12.17$, $p = .001$, $\eta_p^2 = .48$. Pairwise comparisons revealed that peak grip forces were higher for trial 1 compared to trials 5 ($p = .026$), 10, ($p = .017$), 15 ($p = .018$) and 20 ($p = .024$). Peak grip forces were not significantly different across trials 5-20 (all *p*s = 1.00). The interaction between the two factors was not significant, $F(2, 27) = 1.89$, $p = .169$, $\eta_p^2 = .13$.

Peak load force. Mean peak load force values for the large and small bottles across all trials are displayed in Figure S4B. There was neither a main effect of Bottle, $F(1, 13) = 1.41$, $p = .257$, $\eta_p^2 = .10$, nor Trial, $F(2, 28) = 1.45$, $p = .251$, $\eta_p^2 = .10$, and no significant interaction, $F(4, 52) = 0.14$, $p = .966$, $\eta_p^2 = .01$.

Load phase duration. Mean load phase durations for the large and small bottles are displayed in Figure S4C. There was no main effect of Bottle, $F(1, 13) = 3.62$, $p = .079$, $\eta_p^2 = .22$, but there was a significant main effect of Trial, $F(2, 31) = 4.44$, $p = .016$, $\eta_p^2 = .26$. Namely, load phases were longer for trial 1 compared to trial 5 ($p = .010$), whereas there were no significant differences between trials 5-20 (all *p*s > .056). The interaction was not significant, $F(4, 52) = 0.46$, $p = .763$, $\eta_p^2 = .03$.

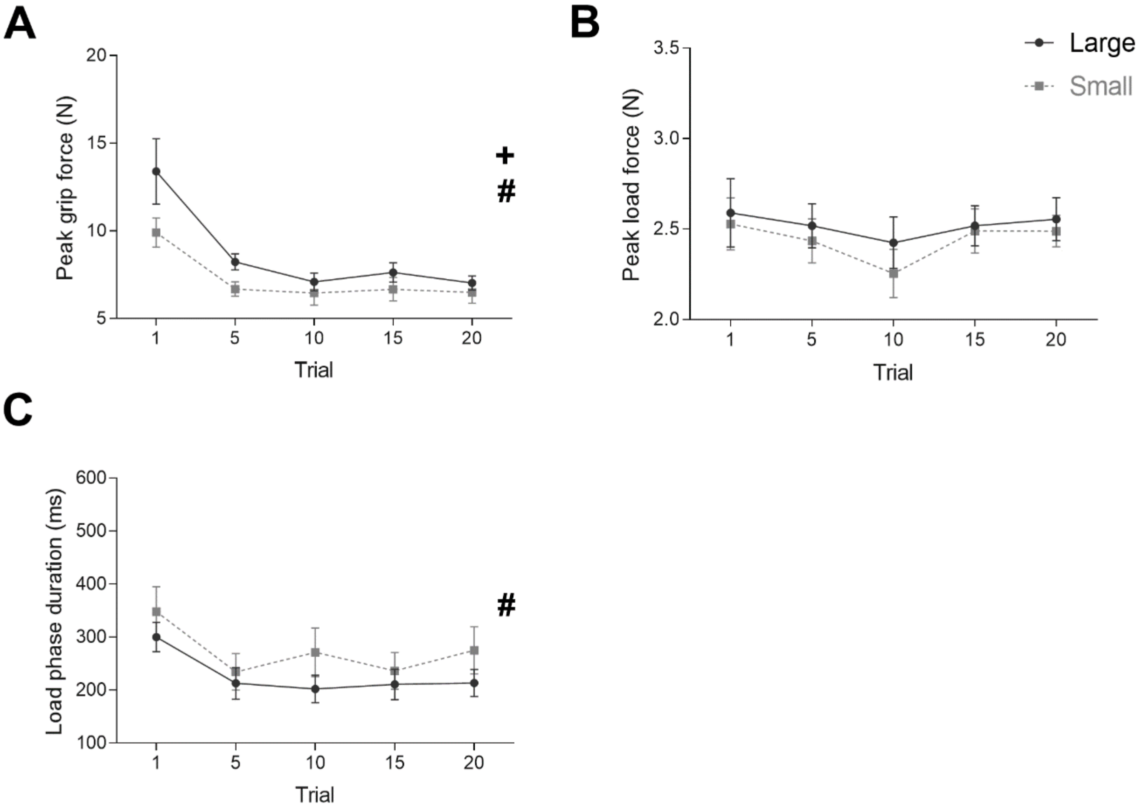

Figure S4. Mean peak grip forces (A), peak load forces (B) and load phase durations (C) for the large and small bottles in Experiment 2b. Error bars denote standard errors of the means. Cross (+) denotes a significant (*p* <.05) main effect of Bottle. Pound (#) denotes a significant (*p* <.05) main effect of Trial.

Table S1

Mean (and standard error) peak grip forces, peak grip force rates, peak load forces, peak load force rates and load phase durations for the full and half-full bottles (Experiments 1a and 2a) and large and small bottles (Experiments 1b and 2b) for trials 1, 5, 10, 15, and 20. In Experiment 1, bottles were lifted via strings. In Experiment 2, bottles were lifted via an attached handle.

Experiment 1a	Full bottle					Half-full bottle				
	Trial 1	Trial 5	Trial 10	Trial 15	Trial 20	Trial 1	Trial 5	Trial 10	Trial 15	Trial 20
Peak grip force (N)	11.11 (1.32)	8.71 (0.87)	8.03 (1.38)	7.19 (0.60)	8.66 (0.87)	9.78 (0.74)	7.73 (1.05)	8.11 (0.81)	8.52 (0.98)	7.80 (0.95)
Peak grip force rate (N/sec)	47.70 (7.29)	41.55 (4.30)	40.16 (5.27)	36.32 (3.55)	48.55 (5.05)	45.76 (4.60)	42.63 (7.22)	46.22 (6.26)	47.57 (7.24)	39.58 (5.89)
Peak load force (N)	2.61 (0.15)	2.72 (0.16)	2.55 (0.09)	2.78 (0.13)	2.52 (0.09)	2.46 (0.18)	2.58 (0.15)	2.49 (0.11)	2.62 (0.09)	2.39 (0.11)
Peak load force rate (N/sec)	16.37 (2.27)	20.04 (2.19)	18.78 (1.99)	19.38 (2.24)	21.79 (2.20)	14.97 (2.06)	19.69 (1.74)	17.93 (1.41)	22.70 (1.88)	18.25 (2.78)
Load phase duration (ms)	494.75 (62.99)	351.04 (34.12)	302.18 (39.14)	340.68 (36.02)	275.79 (31.48)	470.79 (54.44)	303.16 (39.33)	322.38 (29.78)	247.46 (32.81)	353.21 (50.71)

Experiment 1b	Large bottle					Small bottle				
	Trial 1	Trial 5	Trial 10	Trial 15	Trial 20	Trial 1	Trial 5	Trial 10	Trial 15	Trial 20
Peak grip force (N)	13.46 (2.12)	9.78 (1.61)	11.38 (3.23)	10.64 (2.67)	10.08 (2.26)	13.52 (2.69)	11.21 (2.60)	10.11 (2.13)	9.45 (1.75)	10.07 (2.13)
Peak grip force rate (N/sec)	68.71 (13.69)	51.10 (8.75)	51.45 (8.97)	42.44 (5.58)	44.11 (6.72)	51.42 (7.79)	53.11 (10.11)	44.06 (8.11)	42.16 (3.83)	45.61 (6.81)
Peak load force (N)	2.86 (0.19)	2.85 (0.25)	2.53 (0.13)	2.64 (0.10)	2.54 (0.14)	2.57 (0.19)	2.75 (0.19)	2.39 (0.16)	2.56 (0.13)	2.72 (0.15)
Peak load force rate (N/sec)	20.96 (2.02)	20.79 (2.61)	20.50 (2.05)	18.66 (2.08)	18.90 (1.82)	18.72 (3.63)	23.26 (3.44)	16.91 (1.57)	19.15 (1.56)	19.69 (1.63)
Load phase duration (ms)	380.86 (38.18)	289.79 (32.99)	296.00 (38.38)	320.68 (50.44)	366.93 (56.07)	477.86 (48.19)	246.29 (75.12)	338.32 (45.14)	265.64 (67.59)	281.57 (79.36)

Experiment 2a	Full bottle					Half-full bottle				
	Trial 1	Trial 5	Trial 10	Trial 15	Trial 20	Trial 1	Trial 5	Trial 10	Trial 15	Trial 20
Peak grip force (N)	10.47 (1.47)	7.98 (1.47)	7.57 (1.59)	7.80 (1.15)	7.32 (1.06)	10.12 (1.61)	8.21 (1.28)	8.76 (2.67)	7.27 (1.02)	7.28 (1.19)
Peak grip force rate (N/sec)	51.78 (8.33)	39.67 (7.19)	43.71 (9.11)	44.65 (9.91)	37.69 (6.00)	46.70 (9.59)	42.54 (8.34)	48.87 (15.12)	37.22 (3.43)	35.23 (4.83)
Peak load force (N)	2.34 (0.11)	2.37 (0.11)	2.26 (0.10)	2.34 (0.11)	2.32 (0.12)	2.18 (0.09)	2.23 (0.11)	2.25 (0.09)	2.27 (0.09)	2.39 (0.15)
Peak load force rate (N/sec)	19.43 (2.33)	17.40 (1.43)	18.40 (1.39)	20.21 (2.32)	18.84 (1.96)	15.33 (1.22)	17.75 (1.71)	17.71 (1.83)	18.46 (1.87)	19.46 (1.62)
Load phase duration (ms)	303.25 (41.57)	274.71 (26.90)	199.46 (33.93)	265.71 (52.38)	266.50 (36.94)	439.50 (45.37)	312.46 (56.33)	252.61 (33.71)	252.04 (31.88)	259.57 (26.31)
Experiment 2b	Large bottle					Small bottle				
	Trial 1	Trial 5	Trial 10	Trial 15	Trial 20	Trial 1	Trial 5	Trial 10	Trial 15	Trial 20
Peak grip force (N)	9.90 (0.83)	6.68 (0.41)	6.45 (0.68)	6.67 (0.67)	6.49 (0.61)	13.39 (1.86)	8.23 (0.45)	7.09 (0.51)	7.63 (0.55)	7.03 (0.40)
Peak grip force rate (N/sec)	50.72 (6.79)	42.66 (5.71)	42.33 (5.20)	39.63 (4.81)	34.75 (3.87)	66.68 (11.13)	47.46 (3.91)	49.58 (7.55)	40.96 (4.61)	41.45 (3.15)
Peak load force (N)	2.53 (0.14)	2.44 (0.12)	2.26 (0.13)	2.49 (0.12)	2.49 (0.09)	2.59 (0.19)	2.52 (0.12)	2.43 (0.14)	2.52 (0.11)	2.56 (0.12)
Peak load force rate (N/sec)	23.88 (3.85)	22.43 (2.87)	22.16 (2.78)	21.84 (2.46)	21.02 (2.67)	23.41 (3.34)	24.86 (2.45)	22.27 (2.82)	24.30 (2.35)	23.21 (1.82)
Load phase duration (ms)	348.25 (46.90)	234.57 (34.54)	271.36 (45.83)	236.50 (34.44)	275.18 (44.52)	300.32 (27.70)	213.00 (29.74)	202.29 (25.96)	210.93 (28.44)	213.43 (25.35)