Appendix A.

Qualitative responses to the question which scale participants ***liked*** the most. These consist of responses in Dutch and English as both national (Dutch) and international students participated. Dutch responses are translated and printed *italic*

|  |  |  |  |
| --- | --- | --- | --- |
| Original | The descriptions made it more clear what each score meant | Rater 1 | Rater 2 |
|  | deze schaal gaf letterlijk aan wat welk cijfer betekent, wat het makkelijker maakte om te beslissen welk cijfer ik gaf  *This scale literately showed what each number represents, which made it easier to decide which number to choose.* | c | c |
|  | I found it the easiest to indicate my answer on this scale. The other ones were too abstract and I wasn't always sure which option was the most suitable. | c | c |
|  | It is most clear as to what each answer means | c | c |
|  | Het was het duidelijkste voor mij. Doordat er boven stond wat elk punt betekende.  *It was the clearest one for me.* *Because above it was written what each point meant.* | c | c |
|  | Je ziet hier per categorie waar het precies voor staat. Naar mijn mening kan je hier beter aangeven wat je van een riddle vond dan wanneer je aangeeft hoeveel kilo voor de lastigheid ervan staat.  *One can see per category what it represents. In my opinion, it is better to assess by using a riddle than when you mention how many kilos the difficulty weights.* | c | c |
|  | omdat ik hiermee heel duidelijk aan kan geven wat ik van de vraag vond  *because with this, I can very clearly indicate what I think of the question* | c | c |
|  | It's the most specific scale | a | n/a |
|  | It seemed more clear and there is little room for misinterpretation | c | c |
|  | Duidelijke beschrijving erbij van waar de cijfers precies voor staan  *Clear description of what the numbers represent* | c | c |
|  | Daar ben ik het meest bekend mee, en het in woorden verwoorden vind ik het makkelijkst.  *I am most familiar with this, and I find translating it to words is the easiest* | r | r |
|  | De emoticons en de gewichten vind ik wat abstract. De 9 puntsschaal vind ik duidelijker dan de VAS omdat ik liever kies tussen Veel/zeer veel dan 80-90  *I find the emoticons and weights a bit abstract. I find the 9 point scale clearer than the VAS because I prefer to choose between much/very much than 80/90* | c | c/r |
|  | Omdat hier een beschrijving bij stond van wat de score precies inhield.  *Because this one had a description on what the scores represented exactly* | c | c |
|  | It was easy to give an accurate estimation | c | c |
|  | It is more clear to indicate. | c | c |
|  | Seems "adult" and easy to push a button instead of deciding a place on a moving scale. | c | c |
| VAS | Niet één keuze, maar je kan slepen.  *Not one choice but you can swipe* | n | n |
|  | Gaf de nuances het meest verfijnd neer (continue schaal)  *Represented the nuances most refinedly (continuous scale)* | n | n |
|  | Because the rating is not fixed, it can give flexibility to how one perceives the task. | n | n |
|  | It is most accurate | c | c |
|  | It is the easiest way to weight the difficulty | c | c |
|  | There's less specification - it's an actual scale so you have more room for nuance | n | n |
|  | It allows for values between the ones given | n | n |
|  | It is the most sensitive scale. | n | n |
|  | hiermee is er niet echt een ´vaststaand´ punt, maar kun je ook in een grijs gebied zitten. Meer mogelijkheden  *with this there is not really an “ absolute” position, but one van be in a grey area. More possibilities* | n | n |
|  | Easiest to express my opinion on a scale that isn't limited to specific number (9 points scale). The emoticons felt unrelated and the weights were just annoying. | n | c/n/r |
|  | It was easier to how difficult the question was/how much effort you put into it | c | c |
|  | Omdat ik niet hoefde te lezen. De schaal sprak voor zich.  *Because I did not need to read. The scale spoke for itself* | c | c |
|  | I found the visual analogue scale better because it wasn't as strict and specific as the other scales | n | n |
|  | Die met veel moeilijk (dus de woorden) is niet duidelijk geformuleerd. Die met de emoji's zijn het onduidelijkst. Die met alleen de getallen zijn ook fijn, maar de komma getallen maken het weer moeilijker. Die met de gewichten zijn duidelijk en zorgen dat je snel je gedachtes aan kan klikken ipv dat je ook over de komma getallen moet denken.  *The one with “very difficult” (so the words) is not clearly formulated. The one with the emoji’s are the most unclear. The ones with the weights are clear and enable you to quickly click on your thoughts instead of having you think about the decimal numbers.* | c | c |
|  | This was the best representative scale | r | r |
| Emoticons | Very visual, exactly knew how to feel | r | r |
|  | The emoticons one I found the most interesting | r | r/a |
|  | de poppetjes representeren mijn uitdrukking | r | r |
|  | It is easier to pair the difficulty visually rather than assigning numbers because I can compare the difficulty to other problems and anchor my answers around it. | c | c |
|  | because it was a better illustration of how i felt | r | r |
|  | It helps me not have to think about words or descriptions, and instead let me interpret the difficulty with the way I am feeling. These questions made me feel frustrated and confused, so it was easiest to indicate on a scale of what I am feeling with the colors and emoticons. | c | c |
|  | Most visually appealing and quick to use | a | a |
|  | De emoticons gaven een concreet beeld van wat de beoordelingen betekenen.  *The emoticons gave a concrete image of what the assessments meant* | c | c |
|  | deze zag er duidelijker uit.  *This one looked clearer* | a | a |
|  | Describes the emotion felt when solving the problem | c/r | r |
|  | Door de kleuren  *Because of the colors* | r | r/a |
|  | Door de uitdrukking, geeft dit een duidelijk beeld.  *Because of the expression, that provides a clear image* | c | c |
|  | I was feeling bad for not being able to answer to questions, but the smileys made me smile. The rest of the scales were depressing. | r | r/a |
| Weights | It felt most physically relatable and real instead of abstract | c/r | r |
|  | To me, it is the most intuitive because I know what weight feels like. With the visual analogue scale for example I tend to doubt whether I am accurately describing my experience or if I would think totally differently had I had other problems for comparison | r | r |
|  | Gaf duidelijk het verschil aan tussen de verschillende waarden | c | c |

*Note.* c = clarity, r = relatedness, n = nuance and, a = appearance

Appendix B.

Qualitative responses to the question which scale participants ***disliked*** the most. These consist of responses in Dutch and English as both national (Dutch) and international students participated.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Rater 1 | Rater 2 |
| Original | It felt a bit too basic for me, and too limited. It felt like I had to try to figure out what was meant with the scale points. It felt too close-mindedly open-ended..... if that makes any sense. I think that compared to the other ones which let my creative mind indicate what I felt. The 9 point scale just felt a bit more neutral at every point of the scale compared to the other scales. | c/r/n | r |
|  | The Text was overly complicated for the answers I wanted to give | c | c |
|  | Omdat je hier moest lezen wat alles betekende.  *Because you had to read here to find out what everything meant* | c | c |
|  | Er wordt daarbij te veel gebruik gemaakt van tekst  *There is too much text used* | c/r/a | c |
|  | The descriptions feel too definite. Maybe I am too indecisive but it is hard for me to decide on one description | r | c/r |
|  | It is boring and less accurate | c/n | c |
|  | It's hard to distinguish between "low", "very low" and "very very low" | c/n | n |
|  | Veel moeilijk en zeer veel moeilijk klinkt niet als goed Nederlands, daarom moet ik eerst nadenken wat er bedoeld word.  *Very much difficult and very much difficult does not sound as proper Dutch, therefore I need to think first what is meant* | c | c |
| VAS | Boring | a/r | r |
|  | It is too ambiguous | c | c |
|  | Weinig overzicht  *Little overview* | c | c |
|  | It allows for values between the given numbers | n | n |
|  | It is least clear as to how my answer will be interpreted | c | c |
|  | Het voelt een beetje willekeurig om hierbij in te vullen hoe lastig je de opdracht vond.  *It feels a bit random to fill in here how difficult the assignment was* | c/n | c/n/r |
|  | There is the most room for misinterpretation | c | c |
| Emoticons | It felt annoying | r | r |
|  | It is the hardest way to weight the difficulty | c | c |
|  | The emoticons aren't clear as to what they exactly represent | a | c/a |
| Weights | te subjectief  *too subjective* | c | c/r |
|  | It makes no sense. | c/r | r |
|  | Kilo past niet bij de beoordeling  *Kilo does not match to assessment* | r | r |
|  | incongruent with what the scale is asking | r | r |
|  | Moeilijk te koppelen aan moeilijkheidsgraad  *Hard to link to difficulty* | r | c/r |
|  | It made me feel as if I was actually carrying the weight. | r | r |
|  | Vaag ten opzichte van de betekenis  *Vague with regard to the meaning* | c | c/r |
|  | Vond het irritant dat de gewichten steeds groter werden. Zorgt voor weinig overzicht.  *Found it annoying that the weights became larger. Presents little overview*. | a | a |
|  | The scale was huge, did not fit onto my screen, the kilograms did not seem relevant to the mental task at hand. | a | a |
|  | It is very ugly and not very meaningful | a | a |
|  | Voelde het minst logisch, ik vond dat de andere lijsten de verschillende duidelijker aanduidden, terwijl de lijsten op precies dezelfde manier werkten.  *Felt the least logical, I found the other lists pointed out the differences clearer, despite that the lists work the same way* | r | r |
|  | They were annoying to maneuver and also felt quite related. emotions close second. | r | r |
|  | The pictures made it look really heavy and overwhelming | a | a |
|  | ik had het gevoel dat ik vaker de extreme waardes koos.  *I felt as if I chose the extreme values more often* | n | n |
|  | Unsure | c | c |
|  | de gewichten zeiden mij veel minder over de moeilijkheid van deze vraagstukken en de moeite die het koste, de andere schalen voelden logischer en makkelijker om in te vullen  *The weights told me less about the difficulty on these matters and the effort it cost, the other scales felt more logical and easier to fill in.* | r | r |
|  | It was hard to indicate my answer on this one because the weights made me feel like even 5 kg already had too much weight (so that it indicated a lot of mental effort and difficulty) and therefore I chose the smaller scales even for higher mental effort | r/n | c/r |
|  | Ik vond het niet relevant dat het over gewichten ging en snapte de toevoeging ervan niet  *I did not find it relevant that it was about weights and did not understand the added value of it* | r | r |
|  | dit voelt heel abstract. (en 9 kg zie ik niet als heel zwaar) Dus een moeilijke vraag met gewicht 9 voelt niet als heel moeilijk  *This feels very abstract. (and I do not consider 9 kg as very heavy) So a hard question with a weight of 9 does not feel as very hard* | c/r | r |
|  | Gewichten zijn nooit gebruikt in andere vragenlijsten die ik gemaakt heb waardoor het misschien even wennen was  *Weight are never used in other questionnaires that I have encountered, so maybe I had to get used to it* | r | r |
|  | Ik ben niet gewend om een moeilijkheidsgraad in kilo's uit te drukken.  *I am not used to express the level of difficulty in kilo’s.* | r | r |
|  | Meest abstract in deze context.  *More abstract in this context.* | r | c/r |
|  | Omdat dit niet op mijn scherm paste, dus ik moest steeds scrollen.  *Because it not fitted on my screen, I had to scroll all the time* | a | a |
|  | Difficult to combine mental difficulty with weight | c | c/r |
|  | It cannot indicate | n | c/n |
|  | Makes me feel "brainwork" is like exercising. | r | r |

*Note.* c = clarity, r = relatedness, n = nuance and, a = appearance