

Resultaten

Uit de enquêtes die door middel van het onderzoek zijn ingevuld hebben we een schematisch overzicht gemaakt (zie Appendix) van de verkregen data. Deze worden hieronder verder geanalyseerd. Eerst wordt er apart gekeken naar de twee verschillende ontmoetingen met AMIGO en daarna worden deze twee benaderingen (Real-life VS Beeld) met elkaar vergeleken. Om het makkelijk te maken, definiëren we de groepen als volgt:

- Groep 1.0 is de groep die de real life AMIGO nog moet ontmoeten;
- Groep 1.1 is de groep die de real life AMIGO heeft ontmoet;
- Groep 0.0 is de groep die de visuele AMIGO nog moet ontmoeten;
- Groep 0.1 is de groep die de visuele AMIGO heeft ontmoet;

Let op: we hebben dus te maken met 2 verschillende groepen en 2 verschillende situaties per groep.

'Real-life' ontmoeting met AMIGO (groep 1.X)

Als eerst bekijken we of er sprake is van een normale verdeling bij de items(vragen en taken). Hiervoor gebruiken we de Shapiro-Wilk test. In dit onderzoek maken we gebruik van 95% betrouwbaarheidsintervallen en een alpha waarde van 0.05. Uit de appendix is af te leiden, dat de items (vragen) 3, 10, 12, 14, 15, 23, 27, 31, 34 en 36 een waarde aangeven dat kleiner is dan 0.05. Dit betekent dat de H_0 hypothese (de data is normaal verdeeld) verworpen mag worden. Dus voor deze items is er geen sprake van een normale verdeling. Voor de taken wordt er op dezelfde wijze afgeleid dat er ook enkele van deze taken zijn waar er geen sprake is van een normale verdeling (zie appendix).

Met behulp van deze informatie kunnen we later rekening houden met het kiezen van de soort test, o.a. 'ttest', 'multiple regression' of 'Ranksum'.

Nu gaan we verder met het bekijken van de 'Cronbach's alpha' die zowel bij de taken als bij de vragen 1 tot en met 38 horen. Let op, sommige vragen die gesteld zijn in de enquête hebben we uiteindelijk geschrapt. Deze vonden we niet bij het onderzoek passen.

De verdeling van de vragen in de eerder genoemde categorieën (zie opzet).

- Uiterlijk: 11, 14 en 23
- Voorkennis/interesse: 1, 2, 3, 13 en 30
- Normen en waarden: 4, 7, 10, 15, 17, 19, 20, 22, 24, 26, 28, 29, 33, 34, 35 en 38
- Toekomst: 5, 6, 8, 9, 16, 18, 21, 25, 31, 32, 36 en 37
- Taken: Alle items onder het kopje 'taken' in de enquête.

Het meten van 'angst' gebeurt in principe door middel van de bovengenoemde categorieën. Als extra, gaan we op het eind nog testen of de mate van voorkennis/interesse invloed heeft op de hoeveelheid angst bij de proefpersonen. Hier komen we later nog op terug.

Voor de 'Cronbach's alpha' die hoort bij de taken en vragen voor zowel voor en na interactie, kunt u bij de appendix terecht. Deze heeft een waarde van respectievelijk 0.9194 0.9016 0.7992 en 0.4293. Meteen valt er op te merken dat de laatste Cronbach's een lagere waarde aangeeft. Dit betekent dat er verschil is in het soort impact die de robot heeft op de proefpersonen. Het kan ook zo zijn dat de proefpersonen de vragenlijst niet meer serieus in hebben gevuld om zo snel mogelijk klaar te zijn. Omdat de vragen namelijk precies hetzelfde zijn, moeten ook deze na de interactie nog steeds angst meten.

De volgende tests bepalen of de scores voor taken en de angstvragen, die door de alpha functie in stata verkregen zijn, normaal verdeeld zijn en of deze gelijke varianties onder de groepen (groep 1.0 en groep 1.1) aannemen.

Het commando 'swilk' (oftewel de Shapiro-Wilk test) test of er sprake is van een normale verdeling. Bij groep 1.0 (taken) hebben we een p-waarde van 0.12886 en bij de 1.1 een waarde van 0.85596 (voor alle resultaten kunt u weer bij de appendix terecht). Beide waardes zijn groter dan 0.05 dus de hypothese dat de data normaal verdeeld is kan niet worden verworpen. Voor de angstvragen zijn dat respectievelijk 0.03691 en 0.00150. Hieruit valt op te maken dat er geen sprake is van normale verdeling bij groep 1.0 en 1.1 als het gaat om angst.

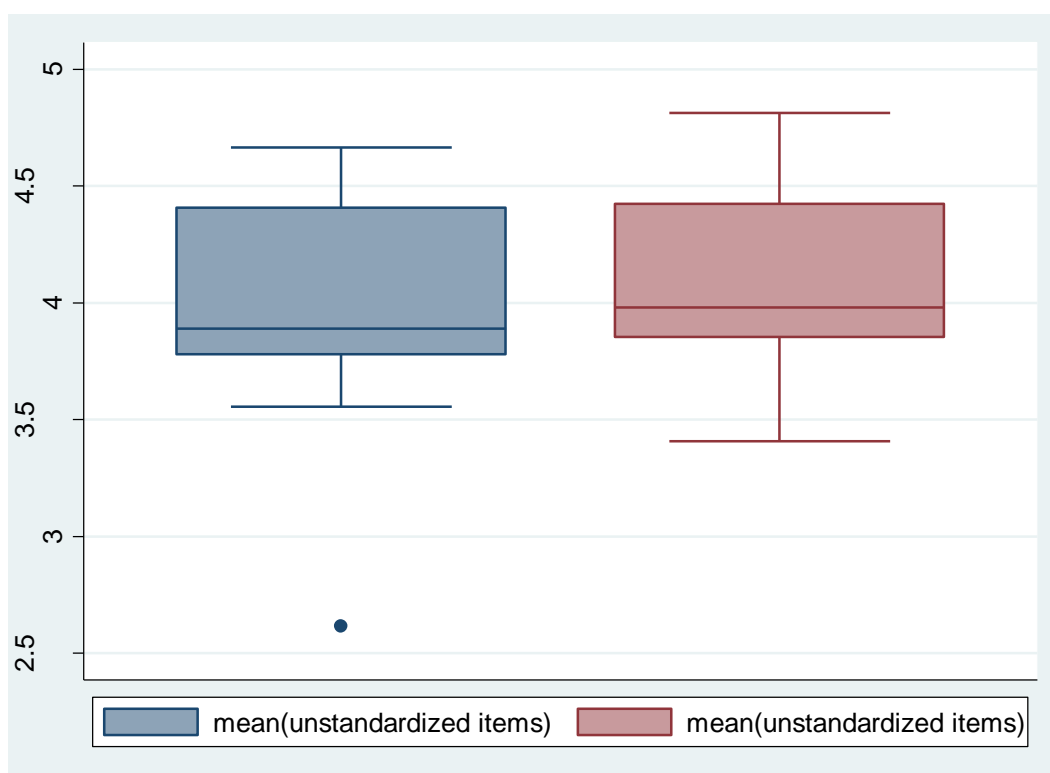
Nu gaan we testen of er sprake is van gelijke variantie tussen de twee groepen. Dit testen we met het commando 'sctest' in stata. Weer kijkend naar de p-waardes, kunnen we concluderen dat de nulhypothese kan worden verworpen bij de vragen. Dus ook geen gelijke variantie.

Onder de taken wordt zowel normale verdeling als gelijke variantie niet verworpen.

Als aangenomen is dat we te maken hebben met een normale verdeling en gelijke variantie tussen twee groepen, dus voldaan is aan de aannames, kan er een ttest uitgevoerd worden. Deze test bepaalt of er sprake is van gelijke gemiddelde en dus of er een verschil is tussen twee groepen.

In ons geval is de aanname van normale verdeling bij de angstmeting geschonden en daar moeten we rekening mee houden.

Laten we eerst de taken bekijken, zie box plot hieronder.

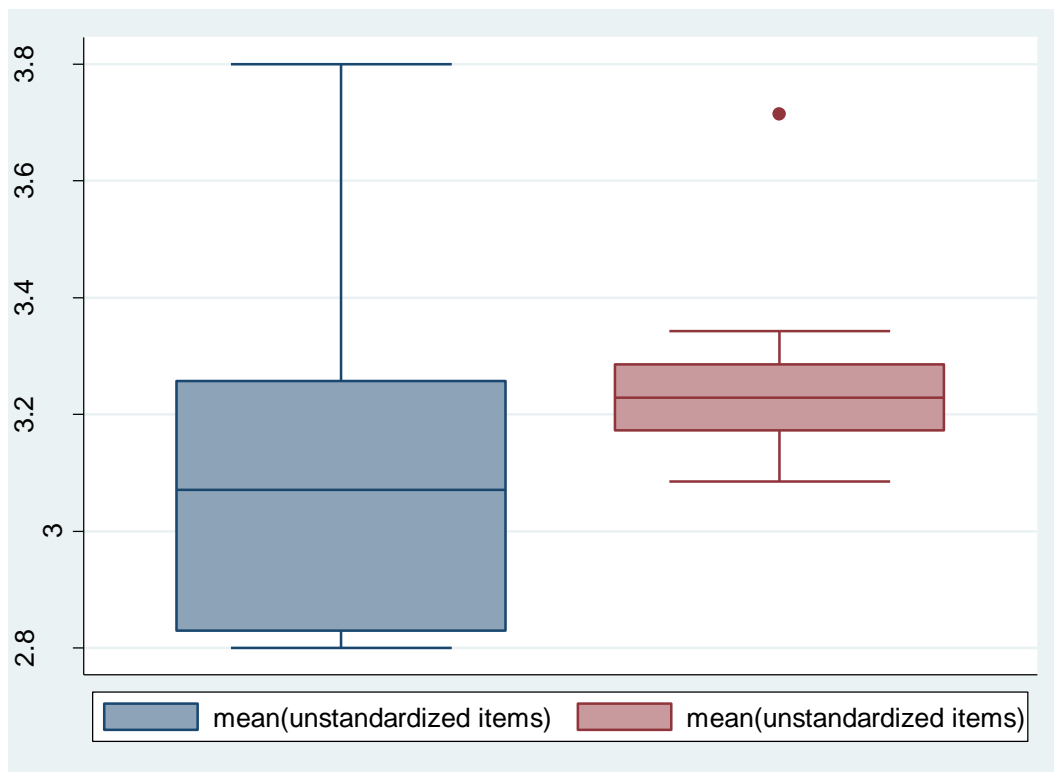


Het blauwe boxje hoort bij de vragen voor het onderzoek (taak10) en de rode bij de vragen erna (taak11). De verticale schaal stelt het effectgrootte voor (mate van behoefte).

Op te merken is, dat er een duidelijk aanwezige overlap tussen de twee gemiddeldes aanwezig is. Dit betekent dat er geen verschil aanwezig is tussen de taakgroep 1.0 en 1.1.

Laten we hiervoor een test in stata uitvoeren. We weten dat bij de taken normale verdeling niet wordt geschonden, dus wordt de ttest uitgevoerd. Uit deze test blijkt dat we voor de taken geen verschil aan kunnen tonen (p-waarde 0.4568). Dit bevestigde het vermoeden uit de box plot.

Hetzelfde analyse maar dan voor de angstvragen. Zie weer hieronder de bijbehorende box plot.



Zoals te zien is in de grafiek, is de spreiding bij de blauwe box groter.

Maar als we kijken naar de angstvragen, dan kunnen we met 95% zekerheid zeggen dat er een verschil bestaat tussen groep 1.0 en 1.1. Hiervoor hebben we de ranksum-test gebruikt waarbij een p-waarde van 0.0499 uitkwam.

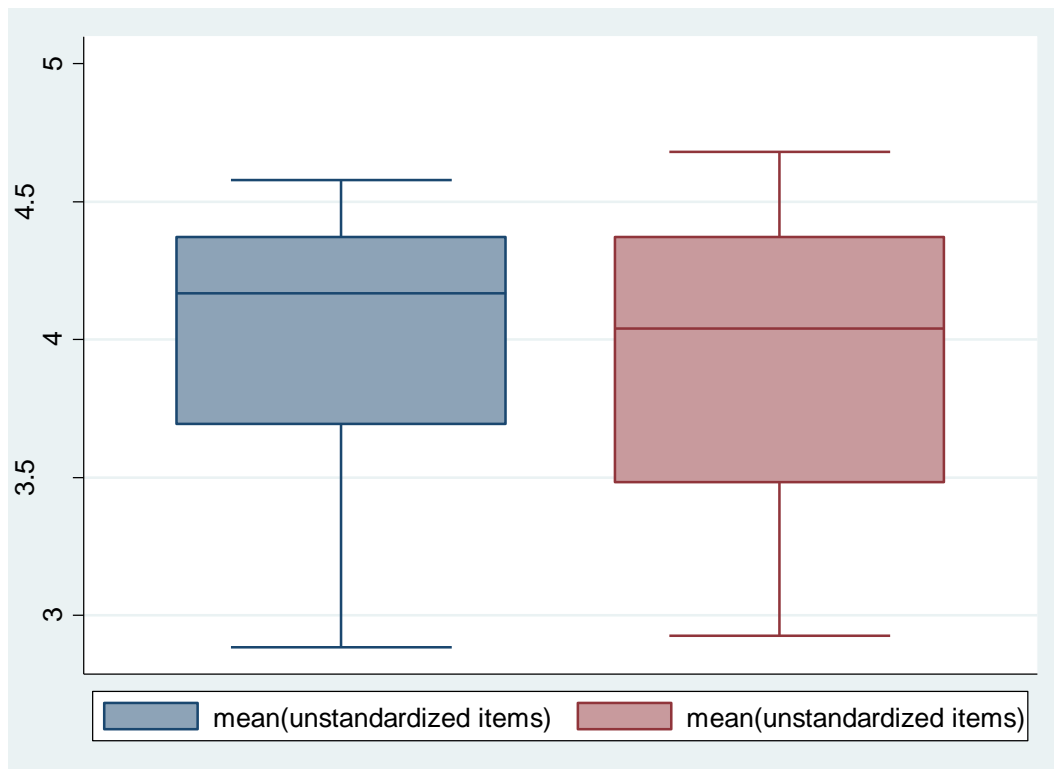
De groep na interactie met AMIGO bleek iets minder bang te zijn voor robots dan de groep voor interactie. Het gemiddelde voor interactie is 3.08 en het gemiddelde na 3.25.

'Visuele ontmoeting met AMIGO

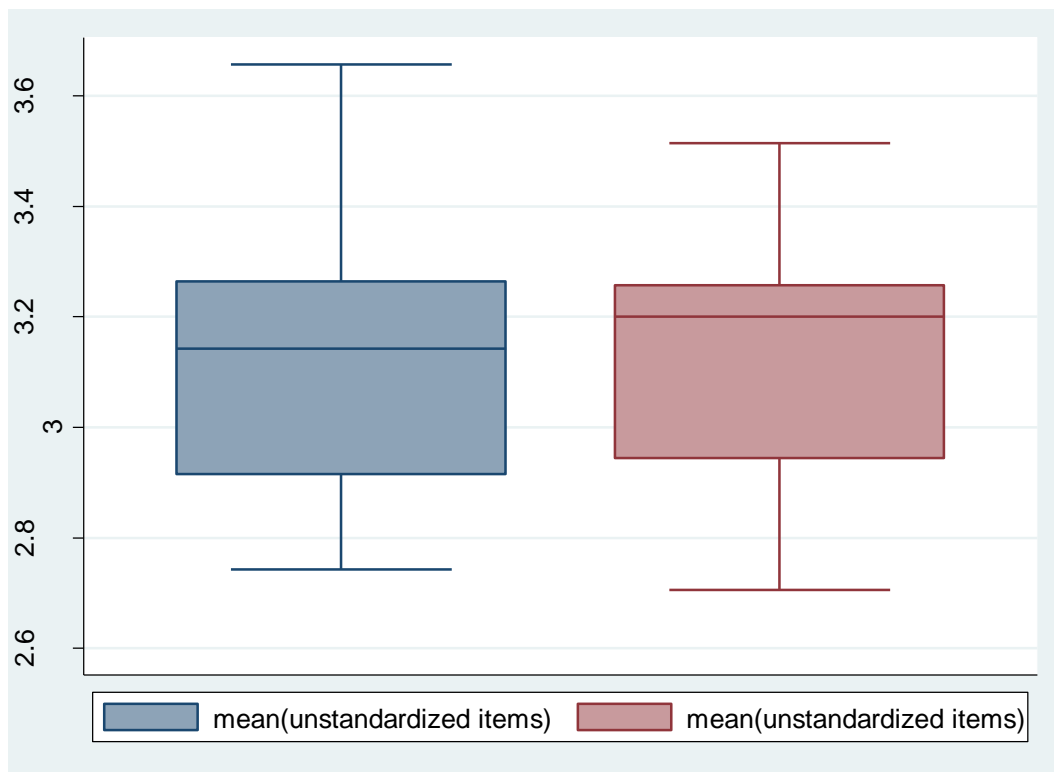
We hebben in deze situatie telkens hetzelfde aanpak als bij de 'real-life' ontmoeting met AMIGO. Zie weer de Appendix voor de verkregen data.

Uit de visuele ontmoeting blijkt dat we een ttest uit kunnen voeren (dus er is voldaan aan de aannames). Uit deze test blijkt dat er geen verschil is tussen de groep voor interactie en de groep na interactie. Dit geldt voor zowel de taken als voor de angstvragen. De p-waardes zijn respectievelijk 0.9605 en 0.9446.

Box plot, horend bij de taken (voor en na interactie).



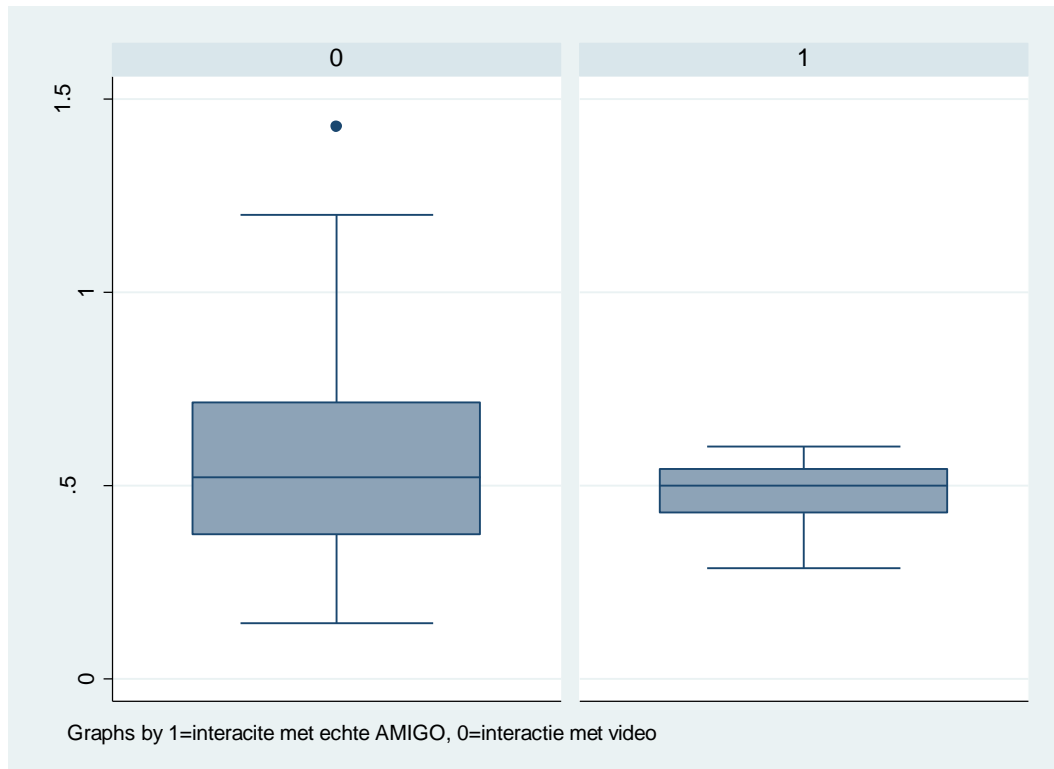
Box plot horend bij de angstvragen (voor en na ontmoeting).



Real-life VS Visuele ontmoeting

Om de hoofdvraag te beantwoorden is het belangrijk naar het verschil te kijken in de twee aanpakken die we hebben toegepast. Bij het vergelijken van de taken met elkaar, die aan de ene kant de real-life invloed meten en aan de andere kant de visuele, komen we op een p-waarde van

Laten we eerst kijken naar de box plot van de twee verschillende interacties (groep 0.1 en 1.1).

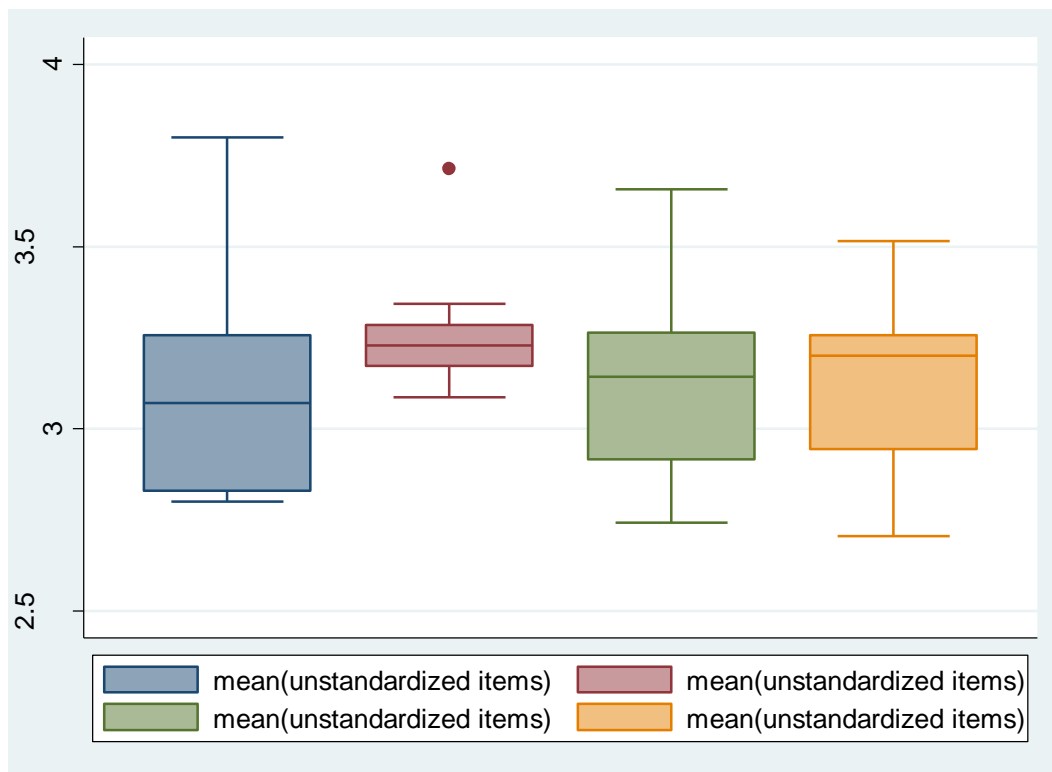


Er valt op te merken dat de spreiding bij de visuele interactie veel groter is. Dit zou kunnen betekenen dat de impact van de real life AMIGO groter is. Het kan wel nog steeds zo zijn dat de gemiddeldes van deze twee groepen gelijk zijn. Een andere mogelijkheid is, dat de spreiding zo groot is omdat de omstandigheden waarin de enquêtes gemaakt zijn (bij het visuele gedeelte) niet gelijk waren per proefpersoon en dus voor een grote spreiding zorgde.

Er wordt niet voldaan aan de assumptie voor normaal verdeeldheid, dus gaan we nu de ranksum test toepassen op deze twee groepen.

Opmerkelijk is, dat we uit deze test niet kunnen concluderen of er een verschil is of niet. Met een p-waarde van 0.5044 kunnen we hier dus niets over zeggen.

Hieronder nogmaals de box plot, met alle 4 gevallen ineen.



De blauwe box geeft de het geval vóór een interactie met de robot.
De boxen geven de groepen 1.0 1.1 0.0 0.1 in deze volgorde weer.

Tussen de data lijkt er geen verschil te zitten. Maar dit zouden we dus moeten testen met een grotere sample size, want het kan wel zo zijn dat er alsnog verschil is tussen de bovenstaande gevallen.

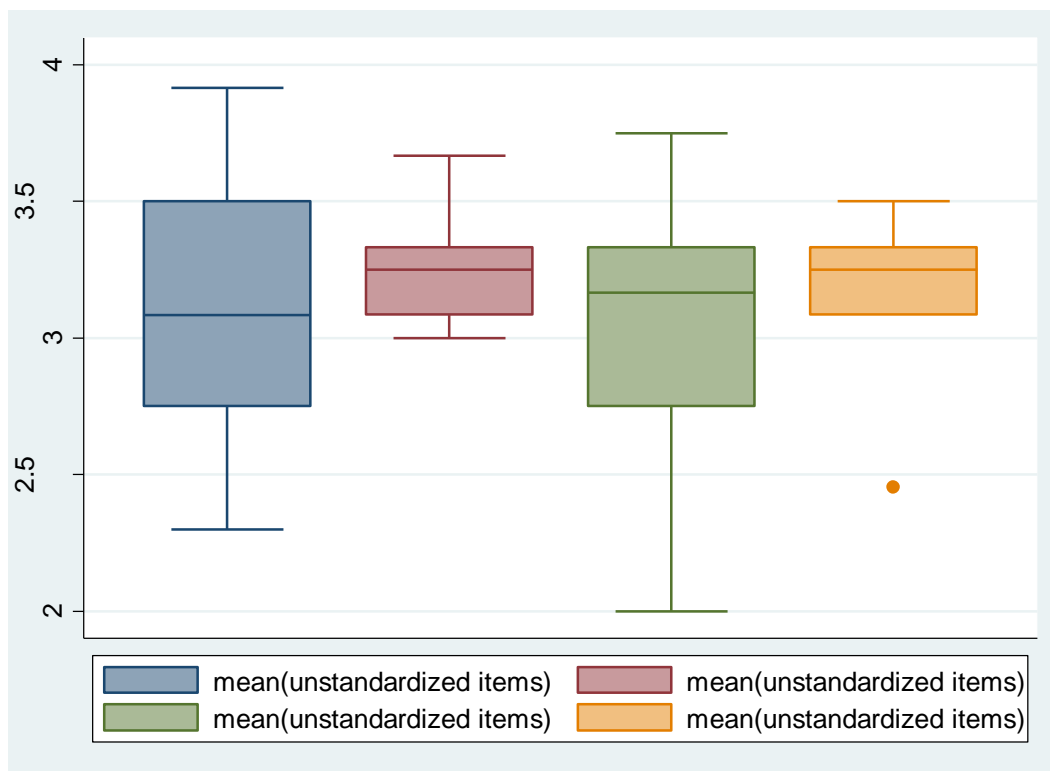
Hieronder nog even de gemiddeldes in getallen:

Variable	Obs	Mean	Std. Dev.	Min	Max
angst10	14	3.084079	.2822434	2.8	3.8
angst11	14	3.24898	.1513637	3.085714	3.714286
angst00	14	3.151478	.2958119	2.742857	3.657143
angst01	14	3.144205	.2513399	2.705882	3.514286

Hoe hoger de waarde, hoe kleiner de angst en hoe groter het vertrouwen in robots. De waarde 3 geeft een neutrale stand aan. We hadden eerder aangetoond dat er een significant verschil is tussen de groepen 1.0 en 1.1 bij de angstvragen. Hierboven kun je aflezen dat deze verschil bestaat, maar niet al te groot is. Na de interactie met de real-life AMIGO blijken de proefpersonen minder bang te zijn.

Hoe denken de mensen als het gaat om de toekomst?

We hebben gekeken naar de resultaten per categorie. Hieronder ziet u de resultaten die horen bij de categorie toekomst.



Groep 1.0 1.1 0.0 0.1 (in deze volgorde).

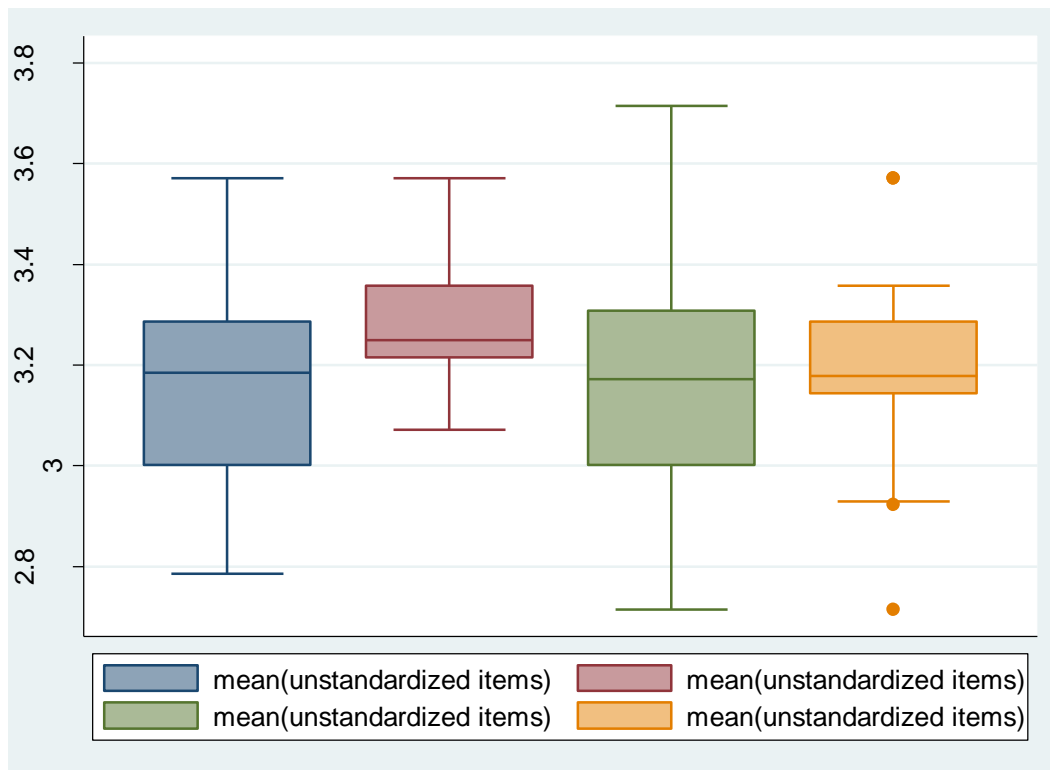
De bijbehorende gegevens:

Variable	Obs	Mean	Std. Dev.	Min	Max
toekomst10	14	3.08438	.4627357	2.3	3.916667
toekomst11	14	3.267857	.1735089	3	3.666667
toekomst00	14	3.055952	.4214143	2	3.75
toekomst01	14	3.199134	.2644334	2.454545	3.5

Na het uitvoeren van de ttest, kunnen we geen significant verschil aantonen (zoals verwacht).

Vindt er verandering plaats in de normen en waarden?

Voor de categorie normen en waarden ziet u hieronder de bijbehorende box plot.



Groep 1.0 1.1 0.0 0.1 (in deze volgorde).

Ook hier geldt dat de assumpties voor een ttest (normaal verdeeld en gelijke variantie) niet geschonden worden. De ttest uitvoeren levert een verschil van groep 1.0 en 1.1 met 94% zekerheid (p-waarde ongeveer 0.6).

Hieronder even de samenvatting van de gegevens.

Variable	Obs	Mean	Std. Dev.	Min	Max
toekomst10	14	3.08438	.4627357	2.3	3.916667
toekomst11	14	3.267857	.1735089	3	3.666667
toekomst00	14	3.055952	.4214143	2	3.75
toekomst01	14	3.199134	.2644334	2.454545	3.5

De groep na de real-life AMIGO interactie lijkt een meer positieve kijk te hebben dan ervoor, als het gaat over normen en waarden.

Wiskundig model

Om de mate van angst te bepalen uit de verschillende vragen, wordt Cronbach's alpha vaak gebruikt. Het is een maat om de validiteit van de data te testen. M.a.w. de waarde van alpha meet in hoeverre de items (vragen) hetzelfde concept meten. Deze alpha waarden wordt als volgt berekend:

$$\alpha = \frac{N}{N-1} \left(1 - \frac{\sum_{i=1}^N S^2_{Y_i}}{S^2_X} \right)$$

Daarin stellen de S^2 -en de steekproefvarianties voor van de totale score X en de scores op de items Y_i . Hierbij is $X = \sum_{i=1}^N Y_i$ met N het aantal items.

Verder geldt als Y_{ik} de score is van testpersoon k op item i en er zijn m testpersonen:

$$S^2_{Y_i} = \frac{1}{m-1} \sum_{k=1}^m (Y_{ik} - Y_{i\cdot})^2$$

Waarin de gemiddelde score op item i als volgt gedefinieerd is:

$$Y_{i\cdot} = \frac{1}{m} \sum_{k=1}^m Y_{ik}$$

en

$$S^2_X = \frac{1}{m-1} \sum_{k=1}^m (X_k - \bar{X})^2$$

waarin

$$X_k = \sum_{i=1}^N Y_{ik}$$

de testscore van testpersoon k is.

Cronbach's alpha:

http://en.wikipedia.org/wiki/Cronbach%27s_alpha

Shapiro-Wilk:

http://en.wikipedia.org/wiki/Shapiro%E2%80%93Wilk_test