

Dieptedronkenschap: ook getest op duikers

De wetenschappelijke uitleg

In Hippocampus nr. 238 (maart/april 2012) staat het artikel over het dieptedronkenschap-experiment in Nemo33. Om het artikel niet te overladen werd de wetenschappelijke uitleg weggelaten. Voor wie er toch in geïnteresseerd is, kan deze hier lezen.

Speciaal voor de lezer die iets van statistiek afkent. Het verschil in prestaties tussen de verschillende diepten en de invloed van ervaring en geslacht, werd getest met een 2 (diepte: 2,5 m versus 35 m) x 2 (ervaring: Instructeur versus 2* en 3*D) x 2 (geslacht: vrouw versus man) herhaalde metingen ANOVA met diepte als binnensubjectvariabele (elke deelnemer wordt getest op elke diepte: de variabele diepte varieert dus binnen het subject) en met ervaring als tussensubjectvariabele 1 (elke deelnemer is ofwel Instructeur, ofwel 2*D/3*D: de variabele ervaring varieert dus tussen de subjecten) en geslacht als tussensubjectvariabele 2. Voor de drie onderzoeksvragen (snelheid, juistheid en vissengeheugen) hebben we respectievelijk 3 afhankelijke variabelen: aantal vragen ingevuld, aantal vragen juist ingevuld, aantal kenmerken herkend.

ANOVA-statistiek (van het Engelse 'Analysis of variance'), ook bekend als variantie-analyse, is een toetsingsprocedure om na te gaan of de populatiegemiddelden van twee of meer groepen van elkaar verschillen.

Als er een significant hoofdeffect is van de variabele diepte ($F > 1,0$ en $p < 0,05$), kunnen we besluiten dat prestaties op 2,5 m verschillen van die op 35 m. Om na te gaan of die verschillen anders zijn voor Instructeurs dan voor 2*/3*Duikers, kijken we of er een significant interactie-effect is van ervaring en diepte ($F > 1,0$ en $p < 0,05$). Om na te gaan of de verschillen tussen prestaties op 2,5 m en 35 m anders zijn voor mannen dan voor vrouwen, kijken we of er een significant interactie-effect is van geslacht en diepte ($F > 1,0$ en $p < 0,05$).

M = gemiddelde; SD = standaarddeviatie of de gemiddelde afwijking van het gemiddelde (de deelnemers hebben een andere score dan het gemiddelde); SD geeft aan hoe die andere scores afwijken van het gemiddelde; F = de waarde van de ANOVA, p = probabiliteit (deze waarde geeft indicatie van 'significantie').

Als $p < 0,05$ dan is het resultaat significant. Wanneer er wordt gekeken of een resultaat significant is, wordt er nagegaan of het resultaat dat gevonden is in de groep die het experiment heeft gedaan (in dit geval het gevonden gemiddelde en de individuele afwijkingen van het gemiddelde, de gevonden M en SD dus) overeenkomt met wat het resultaat zou zijn in 'de echte wereld' (in dit geval: als alle duikers de test uit het experiment zouden doen). Alle individuele scores op beide diepten, beide gemiddelden (9,56 voor 2,5 m en 5,62 voor 35 m) en beide SD's worden daarom in een ingewikkelde formule gestopt om de F te bereken. Als $F = 0$, is er geen verschil tussen de gemiddelden op 2,5 m en 35 m. Er is wel een verschil als $F > 1$.

We zeggen dat het resultaat uit de experimentgroep overeenkomt met die in de echte wereld wanneer $p < 0,05$ (dat is een 'afpraak' die in de statistiek wordt gebruikt). $p < 0,05$ betekent immers dat, wanneer het effect (in dit geval verschil tussen 2,5 m en 35 m) er NIET zou zijn in de echte wereld, we slechts 5% kans (en dat is weinig) zouden hebben om in de experimentgroep een verschil te vinden tussen 2,5 m en 35 m. Dus als we in ons experiment vinden dat $p < 0,05$ dan betekent het dat we kunnen zeggen dat het verschil dat wij vonden tussen de prestaties op 2,5 m en 35 m niet zomaar een 'toevallig verschil' is, maar dat de kans heel groot is (nl. 95%) dat er ook in het echte leven een verschil is tussen 2,5 m en 35 m (en dat dit verschil ongeveer even groot zal zijn als wat we in het experiment vonden).

Snelheid (aantal vragen ingevuld): er is een significant effect van diepte, $F(1,12) = 4,76$ en $p = 0,04$: er worden gemiddeld meer vragen ingevuld op 2,5 m ($M = 16,81$ en $SD = 7,88$) dan op 35 m ($M = 12,25$, $SD = 7,52$). Er zijn geen significante interactie-effecten, noch tussen diepte en ervaring, $F(1,12) = 0,06$ en $p = 0,82$, noch tussen diepte en geslacht, $F(1,12) = 0,39$ en $p = 0,54$: het effect van diepte is dus hetzelfde voor instructeurs dan voor 2*D/3*D en voor mannen dan voor vrouwen.

Juistheid (aantal vragen correct ingevuld): er is een significant effect van diepte, $F(1,12) = 16,11$ en $p = 0,002$: er worden gemiddeld meer vragen correct ingevuld op 2,5 m ($M = 9,56$ en $SD = 6,09$) dan op 35 m ($M = 5,62$ en $SD = 4,21$). Er zijn geen significante interactie-effecten, noch tussen diepte en ervaring, $F(1,12) = 1,79$ en $p = 0,21$, noch tussen diepte en geslacht, $F(1,12) = 1,79$ en $p = 0,21$: het effect van diepte is dus hetzelfde voor Instructeurs dan voor 2*/3*Duikers en voor mannen dan voor vrouwen.

Vissengeheugen (aantal kenmerken herkend): er is geen significant effect van diepte, $F(1,9) = 0,52$ en $p = 0,49$: er worden gemiddeld evenveel kenmerken van de 2,5 m vis als van de 35 m vis herkend. Er zijn geen significante interactie-effecten, noch tussen diepte en ervaring, $F(1,9) = 3,96$ en $p = 0,08$, noch tussen diepte en geslacht, $F(1,9) = 0,30$ en $p = 0,60$: Instructeurs herinneren zich van beide vissen even weinig als 2*/3*Duikers, en mannen herinneren zich van beide vissen even weinig als vrouwen.

Leentje Vervoort