

Standaardisatie.

Standaardisatie is het credo van veel wetenschappers uit de mens- en dierwetenschappen. In een review artikel van Diederich en Giffoy (2006) *Behavioural testing in dogs: A review of methodology in search for standardisation*, bijvoorbeeld werd hier nog eens flink op gehamerd.

Laat ik vooropstellen dat ik de noodzaak van standaardisatie onderschrijf: anders vergelijk je per definitie al appels met peren. Er valt ook heel veel te standaardiseren. Materiaal, testvolgorde, beoordelingsformulieren; het kan allemaal. Het is naar mijn mening echter een illusie om te denken dat je iedere hond dan op dezelfde manier test of een test precies zo kunt herhalen. Er is immens veel dat niet valt te orkestreren, maar dat weldegelijk een invloed heeft op het gedrag van de hond. Zeker als je buiten test, ook al is het steeds op dezelfde plek, zijn er allerlei ongeplande invloeden, zoals bijvoorbeeld het weer en de omgevingsgeluiden (overkomende vliegtuigen, passerend verkeer) Blaffende honden in de omgeving hebben een enorme invloed op de emoties en reacties van de hond (vooral bij pups) die getest wordt. Je zult naar mijn mening dat soort dingen dan ook moeten noteren en op de een of andere manier moeten verwerken in je beoordeling. (en vooral níét doen alsof het zich niet heeft voorgedaan) Ik beschouwde en scoorde dergelijke dingen (of liet ze scoren) als een soort extra test die vermeld werd op het formulier en vatte ze, als dat kon, onder bestaande testonderdelen, zoals reactie op een akoestische prikkel, of gedrag t.o.v andere honden. Omdat standaardisering voor mij geen doel is om 'objectiviteit' (=intersubjectiviteit) te realiseren, maar een middel, acht ik dat geen probleem. Voor het in kaart brengen van eigenschappen (*traits*) is zo'n 'ongeplande extra test' een toegevoegde waarde omdat hij de *intra-test reliability* alleen maar vergroot.

Wat zich in mijn ervaring niet goed laat standaardiseren is het gedrag van de testers, hoe goed je ze ook traint. Er gaat altijd wel wat mis. Maar afgezien daarvan kun je de emoties van testers niet beheersen. Slecht weer (harde regen en/of wind) heeft bijvoorbeeld vaak een invloed op het humeur en de bewegingen van de testers. Testers hebben emotionele voorkeuren en afkeuren voor honden, die hun gezichtsuitdrukking en stem onwillekeurig beïnvloeden en door honden worden opgemerkt omdat je fungeren als primaire bekrachtigers/straffers (zie bv Rolls).

Door onderzoekers wordt vaak in hun streven naar 'objectiviteit' een oplossing gezocht in de richting van mensen eindeloos trainen in het maken van dezelfde bewegingen of zelfs in de mogelijkheid om robots of zelfs puntlichtfiguren in te zetten. Ik vraag me af of dit wel de richting is waarin oplossingen moeten worden gezocht. Daarmee wordt namelijk eraan voorbijgegaan dat honden sociale en emotionele dieren zijn die zelf ook bewust een sociale interactie aangaan en bepaalde verwachtingen hebben bij interactie met een mens. Als mensen zich onnatuurlijk, robotachtig gedrag vertonen werkt dit vaak al gauw bevreemding of zelfs angst op die reacties weer beïnvloeden en vertekenen. Beter lijkt het mij dat testers zo gewoon mogelijk doen, zodat hun gedrag binnen het referentiekader van de hond valt. Dat levert naar mijn mening minder vertekening op dan robotachtig gedrag.

Ook voor het welzijn van de hond, een factor die naar mijn mening voorop dient te staan, is afstemming op het gedrag van de hond/pup een vereiste. Ik heb video-opnames gezien van de UU agressietest bij Golden Retrievers die mij op welzijnsgebied met afgrijzen vervullen. In alle FRC testen is in de instructies aan de testers en beoordelaars opgenomen dat een testonderdeel onderbroken of overgeslagen moet worden of aan de hond aangepast moet

worden als er sprake is van grote angst of als de hond dichtslaat. Ook worden testers geïnstrueerd dat zij moeten zorgen voor afdoende herstel voordat verder gegaan mag worden met de test. Herstelt de hond onvoldoende dan wordt de hele test afgebroken en wordt de eigenaar/fokker begeleid om te zorgen dat de hond weer in een goede stemming weggaat. Omdat de FRC testen niet al te confronterend zijn, is iets dergelijks zelden aan de orde. Maar toch. Geen een hond mag de dupe worden van de test.

Voor onderzoekers is het gangbaar om testen waarin afgeweken is van het protocol niet te gebruiken binnen een groter geheel. Bij een onderzoek naar de persoonlijkheid van de hond is dat naar mijn mening niet nodig, mits dit goed aangetekend wordt in de scoreformulieren en in de eindscore. Immers het noodzakelijk zijn van een dergelijke aanpassing zegt wat over de persoonlijkheid van de hond en kan dus net zo goed meetellen, mits juist geadmistreerd en verwerkt in de conclusies.

Wat naar mijn mening ook niet helemaal goed valt te standaardiseren, ook niet met allerlei perfecte formulieren, protocollen en scoremethodes, is de beoordeling van de beoordelaars. Ook zij hebben voorkeuren en afkeuren, waardoor de focus bij de ene hond anders komt te liggen dan bij de andere. De eerste indruk is daarbij vaak bepalend. Een beoordelaar registreert dan vervolgens volgens zijn verwachting meer in een bepaalde richting (ziet bv meer stresssignalen of houdingverlagingen als zijn eerste indruk is: wat een stresskip) dan bij een hond waarvan hij een andere eerste indruk heeft (lekker stabiel) van het een of het ander. Maar het kan ook zijn dat minder gewenst gedrag aan het einde van de test een beoordelaar zo beïnvloedt dat het zijn beeld van eerder (wel gewenst) gedrag van de hond in de test vertekent. Daarom is het beter de eindscore ter plekke door meerdere (minimaal twee) te laten doen die tot een consensus moeten komen. Ook het terugkijken van de video enkele dagen later zou een vereiste moeten zijn, zeker wanneer een hond in de test veel ongewenst gedrag heeft laten zien, omdat dit een minder emotionele beleving voor de beoordelaar is dan de ter plekke situatie waar de eigenaar/fokker aanwezig is die een slecht nieuws bericht krijgt. Het later terugkijken zorgt dan vaak voor een bijstelling van de eerste emotioneel geladen beoordeling. Dit later terugkijken is ook in een ander opzicht zinnig, omdat uit onderzoek is gebleken dat in het geheugen van de beoordelaar bij een beoordeling ter plekke, ook bij een niet emotioneel geladen beoordeling, relatieve accenten worden gelegd. Een beoordeling door derden op basis van de video kan echter naar mijn mening nooit de beoordeling ter plekke helemaal vervangen, omdat allerlei relevante informatie uit de omgeving, die invloed hebben op de test, buiten beeld van de video gebeurt. De meest zuivere beoordeling is dus een combinatie van de twee.