Behavioural testing for aggression in the domestic dog.

Doreen J. U. Planta[1)], Willem J. Netto.
1) Raad van Beheer op Kynologisch Gebied in Nederland, Emmalaan 16-18, P.O. Box 75901, 1070 AX Amsterdam, The Netherlands.

Abstract
This study reports on the development of a behavioural test that detects attack behaviour in dogs. Validation took place by comparing the attacking behaviour during the test with the reported biting history of the dogs. The results show that the test is a valid instrument to detect attacking behaviour towards humans and towards dogs.
The predictability was measured by retesting 37 dogs. The comparison between test and retest shows a highly significant correlation. The resemblance between the scores for attacking behaviour in test and retest is significant for most subtests.
The results from the validation and the predictability show that the developed test is a valid and predictable instrument and therefore a useful tool to detect attacking behaviour of dogs towards humans and towards other dogs.

Introduction
Because of the concern about biting incidents in the Netherlands, the Ministry of Agriculture initiated research into the development of a behavioural test to detect aggressive behaviour. The aim of this test is to select dogs for their tendency to react aggressively to stimuli and to exclude these aggressive dogs from breeding. Secondly, such a behavioural test can also be used as a tool to predict to what extent a dog is a danger for humans and animals.
In the literature different kinds of aggression are classified according to the contexts in which biting occurs (Borchelt, 1983): dominance aggression, fear aggression, possessive aggression, protective aggression, territorial aggression, predatory aggression, sexual aggression, maternal aggression, irritable aggression, pain induced aggression and punishment induced aggression. Other forms of aggression that could be added are frustration aggression and redirecting aggression.
Other situations in which many bite-incidents occur are play, as well as petting or handling the dog (van der Wijk & Klasen, 1981).
Hart (1976) describes a form of idiopathic aggression, which is a form of aggression that is not provoked, but is unpredictable and unexplainable. Dogs that show this kind of aggression are often very friendly and loveable pets for the rest of the time. Idiopathic aggression is more often seen with the popular breeds these days. These breeds are, because of their popularity, indiscriminately bred (Beaver, 1980). The expectation is that the latter form of aggression can not be provoked in a behavioural test, because of its unpredictability.
Aggressive behaviour has multiple causation (Lockwood & Rindy, 1987). Genetic predisposition, bad early socialisation to people or dogs, the lack of training for obedience or the training for fighting, the bad quality of care taking and supervision provided by the owner and the behaviour of the victim may all be causal factors.
Important here is the fact that aggression has a genetic component. Selection for protective functions has resulted in different breeds that tend to show aggressive behaviour more easily than other breeds (Beaver, 1981). So there is little doubt that at least some of the different kinds of

aggression are influenced by genetics (Polsky, 1984). Selection in the breeding program is therefore a means to decrease aggressive behaviour. A behavioural test to be used as a selection tool should therefore be developed. The purpose of this study is to report on the development of such a behavioural test and to describe this test. The test should measure the threshold of which dogs can be provoked to show attack behaviour.

Materials and Methods
Test procedures
The studies were performed in a kennel at the university. During testing, each dog was tied to the wall with a rope, except for subtests in which the owner played which his dog. Different areas were used where the owner sat next to his dog or stood next to his dog. During tests where the owner sat next to the dog the distance between owner and dog was far enough that no redirection aggression could occur. The basket of the dog was put in one area to make it more comfortable to the dog and to create the opportunity for the dog to retreat to something familiar. On the floor a line was painted for security reasons. The person who performed the subtests thus knew how far he or she could approach the dog safely. An artificial arm/hand was used during all subtests in which the dog had to be touched.
Before the final behavioural test was developed performed two pilot studies were performed to try out subtests and the order of testing.
All subtests lasted between 20 -30 seconds. Exceptions were the subtests where the owner played with his dog (duration: 1 minute).
To measure the predictability of the final behavioural test 37 dogs were retested after six months.

Scoring method.
A 5 point Gutmann-scale was used (Devellis, 1991). The higher the score, the more intense the aggression.
The aggression-scale looks as follows:
1. No aggression
2: The dog growls and/or barks
3. The dogs shows his teeth, possible with growling and/or barking
4. The dog makes an unfinished bite, possible with growling and/or barking and/or showing his teeth
5. The dog bites or attacks.

The highest score reached by the dog is noted. If for example, a dog first growled and after that attacked, this dog got score 5.

Results
The first pilot study contained of 80 subtests and took 1.5h to perform. 89 dogs were tested. The dogs were divided into three groups for the purpose of validation. Non-biters: dogs that had not yet bitten (N=34). Biters: dogs that had bitten dogs or humans (N=27). Trained- biters: dogs that were trained to bite. (N=28)
The biting history was compared to the corresponding biting behaviour during the test.
The frequencies of attack by the three categories of dogs differ significantly (Kruskall-Wallis: $p=0.03$). The non-biters showed the least attack behaviour. Non-biters differ significantly from trained dogs (Multiple Comparison Test: $p<0.05$). Other differences are not significant. The

conclusion is that the test detects attack behaviour more often in those dogs that were trained to bite. This test also elicits and detects more attack behaviour from dogs that had previously bitten, although the difference with non-biters is not significantly. Therefore, this behavioural test is only partly validated.

The second pilot study included 38 subtests and took 45 minutes. Most of the subtests were performed outside. After 28 dogs tested a substantial lower level of aggression was reached than during the first pilot study. Therefore, the order of performance of the subtests was switched in such a way that in the second series a number of subtests were first performed indoors and then continued outdoors. The second pilot study therefore consisted of two different test series. 86 dogs were tested that were distinguished into 4 categories. Non-biters (N= 31), dog-biters (N=24), human- biters (of some had also bitten dogs) (N=24) and trained biters (N= 7). The first test series was performed on 15 non-biters, 8 dog biters, 4 human biters and 1 trained biter. The second test series was performed on 16 non- biters, 16 dog-biters, 20 human-biters and 6 trained-biters.
In order to measure the effect of the two different test series, used during the second pilot study, the total frequencies of attack behaviour of all dogs in the first test series was compared with those of all dogs in the second test series. The comparison showed a significantly higher attack behaviour in the second test series (Mann Whitney U test: p=0.0004). One explanation for this could be differences between the populations of dogs being tested. However, there was an evenly distribution over the two test series of dogs belonging to the 4 groups (X^2: p>0.05). Therefore, it's more likely that the explanation is that the dogs reach a higher level of aggression in the second test series than in the first test series. Subtests performed inside cause a stronger increase of the aggression motivation that still plays a role in the behaviour of the dog when tested outside later on. Despite the fact that the dogs had the possibility to recuperate during subtests performed outside, we find that with the increase of aggression motivation caused by the subtests performed inside, more dogs also showed aggressive behaviour outside. Subtests should therefore be performed inside.
The biting history of the dogs was compared to attack behaviour during the test. Trained biters were left out of this comparison.

Table 1: median attack behaviour (corrected for differences in number of subtests) per group of test series 1

	median attack behaviour
Non- biters	0%
Dog- biters	0%
Human (and sometimes dog biters)	3.45%

The frequencies of attack behaviour differ significantly over the three groups (Kruskall-Wallis: p=0.02). However, the multiple comparison test shows no significant differences between the groups. Therefore this behavioural test is not validated.

Table 2: median attack behaviour (corrected for differences in number of subtests) per group of test series 2

	median attack behaviour
Non- biters	0%
Dog- biters	2.7%
Human (and sometimes dog biters)	2.7%

The frequencies of attack behaviour differ significantly over the three groups (Kruskall-Wallis: p=0.01). The multiple comparison test shows a significant difference between non biters and dogs that had bitten humans and sometimes dogs. This behavioural test is therefore a valid tool to predict attack behaviour towards humans, but not towards dogs. There is no significant correlation between test number and number of dogs showing attack behaviour. This means that there is no increase in aggressive behaviour during the test.

The final behavioural test was developed as a result of the two pilot studies. It contained 43 subtests and took approximately 50 minutes. Most of the subtests were performed indoors, and there was little time between the subtests, so the opportunity for recuperation is small. Furthermore the more subtests the more dogs are motivated to show aggressive behaviour. The order of the subtests was performed in such a way, that an increase in aggressive behaviour was likely to occur in the course of the test procedure. The dogs (N=114) were divided into three groups on the basis of previous behavioural experience. Non- biters (N=40), dog- biters (N=51) and human- biters of which some had also bitten dogs (N=23).
The attack behaviour per subtest as performed during this test was measured. The results are shown in figure 1. It shows an increase in number of dogs that show attack behaviour during the test. There is a significant correlation between test identification number and number of dogs that show attack behaviour (Spearman Rank Correlation Coefficient 0.51; p=0.001), whereas in the second pilot study there is no significant correlation between these two parameters. Thus, keeping subtests that do not evoke aggressive behaviour tends to increase the number of dogs that will eventually show attack behaviour. Tests 2, 3 and 4 elicit no attack behaviour at all. Nevertheless, these subtests are necessary because they show the degree of obedience of the dog towards its owner. Obedience may influence attack behaviour of the dog during several other subtests. If obedience is high, it is conceivable that the dog shows little attack behaviour in presence of the owner, whereas in absence of the owner the behaviour of the dog can be completely different. Tests 7, 8 and 9 give the dog the opportunity to habituate to the testing area. Test 10 is used to verify that the dog hasn't been given any sedatives. Test 1 was performed first, because people totally unknown to the dog should do the approach of the car.

Validation
The biting history of the dogs was compared to attack behaviour during the test. Furthermore a comparison was made between non- biters and dog-biters during the 11 subtests involving stimulus dogs. The results are given in figure 2. There is a significant difference (Mann-Whitney U test: p= 0.0448) between the two groups of dogs. The dogs that attacked other dogs previously attacked more frequently in those subtests than do the non- biters. Therefore this test is validated for measuring a tendency to attack dogs.
The result of the comparison over 30 subtests where humans was used between non- biters and human biters is given in figure 3. In this comparison we did not take into account subtest 24 and 29, because the measured attack behaviour could not be performed against a human being. Again

there is a significant difference between the non- biters and the human- biters (Mann-Whitney U test: p=0.0034). Dogs that previously have bitten humans and possible dogs show more attack behaviour than the non-biters and the test is therefore also validated for attack behaviour against humans.

Predictability

To measure the predictability of the behavioural test 37 dogs were retested after six months. The results of test and retest were compared and are shown in figure 4. Since for 10 of the 37 dogs the number of subtests they went through during test and retest were different, these dogs were not taken into account. The total frequency of the attack behaviour per dog in the test is plotted against the total frequency of the attack behaviour per dog in the retest. The attack behaviour during test and retest was highly correlated, therefore the predictability of the test is considered to be high (Spearman Rank Correlation Coefficient = 0.82; p<0.0001; explained variance is 43%).

The predictability per subtest

The predictability per subtest has been measured in two ways. First the number of attacks per subtest that occurred in the test and in the retest is compared. In table 3 this is expressed as the percentage of agreement for attack behaviour between test and retest. Secondly the Kappa-coefficient (Martin & Bateson, 1986), which corrects for similarities based on coincidence, is calculated. For the calculation of the Kappa-coefficients and its significance values (Bishop et.al., 1975) the MATMAN-program (De Vries, et.al., 1993) is used.

Table 3: Predictability per subtest, expressed in the percentage (%) of agreement (corrected for number of dogs), and the Kappa- coefficient and its significance. The owner is absent in test 1, 6, 20, 21, 22, 23, 24, 25, 26, 27, and 28. The owner sat next to his dog during test 10 till 14 and test 17 till 28, and 35 till 43. The owner stood next to his dog during test 30 till 34.

Subtest	**% agreement**	**Kappa-coefficient**	**Subtest**	**% agreeme**	**Kappa-coefficient**
1. Approach car	81.3%	0.46; p<0.0(	23. old woman	88.9%	0.28
2. Walk on lead (by owner)	100%	0	24. doll pulled	97.2%	0
3. Command 'sit' (by owner)	100%	0	25.doll in front of the dog	88.9%	-0.05
4. Command 'dov (by owner)	97.1%	0	26. stare at dog	94.4%	0.72; P<0.001
5. Encounter thre barking d(outside	91.9%	0.72; P<0.0(	27. clapping	77.8%	0.42; P<0.05
6. encounter one barking dog	78.4%	0.54; P<0.0(	28. yelling at dog	88.6%	0.60; P<0.001
7. command 'com (by owner)	100%	0	29. artificial human	100%	0
8. play (by owner	100%	0	30. slow approach	94.6%	0

9. play with unfamiliar object (by owner)	100%	0	31. fast approach	91.9%	0.62; P<0.001
10. lamp	100%	0	32. groom	94.4%	0.47
11. tug of war	94.6%	0	33. threat of owner	74.3%	0.25
12. stroke dog	97.3%	0	34. three barking dogs	75.0%	0.52; P<0.001
13. stare at dog (by owner)	100%	0	35. male stimulus dog	75.0%	0.53; P<0.001
14. holds muzzle (by owner)	94.6%	0.47	36. stroking male stimulus dog	85.3%	0.68; P<0.001
15. pulls dog on i back (by owner)	97.2%	0	37. feeding bowl in presence of male stimulus dog	68.6%	0.37; P<0.05
16. squeezes groi (by owner)	88.6%	0.53; P<0.0(	38. feeding bowl to male stimulus dog	76.5%	0.53; P<0.001
17. pulls away food bowl	89.2%	-0.04	39. female stimulus dog	97.2%	0.79; P<0.001
18. pulls away food bowl (owner)	94.4%	-0.03	40. stroking female stimulus dog	94.6%	0.47
19. shaking hand:	94.6%	-0.03	41. feeding bowl in presence of male stimulus dog	91.8%	0.37
20. rattle	86.5%	0.23	42. feeding bowl to male stimulus dog	86.1%	0.37
21. umbrella	89.2%	0.28	43. stroke the dog	97.2%	0.66; P<0.05
22. plastic bag	94.4%	0.47			

Table 3 shows that most subtests have a high percentage of agreement and a significant kappa-coefficient, so there is a significant resemblance between test and retest. Some subtests show 100% agreement and a kappa- coefficient of 0. These outcomes emerge if one column or a row in the matrix is zero. In these cases the kappa-coefficient cannot be defined.

Those subtests in which stimulus dogs were used (tests 5 and 6 and tests 34 till 42) have a lower predictability than most other subtests (Mann Whitney U test: $p < 0.05$). A reason for this can be the fact that the reactions of the stimulus dogs changed during the study. Since we tried to prevent that the stimulus dogs lost from the tested dogs, the male stimulus dog (tests 35-38) and the three dogs on the leash (test 34) became more aggressively during the study. Most of the tested dogs reacted more aggressively too. The stimulus bitch on the other hand (tests 39 till 42) behaved less aggressively. A number of dogs reacted more aggressively towards her and some less aggressively. The kappa-coefficient is nonetheless significant for those subtests, in which the stimulus male dog was used and for the first subtest (39) in which the stimulus bitch was used.

Discussion

In this paper a devised and tried out behavioural test is proposed. In doing so it is tried to cover as many aggression evoking contexts as possible. The situations in which the dogs were confronted were obviously threatening enough for dogs to react. We concentrated on actual attacking behaviour rather than threatening behaviour. It remains uncertain to what extent threatening behaviour reliably predicts an actual attack. Although a successful threat may lead to attack behaviour in the future, because the dog may learn from this situation (Hart, 1976).

One important result we obtained is that training for attacking behaviour seems to lead to a decrease in the biting threshold. From the results of the first behavioural test one can conclude that those dogs that are trained to attack show even more attack behaviour than those dogs that have a 'bite' history. This suggests that training for this behaviour lowers the threshold to attack, also outside the training situations. Hart described the same phenomena in 1976. He named this form of aggression 'learned' aggression and pointed out that the dogs learn, by shaping, severe forms of aggression. Since their attacks are successful these dogs feel more secure in their attacks. Training can also alter the occurrence of behaviour. Dogs which are trained to bite, bark to the person that they attack (Ochsenbein, 1987).

There can be different reasons for the decrease of attacking behaviour in the second behavioural test compared to the first behavioural test. The first reason could be the dogs that were tested. In the first behavioural test a lot of trained biters were tested, and they account for a lot of attacking behaviour in the first behavioural test. The second reason could be the fact that most subtest from the second behavioural test were performed outside instead of inside. The third reason could be the smaller number of subtests in the second behavioural test. An additional reason to device a more time consuming behavioural test that included many subtests is that this makes it more difficult for dog owners to train their dogs in order to pass the test.

CONCLUSIONS

It's possible to develop a behavioural test that detects attacking behaviour in dogs. The final behavioural test developed here is validated by establishing correlation's between the actual attack behaviour and the a priori division into groups based on attacking behaviour as reported by the owner. The stimuli that are attacked more often in the test, namely dogs or humans, are significantly related to the history of the dog. From these results it can be concluded that the contexts created in the test are representative of the situations in which biting occurs. The predictability of the test is measured by retesting 37 dogs. The comparison between test and retest shows a significant correlation. The resemblance between the scores for attacking behaviour in test and retest is significant for most subtest.
The developed test is therefore a valid and reliable tool to detect attacking behaviour towards dogs and humans.

References

- Beaver, B. V., 1980: Mental Lapse Aggression Syndrome. J. Am. Anim. Hosp. Assoc. 16; 937-939.
- Beaver, B. V., 1981. The genetics of canine behaviour. Vet. Med., Small animal Clin., 10;1423-1424.
- Borchelt, P. L., 1983: Aggressive behavior of dogs kept as companion animals: Classification and influence of sex, reproductive status and breed. Appl. Anim. Ethol. 10; 45-61.
- DeVellis, R. F.,1991: Scale Development. Theory and Applications. Applied Social Research Methods Series, 26, Sage Publications; 121 pp.
- Hart, B. L., 1976: Canine behaviour. Canine practice 2; 10-14.
- Lockwood, R., Rindy, K., 1987: Are "pit bulls" different? An analysis of the pit bull terrier controversy. Anthrozoös, 1(1); 2-8.
- Ochsenbein, U., 1987: Beissverhalten von Hunden unter verschiedenen Bedingungen. Artgemassen Tierhaltung; 135-140.
- Martin, P., Bateson, P., 1986: Measuring behaviour, Cambridge University Press; 86-97.
- Polsky, R. H., 1984: Factors Influencing Aggressive Behaviour in Dogs. Aust.vet. Practit. 14(2); 66-69.
- Vries, H, de, Netto, W. J., Hanegraaf, P. L. H.; 1993: Matman: a program for the analysis of sociometric matrices and behavioural transition matrices. Behaviour, 125; 157-175.
- Wijk, J., van der, Klasen, H. J., 1981: Hondebeten en de omstandigheden waaronder ze vóórkomen. Ned. T. Geneesk. 125(4); 133-135.

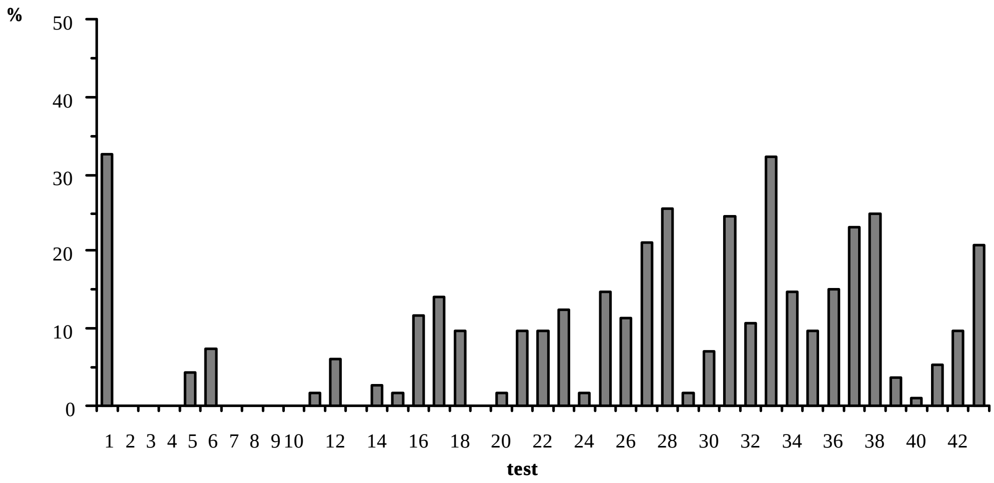

Fig. 1: number of dogs (in percentage of total) that show attack behaviour per test

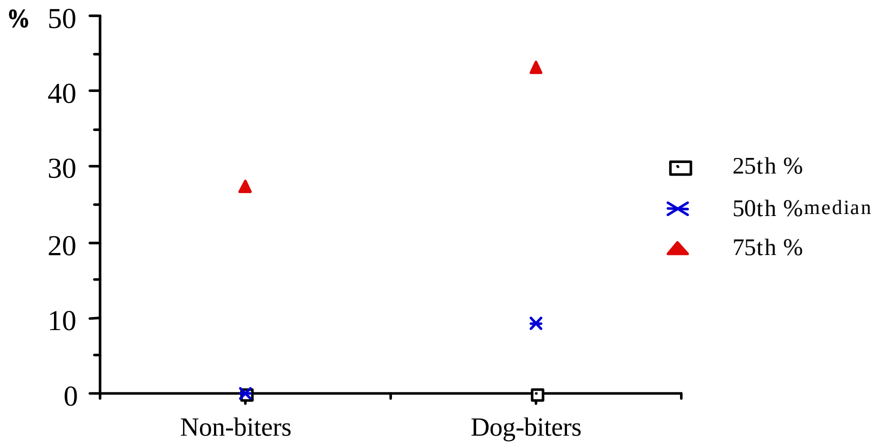

Fig. 2: percentage of attack behaviour of non- biters and dog- biters.

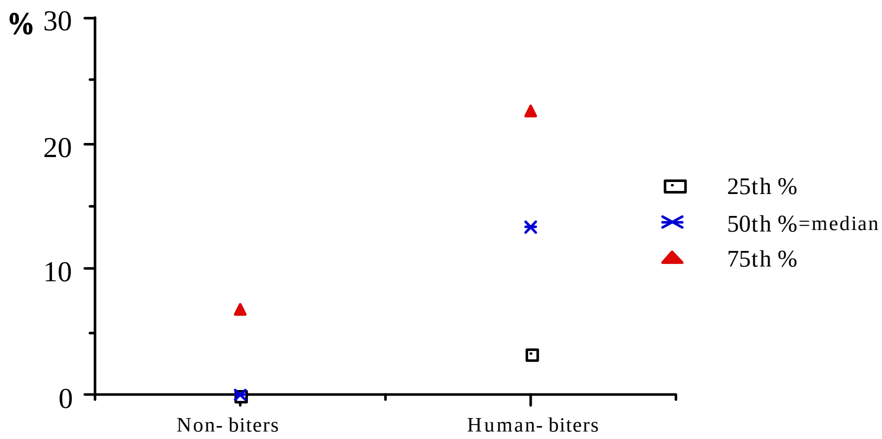

Fig. 3: percentage of attack behaviour of non- biters and human- biters.

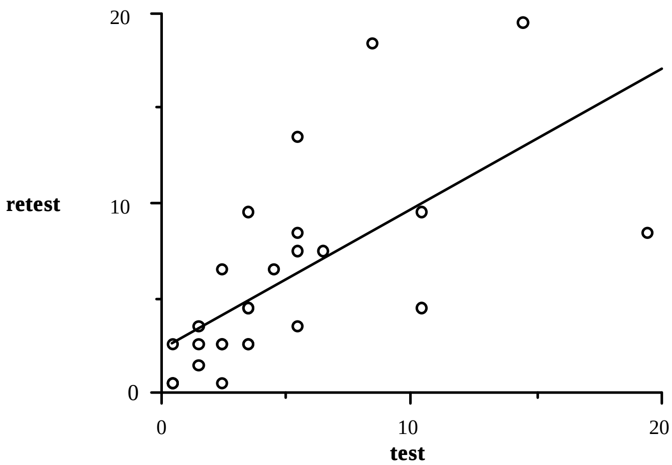

Fig. 4: Regression analysis of the total frequencies of the attack behaviour of 27 dogs over test and retest. The dots
represent the individual dogs. (r^2=0.43)