ANTHROZOÖS VOLUME 25, SUPPLEMENT PP. S139–S151

Construct Validity of Animal-Assisted Therapy and Activities: How Important Is the Animal in AAT?

Lori Marino
Department of Psychology, Emory University, USA

Address for correspondence: Lori Marino, Department of Psychology, Psychology and Interdisciplinary Sciences Building, Suite 488, 36 Eagle Row, Emory University, Atlanta, GA 30322, USA. E-mail: lmarino@emory.edu

ABSTRACT Animal-assisted therapy and animal-assisted activities involve a nonhuman animal as a key therapeutic agent in some kind of intervention that may range from highly specified, as in AAT, to more casual, as in AAA. In this review I address the question: How important is the *animal* in animal therapy? In other words, does the recent literature strongly support the notion that a live animal, as opposed to another novel stimulating component, is specifically necessary for therapeutic success. Two meta-analyses and 28 single empirical studies were reviewed in order to address this issue. I conclude that the effects of AAT and AAA are likely to be moderate and broad at best and that, although improving, the literature has not yet reached an experimentally rigorous enough level to provide a definitive robust conclusion about the effectiveness of these approaches, particularly with regard to the question of whether a live animal is necessary for a therapeutic effect.

Keywords: AAA, AAT, animal-assisted therapy, methods, validity

Animal-assisted activities (AAA) and animal-assisted therapy (AAT) are becoming increasingly popular adjuncts for more traditional forms of therapy in a variety of settings, such as schools, hospitals, and psychiatric centers. The range of nonhuman animals (hereafter referred to as "animals") used and the types of maladies targeted are expansive. Whereas domesticated animals, particularly dogs and horses, are the most popular choice (Fine 2010), there are many facilities and programs that offer "therapy" with less commonly used animals such as fish (Cole and Galinski 1995, 2000) and birds (Davis 1988), as well as exotic animals such as dolphins (Marino and Lilienfeld 1998, 2007) and even elephants (Swanepoel and Odendaal 2005). Disorders targeted range from generalized psychological problems to depression and anxiety to developmental disabilities and infectious disease (Fine 2010).

Various terms have been used to describe AAA and AAT such as pet therapy, pet-facilitated therapy, pet-assisted therapy, and animal-facilitated therapy (Connor and Miller 2000). Pet Partners (formerly Delta Society), a leading international, non-profit organization that provides training for AAA and AAT practice, defines AAA and AAT in the following manner:

> AAA provide opportunities for motivational, educational, recreational, and/or therapeutic benefits to enhance quality of life. AAA is delivered in a variety of environments by specially trained professionals, paraprofessionals, and/or volunteers, in association with animals that meet specific criteria. ("Standards of Practice" 1996)
>
> AAT is a goal-directed intervention in which an animal that meets specific criteria is an integral part of the treatment process. AAT is directed and/or delivered by a health/human service provider working within the scope of practice of his/her profession. ("Standards of Practice" 1996)

AAT is distinguishable from AAA in that it has a specific therapeutic objective and is delivered as part of a highly specific treatment plan. AAA, on the other hand, involves a more casual interaction between animals and humans and can be provided by a broader range of individuals than AAT. Although Pet Partners distinguishes AAA and AAT, in actual practice the two are often conflated and may overlap considerably. In AAT and AAA the animal serves a variety of roles from companionship to caregiver to social facilitator and reinforcement; studies often do not provide enough information about the way the animal was used to differentiate between AAT and AAA. Despite this ambiguity, Pet Partners is clear about the central point that although often used as an adjunct to other forms of therapy that include the presence of a human, AAT and AAA are based on the premise that the animal makes a therapeutic difference to the subject. The purported mechanisms of action include mental stimulation, physical stimulation, physiological effects, and general motivating effects. Therefore, according to Pet Partner's claims, the effects of an animal in the presence of a human on the subject should be differentiable from the effects of just a human on the subject and differentiable from the effects of another novel stimulus and a human.

The implication of animal-assisted interventions of all kinds is that the live animal is a highly specific component of the therapy. That is, without the animal there is no "animal-assisted" activity of any sort. The entire industry is founded on the premise that involvement of a live animal in any kind of activity or therapeutic intervention (no matter how complex or multidimensional) is critical to its identity. Therefore, it is important to determine efficacy and specificity, that is, whether AAT and or AAA result in real therapeutic improvement and, if so, whether the improvement is due to the presence of the animal specifically. Hence, the focus of this review and evaluation is to address the question: How important is the *animal* in animal-assisted therapy?

There is a longstanding general sense among practitioners and researchers that activities and interventions involving animals may have nonspecific positive effects on both patients and normative populations of subjects. Yet, there have also been lingering doubts and controversies about the robustness and specificity of the effect of animals on human health and welfare. In other words, there remain questions about the validity of AAT and AAA. In a 2002 review of the field, Johnson, Odendaal and Meadows recognized the importance of obtaining empirical evidence for AAT and AAA and stated: "Advocates of programs using animals have begun to see empirical documentation as a critical factor in widespread acceptance of animal-assisted activity (visitation programs) and animal-assisted therapy (programs with specific goals for individuals) as beneficial interventions for patients" (pp. 422–423). They cautioned that " …

better designed experiments are needed to enhance the credibility of animal-assisted interventions …" (p. 438). Since then there have been many published peer-reviewed studies of AAT and AAA buttressing claims of therapeutic efficacy for these and related activities. However, it is critical to move from collecting basic data on efficacy to refining our experimental methods in order to better understand the nature of these findings. In other words, we currently need to probe the various components of validity in AAT and AAA. More recently, in a review of the effects of pets on human health, Herzog (2011) stated: "Design problems are common in studies of human–animal interactions" and " … it is often difficult or impossible to eliminate placebo effects via traditional methods such as single- and double-blind experimental and control groups" (p. 238). Although Herzog's comments were made specifically about studies of the health effects of pets, they are relevant to the current literature on AAT and AAA. In order to explore the current status of AAT and AAA, I will examine two meta-analyses as well as individual studies published in peer-reviewed journals since 1995 and will focus upon issues of validity and on construct validity in particular.

Review of Meta-Analytic Studies

In 2007, two major meta-analytic reviews of animal-assisted therapy were published. One, by Nimer and Lundahl (2007), was a comprehensive evaluation of AAT studies from 1973 to 2004, yielding 49 studies that met criteria for inclusion in the meta-analysis. The review covered a wide range of animals, disorders, therapeutic settings, client age categories, and outcome measures. The other, Souter and Miller (2007), consisted of an initial review of 60 studies with only five meeting the inclusion criteria for meta-analysis. Moreover, Souter and Miller focused exclusively on the effect of AAT/AAA on depression.

Taken together, both meta-analyses found moderate positive effects for a range of disorders. Nimer and Lundahl did not exclude studies lacking a control group but they found similar results for uncontrolled and controlled studies in their meta-analysis. However, they noted that because AAT is almost always used as an adjuvant to standard therapies it is very difficult to gain a "universal understanding" (Nimer and Lundahl, p. 235) of exactly what AAT is and what components of it might be effective.

Nimer and Lundahl (2007) and Souter and Miller (2007) showed that many peer-reviewed and published AAT and AAA studies do not meet minimal standards of research design and therefore cannot be employed in a quantitative meta-analysis. Nimer and Lundahl included studies that lacked a proper control group and still found that only 49 of 250 published studies were eligible for inclusion on the basis of simple criteria such as adequate sample size and enough data to calculate effect sizes. Although finding moderate positive effects for depression, Souter and Miller reported that only one study employed AAT strictly, precluding them from differentiating the effects of AAT from AAA and other general procedures involving animals. Moreover, only one study controlled for the role of the human in the animal condition. That is, only one study of the five included a "human only" control group to be compared with the "human and animal" group. Therefore, the role of the animal in the therapies was not investigated in any systematic way.

Therefore, to summarize the two 2007 meta-analyses discussed above, there exist numerous unresolved methodological issues in the AAT and AAA literature. Foremost among these is the lack of research specifically addressing the degree to which the purported positive effects of AAA and AAT are attributable to contact with the human facilitating the animal interaction. Most studies do not include a way to differentiate the animal from the human in AAT and AAA.

Specific Studies

In order to accomplish a more thorough and current evaluation of the status of AAT and AAA, I conducted qualitative methodological analyses of individual empirical studies published from 2005 (where Nimer and Lundahl left off) to the present that were also not included in either meta-analytic study described above. First, I conducted computer searches for peer-reviewed articles from the years 2005 to the present of three of the major journals in AAT and AAA: *Anthrozoös, Applied Animal Behaviour Science, and Society & Animals*. Second, I searched Google and Google Scholar under the terms "animal assisted therapy," "animal therapy," and "animal assisted activities" for studies published in other journals from 2005 to the present. Third, I obtained all empirical studies from peer-reviewed sources from 2005 to the present listed on the Pet Partners website (www.deltasociety.org). Finally, I searched the citation list of each study obtained by the above procedures in order to find additional empirical papers published between 2005 and the present that I may have overlooked in the initial computer search. This led to including 28 empirical studies in my analyses. Although studies on equine-assisted therapy and activities were included, I did not include studies on therapeutic riding or hippotherapy, since these are differentiable from AAT and AAA. I also excluded studies on such issues as the effects of pets in the home and the effects of service-animals, as these are only tangentially related to AAT and AAA.

I assessed the validity of each study according to minimal standard methodological criteria put forth by four sources: Cook and Campbell (1979), Shadish, Cook and Campbell (2002), Kendall and Norton-Ford (1982), and Shaughnessy and Zechmeister (1994). These sources describe a set of threats to experimental validity that should be avoided in experimental research. These include: placebo effects, novelty effects, construct confounding, demand characteristics, and experimenter expectancy effects. The presence of even one major threat to validity can render a study's findings difficult, or in some cases even impossible, to interpret. Additionally, I paid close attention to those studies that did not include a control, in order to address the critical question of whether AAT and AAA can distinguish the roles of the animal versus the human in studies that claim to find a positive therapeutic effect.

Table 1 displays definitions for several important threats to construct validity that were used to evaluate the 28 studies as well as the methodological controls that would minimize those threats. There are a number of dimensions of experimental validity that are critical when evaluating evidence. These include internal validity, that is, the methodological soundness of the measures and, particularly, the ability to infer a causal relationship between two variables, and external validity, that is, the generalizability of these causal inferences from the study sample to the population, respectively. In the present analysis I will focus on a particularly relevant domain of experimental validity, that is, construct validity. Construct validity has to do with whether a scale or instrument measures or correlates with the theorized scientific *construct* that it purports to measure. In the case of AAT and AAA, the construct is the therapeutic value of the animal in animal-assisted therapy or activities.

Most threats to construct validity are generated by nonspecific effects, that is, improvement from effects not specific to the intended treatment. Often these generic effects are shared with many other interventions. Two of the most common types of nonspecific effects are placebo effects and novelty effects. The little-understood placebo effect is well documented and derives from the expectation of improvement (Linde, Fassler and Meissler 2011). AAT and AAA are potentially vulnerable to placebo effects because the nature of the treatment is often evident to the subjects. Related to this situation are demand characteristics, when subjects come to

Table 1. Main threats to construct validity assessed in each of the 28 studies reviewed and methodological controls for (reductions to) them.

Validity Threat	Definition	Possible Controls
Placebo	Improvement from expectation of improvement	Minimization of cues to treatment or hypothesis; deception; sham treatment
Novelty	The effects of energy, excitement and enthusiasm not specific to the intended treatment	Exposure to another novel stimulus; long-term follow-up assessment
Construct Confounding	Failure to take into account the fact that the procedure may include more than one active component	Dismantling procedure
Demand Characteristics	The tendency of participants to alter their responses in accord with their suspicions about the research hypothesis	Minimization of cues to treatment or hypothesis; sham treatment
Experimenter Expectancy Effects	The tendency of the experimenter to unintentionally bias the results in accordance with the hypothesis	Blind rating of subjects

recognize the experimenter's hypothesis and alter their responses in accordance with it. Novelty effects are the general energizing effects of a new, often exciting experience. Because AAT and AAA involve exposure to an often lively, interactive, and engaging animal (and one subjects may not have much experience with), they are particularly prone to novelty effects. Placebo effects and demand characteristics are typically minimized by a blind study that conceals clues about assignment and hypothesis. But these procedures are not easily implemented when the treatment is so apparent and difficult to replicate without the therapeutic component, as in AAT and AAA. They can also be reduced or accounted for with the use of a manipulation check, such as a sham treatment, or by active deception. The proper control for novelty is exposure of the control group to another similar new and exciting stimulus while keeping all else equal or taking follow-up measurements when the excitement of the experimental condition has diminished. Finally, experimenter expectancy effects are due to the tendency for the experimenter to unintentionally bias the results of the study in accordance with his or her hypothesis. These kinds of effects are controlled by employing outcome measures that can be measured by raters blind to condition.

Construct confounding occurs when there is failure to take into account that the experimental procedure may involve more than one active component. In AAT and AAA the experimental treatment typically involves a complex mixture of components in addition to the animal per se, such as increased attention from a therapist, being presented with a novel situation, increased physical activity, and many other aspects of the situation. Moreover, the animal is a complex stimulus that has potentially a number of therapeutic features, such as being soft to the touch or making eye contact. In general, construct confounding is typically minimized or eliminated by dismantling procedures (Kazdin 1994), which require the experimental and control group(s) be exposed to the same or highly similar procedures and stimuli with only the key component as the differential treatment component between groups.

In the present analysis I will briefly identify those studies which will not be considered further due to having not met minimal requirements for methodological soundness and, therefore, the results are uninterpretable at best. I will focus on those who meet these minimal criteria instead.

Table 2. Details of the 28 studies reviewed: subject population and sample size, measured construct(s), design, and findings.

Study	Subject Population (n)	Measured Construct(s)	Design	Findings
Anderson and Olson (2006)	Children with emotional disorders (6)	Emotional stability	Case study	Qualitative improvement in emotional stability
Banks and Banks (2005)	Elderly in nursing home (33)	Loneliness	Pre-post repeated measures (individual vs. group AAT), no control	Significant decrease in loneliness scores in individual AAT group only
Banks, Willoughby and Banks (2008)	Elderly in nursing home (13)	Loneliness	Pre-post; two control groups (robotic dog and no intervention), randomized	Significant decrease in loneliness score in both live and robotic dog conditions
Bergret, Ekeberg and Braastad (2008)	Adults with psychiatric disorders (60)	Self-efficacy, coping ability, quality of life	Pre-during-post, randomized, between subjects	Significant increase in self-efficacy and coping in experimental group
Braun et al. (2009)	Children with chronic pain (18)	Pain level, vital signs	Quasi-experimental, between subjects	Significant decrease in pain and significant increase in respiration rate in experimental group; no change in pulse or blood pressure
Chu et al. (2009)	Adults with schizophrenia (15)	Self-esteem, self-control, other psychological factors	Randomized pre-post, between subjects	Significant increase in self-esteem, self-control, improvement in other emotional symptoms in experimental group
Cole and Gawlinski (2007)	Patients hospital ized with heart failure (26)	Hormonal state, anxiety, vital signs	Three group (volunteer and dog, volunteer, usual care) randomized repeated measures	Significantly greater decrease in systolic and capillary pressure, epinephrine and norepinephrine levels, and state anxiety in experimental group
Colombo et al. (2006)	Institutionalized elderly (48)	Mental state, subjective perception of quality of life, psychopathological symptoms	Pre-post, between subects with two control groups (plant, no intervention)	Significant increase in mood scores in experimental group
Estevez and Stokes (2008)	Children with developmental disabilities (3)	Social behaviors	Single case with repeated measures and replicated effects, only descriptive statistics	Increase in positive behaviors, decrease in negative behaviors, and improved social responsiveness
Fournier, Geller and Fourtney (2007)	Prison inmates (24)	Treatment progress, social skills	Quasi-experimental, pre-post mixed, repeated measures	Significant improvement in treatment progress and social skills in experimental group
Gee, Church and Altobelli (2010)	Developmentally delayed and typical pre-schoolers (12)	Cognition (object categorization task)	Randomized, mixed with three groups (humans and live dog, human and stuffed dog, human only)	Significant decrease in irrelevant choices on task in experimental group
Gee, Harris and Johnson (2007)	Developmentally delayed and typical pre-schoolers (14)	Gross motor skills	Randomized, mixed	Significantly faster completion, both significant increases and decreases in accuracy depending on task, in experimental group

Study	Subject Population (n)	Measured Construct(s)	Design	Findings
Hoffman et al. (2009)	Adults with major depression (12)	State anxiety	Pre-post, crossover	Significant decrease in state anxiety in experimental group
Johnson et al. (2008)	Adults with cancer undergoing radiation therapy (10)	Mood, fatigue, self-perceived health, sense of coherence	Pre-post, mixed, two control groups (human, quiet reading)	No significant differences across groups
Klontz et al. (2007)	Unspecified (31)	Psychological distress and well-being	Pre-post, no control group	Significant decrease in psychological distress, significant increase in well-being
Kovacs et al. (2006)	Adults with schizophrenia (3)	Nonverbal communication	Pre-post, no control group	Some changes in activity but only descriptive results
Krskova, Talarovicova and Olexova (2010)	Children with autism (9)	Frequency of social contacts	Repeated measures	Significant increase in frequency of social contacts in experimental group
LaFrance, Garcia and Labreche (2007)	Adult with aphasia (1)	Communication skills	Case study	Increase in communication with introduction of dog
Lang et al. (2010)	Adults with schizophrenia (14)	State anxiety	Pre-post, crossover	State anxiety significantly reduced in experimental condition
Nathans-Barel et al. (2005)	Adults with schizophrenia (10)	Hedonic tone	Pre-post, between subjects	Significant improvement in hedonic tone in experimental group
Orlandi et al. (2007)	Patients undergoing chemotherapy (89)	Anxiety, depression, hostility, somatic symptoms	Pre-post, between subjects	Significant decrease in depression and increase in arterial oxygen in experimental group
Phelps et al. (2008)	Elderly in a nursing home (5)	Depression, mood, social interaction	Multiple baseline	No improvement in any measures
Sobo, Eng and Kassity-Krich (2006)	Children with post-operative pain (25)	Perceived pain	Pre-post, no control group	Significant decrease in post-operative pain
Sockalingam et al. (2008)	Assault victim with bipolar disorder (1)	Mood	Case study	Qualitative improvement in some psychological factors
Stetina et al. (2009)	Drug offenders in a penal institution (28)	Self-concept, emotional experience	Pre-post, three group	Significant increase in social and emotional competency in experimental group
Stetina et al. (2011)	Typical children (32) and adults (34)	Emotion recognition ability	Pre-post, repeated measures	Significant increase in ability to recognize anger and fear by children and adults
Tsai, Friedmann and Thomas (2010)	Hospitalized children (9)	State anxiety, medical feat, physiological arousal	Quasi-experimental, repeated measures	Significant decrease in systolic pressure in experimental condition; no differences in anxiety or medical fear
Yorke, Adams and Coady (2008)	Physical and psychological trauma victims (6)	Level of trauma healing	Qualitative, phenomenological assessments	Participants reported that relationship with a horse had a positive impact

Results

Table 2 presents each study, the subject population (with sample size for the treatment group), the measured construct(s), design, and findings. All of the studies used a dog or dogs as the therapeutic stimulus with the exception of Berget, Ekeberg and Brastaad (2008) (unspecified "farm animals"), Colombo et al. (2006) (a canary), Klontz et al. (2007) (a horse), Krskova, Talarovicova and Olexova (2010) (guinea pig), and Yorke, Adams and Coady (2008) (a horse). I did not specify whether AAT or AAA was used in all studies because of the difficulty in distinguishing the two.

The 28 studies were first reviewed for threats to internal validity and then the remainder were subjected to further scrutiny for threats to construct validity. Ten of the 28 studies were compromised in terms of internal validity because of one or more of the following characteristics: a) the data were qualitative in nature, b) there were no inferential statistics, c) the design was a case study or a single group study lacking a proper control group. For instance, Banks and Banks (2005) found a significant decrease in loneliness scores in their patient population but only for the group that received AAT individually. The patients that received AAT in a group setting did not improve. Moreover, there was no control group for the presence of the dog and the other components of the treatment. Therefore, it is nearly impossible to interpret the findings of this study and it is not clear whether these results should be considered positive or negative. It would depend upon whether one focused on the individual AAT condition, for which there was improvement, or the group AAT condition, for which there was no improvement.

Krskova, Talarovicova and Olexova (2010) attempted to determine if even a small, less sociable animal than a dog, that is, a guinea pig, might have a beneficial effect. They employed a repeated measures design to assess frequency of social contacts in autistic children. However, they did not employ any control for the effect of order of conditions on the outcome. That is, all nine children experienced the control condition (no guinea pig) first, followed by the experimental condition (with guinea pig). Dependent measures were taken during the first condition and then the second and, as expected, they found that there was a significant increase in the frequency of social contacts during the experimental condition. But this serious methodological weakness makes their results essentially uninterpretable.

The other studies lacking or low in internal validity include Anderson and Olson (2006)—single-case study, no quantitative measures; Estevez and Stokes (2008)—case-study design with no inferential statistics; Klontz et al. (2007)—no control group; Kovacs et al., (2006)—no control group, LaFrance, Garcia and Labreche (2007)—single-case study; Sobo, Eng and Kassity-Krich (2006)—no control group; Sockalingam et al. (2008)—case study, no quantitative measures; and Yorke, Adams and Coady (2008)—no quantitative measures.

Two of the studies, Johnson et al. (2008) and Phelps et al. (2008) were fairly strong methodologically but reported no differences between the experimental group and the control group. In these two studies not only were there no significant differences across conditions but there were no systematic improvements in any of the groups. (There may be a statistical power issue in Phelps et al. as they employed a sample size of only five in the experimental group.)

The remaining 16 studies all reported some positive effects and significant differences across experimental and control groups and therefore claimed to provide some evidence of the therapeutic value of AAT or AAA. Of those studies that presented positive effects of AAT or AAA, several of them reported mixed results. That is, while some dependent measures were positively impacted by the experimental intervention, other factors expected to be positively affected were not. One example is Tsai, Friedman and Thomas (2010), which employed a quasi-experimental

repeated measured design to determine the effects of AAT on cardiovascular responses, state anxiety, and medical fear in hospitalized children. They found that systolic blood pressure decreased during and after AAT but there was no improvement in state anxiety or medical fear. This study contained several weaknesses, including lack of randomization. Also, the authors noted that they would have needed a sample size of at least 40 for the estimated effect size; their sample size for the experimental group was only nine. Braun et al. (2009), also employing a quasi-experimental design, measured the effects of AAT on pain relief in children and found significantly decreased pain and increased respiration rate in the experimental group, but no change in pulse or, unlike Tsai, Friedmann and Thomas (2010), blood pressure.

Orlandi et al. (2007) conducted a pre-post between-subjects study of the effects of AAT on various measures in patients undergoing chemotherapy. They found mixed results. There was a significant decrease in depression and an increase in arterial oxygen in the experimental group but there were similar decreases in anxiety, hostility, and blood pressure in both the experimental and control groups. Chu et al. (2009) conducted a randomized pre-post between subjects study with the additional validity check of raters blind to condition. Their subjects were adults with schizophrenia. Compared with the control group, the treatment group showed significant improvement on all measures except for social support and negative psychiatric symptoms.

Cole et al (2007) conducted a well controlled three-group randomized study of various dependent measures in patients hospitalized with heart failure. They found that AAT significantly improved cardiopulmonary function, neurohormone levels and anxiety levels, although other measures and specific components of these measures were not affected. Stetina et al. (2009), also, reported that many but not all measures of social and emotional health in their sample of drug offenders improved in the experimental group.

Stetina et al. (2011) found significant improvements in the ability to recognize fear and anger emotions in a pre-post design, but findings for the other emotions were mixed. Moreover, Stetina et al. (2011) used a control group that consisted of no intervention and an experimental group of AAT. Given that there was no other control for the multiple components of AAT, it is not known whether their positive findings are due to the animal or the novelty of the AAT situation or some other aspect that was not controlled or measured. Berget, Ekeberg and Braastad (2008) incorporated a six-month follow up measure which typically controls for novelty effects. However, their results were not only mixed but also somewhat difficult to interpret. In their sample of adults with psychiatric disorders there was a significant increase in self-efficacy in the treatment group but not in the control group from before intervention to six months follow-up and from end of intervention to follow-up. There was a significant increase in coping ability within the treatment group between before intervention and follow-up whereas no changes in quality of life were found. It is unclear how or why there would be significant improvement after six months but not immediately after the intervention.

Hoffman et al. (2009) conducted a fairly well controlled crossover study of the effects of therapy with dogs on patients with depression. They found that there was a significant decrease in state anxiety in their experimental group. However, given that the patients were suffering from depression, and not anxiety, it isn't clear to what extent these findings should be considered a positive result.

Finally, Gee, Harris and Johnson (2007) measured the effects of AAT on gross motor skills in children and found significantly faster completion of many tasks in the experimental group compared with the control group but also both increases and decreases in accuracy

depending upon the task. They did not find an expected speed-accuracy trade-off, as a power analysis indicated that, for the expected effect size, their sample size was too small. Therefore, for all three of these studies with mixed results there are methodological weaknesses that may account for the mixed results and/or also provide some doubt about the overall validity of the findings. Certainly these studies, while potentially interesting, need to be replicated with a stronger methodological approach.

The remainder of the studies, seven in all, reported more consistent positive effects (significant improvement in the experimental versus control group/condition) for AAT and AAA than those discussed above. These studies ranged from poor to moderately strong in construct validity. The weakest of these was Fournier, Geller and Fortney (2007), which utilized a quasi-experimental pre-post repeated measures design and was subject to selection bias because they did not employ random assignment. Additionally, the control group was quite different from the treatment group and there were pre-test baseline difference across the groups. They reported significant improvement in treatment progress but admit that given all of these weaknesses their conclusions need to be revisited. Colombo et al. (2006) was the only study in this remaining group that incorporated blind raters in order to minimize experimenter expectancy effects, thus strengthening their finding that there were significant increases in mood scores in the experimental group. Lang et al. (2010) and Nathens-Barel et al. (2005) both reported significant improvements in dependent measures in the experimental group only and both were moderately strong studies. However, neither of them employed any special controls for non-specific effects such as novelty.

One of the strongest studies in this group, ironically, suggests that a live dog is not necessary in AAT under some circumstances (Banks, Willoughby and Banks 2008). In this study the researchers utilized random assignment and two control groups (one in which a robotic dog was presented and one in which no intervention was given). The dependent measure was obtained through self-report. The robotic dog condition minimized many of the potentially confounding components of AAT such as placebo, demand characteristics, and novelty. Construct confounding was minimized by using a robotic dog with many of the same characteristics of a live dog. However, when this was done Banks, Willoughby and Banks (2008) found that there was no distinction between the live dog (AAT) and robotic dog scores. That is, there was a significant decrease in loneliness scores in both conditions. This study involved dismantling the construct of the live dog by including a robotic dog control. Once accomplished, there was no difference between the two conditions, suggesting that a live dog is not necessary for some improvement effects. Only one other study included a similar control. Gee, Church and Altobelli (2010), employing a randomized mixed design, used a stuffed dog as one of the control conditions for a live dog in their study of the performance of preschoolers on an object classification task. In this case the group exposed to the live dog made significantly fewer irrelevant choices on the task than either the group with the stuffed dog or another control group with just a human present. These results, taken together with that of Banks, Willoughby and Banks (2008), suggest that one or more of the important characteristics of a dog (or other animal) in AAT is movement, vocalization, warmth, heartbeat, and social interaction and that a stuffed dog, lacking these features, is not a powerful enough stimulus for treatment effects (at least in preschoolers) but that a moving vocalizing toy dog may indeed be for some patient populations despite its lacking the other components.

In summary, with the exception of those studies with poor internal validity, the remaining studies were stronger in internal validity but were generally only fair in terms of strength of construct validity. With only a few exceptions, there was no incorporation of specific procedures

for minimizing or accounting for nonspecific effects such as placebo, demand characteristics, experimenter expectancy effects, and novelty. However, as mentioned earlier, it is admittedly difficult to devise controls for placebo effects and demand characteristics given the nature of the therapeutic intervention in AAT and AAA. Novelty effects can be minimized through the use of long-term follow up measures but only one study of those examined utilized such a measure. Novelty effects can also be minimized with proper controls that come as close as possible to the novelty of the experimental intervention without giving it. The best example of this among these examined studies is Banks, Willoughby and Banks (2008). Novelty effects are often intrinsically related to construct confounding. Construct confounding requires a dismantling procedure that separates the various potentially effective components of the therapeutic agent. Banks, Willoughby and Banks (2008) came closest to accomplishing this by including a robotic dog as a control for some of the elements of a live dog. When this was done, however, there were no differential effects of a live dog, leading to more questions about the "animal" in AAT and AAA.

Conclusions

The conclusions from the current analysis are consistent with the meta-analyses conducted by Nimer and Lundahl (2007) and Souter and Miller (2007). Both sets of authors concluded that the legitimate effects of AAT are likely to be moderate at best and that there were substantial methodological weaknesses in the AAT and AAA literature that prevent a firm conclusion about the effectiveness of these approaches, particularly with regard to the question of whether a live animal is necessary for a therapeutic effect.

Admittedly, this review has limitations of its own, including the fact that, unlike the two meta-analyses, it is qualitative rather than quantitative. Moreover, it may be that it is not entirely realistic to expect a complete dismantling procedure to be performed. This would often involve the animal component alone without the handler and this may not be feasible under most conditions. However, another way to probe construct validity is to include controls for novelty and other aspects of the intervention. This may not entirely address the construct validity question but it would most certainly get us much closer to determining whether the animal in the animal–human team has a more therapeutic effect than some other novel item in combination with the human.

Where do we go from here? Certainly the literature suggests that there may be an effect worth revealing in AAT and AAA. Most critical to strengthening the literature will be the strengthening of construct validity. This can only be done by utilizing more rigorous controls that dismantle the potential components of the AAT and AAA treatments. Currently, the question of "How important is the animal in AAT?" remains unanswered.

References

Anderson, K. L. and Olson, M. R. 2006. The value of a dog in a classroom of children with severe emotional disorders. *Anthrozoös* 19(1): 35–49.

Banks, M. R. and Banks, W. A. 2005. The effects of group and individual animal-assisted therapy on loneliness in residents of long-term care facilities. *Anthrozoös* 18(4): 396–408.

Banks, M. R., Willoughby, L. M. and Banks, W. A. 2008. Animal-assisted therapy and loneliness in nursing homes: Use of robotic versus living dogs. *Journal of the American Medical Directors Association* 9: 173–177.

Berget, B., Ekeberg, O. and Braastad, B. O. 2008. Animal-assisted therapy with farm animals for humans with psychiatric disorders: Effects on self-efficacy, coping ability and quality of life, a randomized controlled trial. *Clinical Practice and Epidemiology in Mental Health* 4: 1–7.

Braun, C., Stangler, T., Narveson, J. and Pettingell, S. 2009. Animal-assisted therapy as a pain relief intervention for children. *Complementary Therapies in Clinical Practice* 15: 105–109.

Chu, C. I., Liu, C. Y., Sun, C. T. and Lin, J. 2009. The effects of animal-assisted activity on inpatients with schizophrenia. *Journal of Psychosocial Nursing* 47(12): 42–48.

Cole, K. and Gawlinski, A. 1995. Animal assisted therapy in the intensive care unit. *Research Utilisation* 30(3): 529–536.

Cole, K. and Gawlinski, A. 2000. Animal-assisted therapy: The human–animal bond. *AACN Clinical Issues Advanced Practice in Acute Critical Care* 11(1): 139–149.

Cole, K., Gawlinski, A., Steers, N. and Kotlerman, J. 2007. Animal-assisted therapy in patients hospitalized with heart failure. *American Journal of Critical Care* 16: 575–585.

Colombo, G., Buono, M. D., Smania, K., Raviola, R. and De Leo, D. 2006. Pet therapy and institutionalized elderly: A study on 144 cognitively unimpaired subjects. *Archives of Gerontology and Geriatrics* 42: 207–216.

Connor, K. and Miller, J. 2000. Animal-assisted therapy: An in-depth look. *Dimensions of Critical Care Nursing* 19(3): 20–26.

Cook, T. D. and Campbell, D. T. 1979. *Quasi-Experimentation: Design and Analysis for Field Settings.* Boston: Houghton Mifflin.

Davis, J. H. 1988. Animal-facilitated therapy in stress mediation. *Holistic Nursing Practice* 2: 75–83.

Estevez, S. W. and Stokes, T. 2008. Social effects of a dog's presence on children with disabilities. *Anthrozoös* 21(1): 5–15.

Fine, A. H. ed. 2010. *Handbook on Animal-Assisted Therapy: Theoretical Foundations and Guidelines for Practice.* 3rd edn. San Diego: Academic Press.

Fournier, A. K., Geller, E. S. and Fortney, E. V. 2007. Human–animal interaction in a prison setting: Impact on criminal behavior, treatment progress, and social skills. *Behavior and Social Issues* 16: 89–105.

Gee, N. R., Church, M. T. and Altobelli, C. L. 2010. Preschoolers make fewer errors on an object categorization task in the presence of a dog. *Anthrozoös* 23(3): 223–230.

Gee, N. R., Harris, S. L. and Johnson, K. L. 2007. The role of therapy dogs in speed and accuracy to complete motor skills tasks for preschool children. *Anthrozoös* 20(4): 375–386.

Herzog, H. 2011. The impact of pets on human health and psychological well-being: Fact, fiction, or hypothesis. *Current Directions in Psychological Science* 20: 236–239.

Hoffmann, A. O. M., Ah Hyung L., Wertenauer, F., Ricken, R., Jansen, J., Gallinat, J. and Lang, U. E. 2009. Dog-assisted intervention significantly reduces anxiety in hospitalized patients with major depression. *European Journal of Integrative Medicine* 1(3): 145–148.

Johnson, R. A., Meadows, R. L., Haubner, J. S. and Sevedge, K. 2008. Animal-assisted activity among patients with cancer: Effects on mood, fatigue, self-perceived health, and sense of coherence. *Oncology Nursing Forum* 35(2): 225–232.

Johnson, R. A., Odendaal, J. S. J. and Meadows, R. L. 2002. Animal-assisted interventions research: Issues and Answers. *Western Journal of Nursing Research* 24: 422–440.

Kazdin, A. E. 1994. Methodology, design and evaluation in psychotherapy research. In *Handbook of Psychotherapy and Behavior Change*. 4th edn, 19–71, ed. A. E. Bergin and S. L. Garfield. New York: Wiley.

Kendall, P. C. and Norton-Ford, J. D. 1982. Therapy outcome research methods. In *Handbook of Research Methods in Clinical Psychology,* 429–460, ed. P. C. Kendall and J. N. Butcher. New York: John Wiley and Sons.

Klontz, B. T., Bivens, A., Leinart, D. and Klontz, T. 2007. The effectiveness of equine-assisted experiential therapy: Results of an open clinical trial. *Society & Animals* 15: 257–267.

Kovacs, Z., Bulucz, J., Kis, R. and Simon, L. 2006. An exploratory study of the effect of animal-assisted therapy on nonverbal communication in three schizophrenic patients. *Anthrozoös* 19(4): 353–364.

Krskova, L., Talarovicova, A. and Olexova, L. 2010. Guinea pigs—The "Small Great" therapist for autistic children, or: Do guinea pigs have positive effects on autistic child social behavior? *Society & Animals* 18: 139–151.

LaFrance, C., Garcia, L. J. and Labreche, J. 2007. The effect of a therapy dog on the communication skills of an adult with aphasia. *Journal of Communication Disorders* 40: 215–224.

Lang, U., Jansen, J. B., Wertenauer, F., Gallinat, J. and Rapp, M. A. 2010. Reduced anxiety during dog assisted interviews in acute schizophrenic patients. *European Journal of Integrative Medicine* 2:123–127.

Linde, K., Fassler, M. and Meissner, K. 2011. Placebo interventions, placebo effects and clinical practice. *Philosophical Transactions of the Royal Society B: Biological Sciences* 366: 1905–1912.

Marino, L. and Lilienfeld, S. 1998. Dolphin-assisted therapy: Flawed data, flawed conclusions. *Anthrozoös* 11(4): 194–199.

Marino, L. and Lilienfeld, S. 2007. Dolphin assisted therapy: More flawed data, more flawed conclusions. *Anthrozoös* 20: 239–249.

Nathans-Barel, Feldman, P., Berger, B., Modai, I. and Silver, H. 2005. Animal-assisted therapy ameliorates anhedonia in schizophrenia patients. *Psychotherapy and Psychosomatics* 74(1): 31–35.

Nimer, J. and Lundahl, B. 2007. Animal-assisted therapy: A meta-analysis. *Anthrozoös* 20(3): 225–238.

Orlandi, M., Trangeled, K. A., Mambrini, A., Tagliani, M., Ferrarini, A., Zanetti, L., Tartarini, R., Pacetti, P. and Cantore, M. 2007. Pet therapy effects on oncological day hospital patients undergoing chemotherapy treatment. *Anticancer Research* 27: 4301–4304.

Phelps, K. A., Miltenberger, R. G., Jens, T. and Wadeson, H. 2008. An investigation of the effects of dog visits on depression, mood, and social interaction in elderly individuals living in a nursing home. *Behavioural Interventions* 23: 181–200.

Shadish, W. R., Cook, T. D. and Campbell, D. T. 2002. *Experimental and Quasi Experimental Designs for Generalized Causal Inference.* Boston: Houghton Mifflin.

Shaughnessy, J. J. and Zechmeister, E. B. 1994. *Research Methods in Psychology.* New York: McGraw-Hill.

Sobo, E. J., Eng, B. and Kassity-Krich, N. 2006. Canine visitation (pet) therapy. Pilot data on decreases in child pain perception. *Journal of Holistic Nursing* 24(1): 51–57.

Sockalingam, S., Li, M., Krishnadev, U., Hanson, K., Balaban, K., Pacione, L. R. and Bhalerao, S. 2008. Use of animal-assisted therapy in the rehabilitation of an assault victim with a concurrent mood disorder. *Issues in Mental Health Nursing* 29: 73–84.

Souter, M. A. and Miller, M. D. 2007. Do animal-assisted activities effectively treat depression? A meta-analysis. *Anthrozoös* 20(2): 167–180.

Standards of Practice for Animal-Assisted Activities and Animal-Assisted Therapy. 1996. Renton, WA: Delta Society.

Stetina, B. U., Kuchta, B., Gindl, B., Lederman Maman, T., Handlos, U., Werdinech, W. and Kryspin-Exner, I. 2009. Changes in emotional competences of drug offenders during dog-assisted group training (MTI). *Journal of Veterinary Behavior* 4(2): 99–100.

Stetina, B. U., Turner, K., Burger, E., Glenk, L. M., McElheney, J. C., Handlos, U. and Kothgassner, O. D. 2011. Learning emotion recognition from canines? Two for the road. *Journal of Veterinary Behavior* 6: 108–114.

Swanepoel, H. C. and Odendaal, J. S. J. 2005. Elephant-facilitated psychotherapy—A clinical evaluation. *Pakistan Journal of Social Sciences* 3(): 205–209.

Tsai, C-C., Friedmann, E. and Thomas, S. A. 2010. The effect of animal-assisted therapy on stress responses in hospitalized children. *Anthrozoös* 23(3): 245–258.

Yorke, J., Adams, C. and Coady, N. 2008. Therapeutic value of equine–human bonding in recovery from trauma. *Anthrozoös* 21(1): 17–30.