



PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO

**PROGRAMA DE ESTUDOS PÓS-GRADUADOS EM PSICOLOGIA
EXPERIMENTAL: ANÁLISE DO COMPORTAMENTO**

Validade Social:

**Implicações da Proposição de um Conceito
para a Análise do Comportamento**

Regina Célia Alves Barreira

São Paulo

2006

Regina Célia Alves Barreira

**Validade Social:
Implicações da Proposição de um Conceito
para a Análise do Comportamento***

Dissertação apresentada à Banca Examinadora da Pontifícia Universidade Católica de São Paulo, como exigência parcial para a obtenção do título de Mestre em Psicologia Experimental: Análise do Comportamento, sob orientação do Prof. Dr Sérgio Vasconcelos de Luna.

* Projeto parcialmente financiado pela CAPES

**São Paulo
2006**

Banca Examinadora

AGRADECIMENTOS

À minha família. Por todo amor e carinho. Por sempre disporem de tudo que foi possível para me ajudar. Por me compreenderem...

Leo, Álvaro, Cláudio, Philomena e Alvarino.

Aos meus professores Amália, Nilza Roberto e Mari. Por tanto que me ensinaram. Pela dedicação e comprometimento. Por se tornarem exemplos para mim...

Amália, Nilza, Roberto, Mari

...e muito especialmente...

Téia.

Àqueles que correspondem ao que entendo por “amigos”. Os que, apesar do mundo nos distanciar por muitas vezes, sempre se fizeram presentes nos momentos mais importantes me ajudando, me apoiando, me fazendo rir quando pensava ser impossível, me ouvindo ou respeitando meu silêncio...

Borny, Quel e Joe.

Àqueles que tornaram meus ambientes de trabalho agradáveis. Pela adorável convivência...

Helena, Didi, Ceição, Maurício e Neusa.

Aquela que por tanto tempo esteve ao meu lado e que sempre terá meu carinho e admiração...

Maria Ester.

Aos meus colegas. Pelo auxílio, preocupação, por me darem o prazer de compartilhar de seus trabalhos e por tantos momentos de descontração...

Thais Sales, Maria Paula, Fábio, Carolzinha, Rodrigo,
Carolina Perrone, Maria, Luciana, Marcelo, Cândido,
Verônica, Thais Nogara, Leila, Ghoerber, André,
Ana Beatriz, Aline, Renata, Paulo, Maria Elisa.

Àquele que sempre foi muito mais do que se poderia esperar de um orientador. Pelo apoio, amparo, disponibilidade e paciência. Por me fazer acreditar em mim mesma. De quem jamais terei todas palavras necessárias para qualificar e agradecer...

Sérgio.

Barreira, R. C. A. (2006). *Validade social: Implicações da proposição de um conceito para a análise do comportamento*. Dissertação de Mestrado. Programa de Estudos Pós-graduados em Psicologia Experimental: Análise do Comportamento. Pontifícia Universidade Católica de São Paulo.

RESUMO

A revisão de literatura em determinada área de aplicação da Análise do Comportamento e o interesse na avaliação das implicações ético-sociais dos trabalhos resultantes de tal revisão levaram ao contato com o conceito de validade social. Dificuldades encontradas na produção teórico-conceitual sobre validade social e a relevância das questões abordadas sob tal rótulo afastaram o presente trabalho de seu interesse original, levando a pesquisadora a procurar respostas às seguintes questões: Que conceito de validade social pode ser extraído da literatura que aborda o tema? Quais são os procedimentos de validação social e quais são seus aspectos metodológicos? Como o constructo “validade social” se relaciona com o Behaviorismo Radical e com a Ciência do Comportamento? Partiu-se da seleção de material sobre validade social através da busca por palavras-chave na base de dados PsycINFO. Os registros resgatados da base de dados foram inseridos em um banco de dados e organizados em categorias de acordo com sua relação com o tema “validade social”. Dos resumos contidos no banco de dados, foram derivados descritores que guiassem a escolha de textos a serem analisados. Foram selecionados textos que descrevessem o contexto histórico da proposta do conceito de validade social e, dentre os trabalhos do banco de dados incluídos nas duas categorias consideradas mais relevantes para os objetivos da pesquisa, selecionaram-se os textos que melhor abrangiam as questões mencionadas. Baer, Wolf e Risley (1968) diferenciam pesquisa básica de pesquisa aplicada e definem o domínio da análise do comportamento abordando questões relacionadas à importância social. Winnett e Winkler (1972), assim como Holland (1983/1978), tecem críticas à prática do analista do comportamento questionam os valores por trás dos objetivos adotados nas intervenções e atribuem ao analista do comportamento o papel de mudar o ambiente social. Dentro deste contexto, Wolf (1978/1976) propõe o conceito de validade social, que tem por objetivo perguntar à sociedade se as aplicações estão fazendo algo de importante. Em seguida, sistematizações de procedimentos de validação social (e.g. Kazdin, 1977) e extensões metodológicas destes (e.g. Van Houten, 1979) começaram a surgir na literatura. Levando em conta o atual estado da produção teórica, conceitual e metodológica sobre validade social, é possível considerar inevitável a adesão à validade social pelo analista do comportamento como pesquisador e prático no que diz respeito à reflexão das questões embutidas no conceito e aplicação de procedimentos que evidenciem valores que embasam sua prática. Entraves metodológicos devem ser encarados como propulsores do desenvolvimento de melhores sistemas de avaliação da importância social das intervenções ao invés de impeditivos do uso de procedimentos de validade social.

Palavras chave: análise do comportamento aplicada; validade social; procedimentos de validação social

Barreira, R. C. A. (2006). *Social Validity: implications of this concept to behavior analysis*. Master Thesis. Programa de Estudos Pós-graduados em Psicologia Experimental: Análise do Comportamento. Pontifícia Universidade Católica de São Paulo.

ABSTRACT

Review of literature pertaining to an applied area of Behavior Analysis and the interest in evaluating the ethical and social implications of works resulting of such a review led into contact with the concept of Social Validity. Controversies found in theoretical and conceptual works on social validity and the relevance attributed to the issue carried the present study away from its original objective. Thus, the questions that guided this study were: What conceptions of social validity can be gathered from the literature regarding this subject? What social validity procedures are available and what are its methodological features? How does the concept of social validity relate to Radical Behaviorism and Behavior Analysis? Literature regarding social validity was selected through search on PsycINFO database. The retrieved records were registered in a database and organized into categories according to their relation to "social validity". Descriptors were taken from the abstracts in order to guide the selection of texts for further analysis.. The criteria for the selection of texts were: a) texts that described the historical context in which the concept was formulated and b) those that best approached the questions of interest in this study, according to the two categories considered the most relevant to the questions. Baer, Wolf, and Risley (1968) differentiated basic research from applied research and defined behavior analysis domain as approaching questions of social relevance. Winnett and Winkler (1972) as well as Holland (1983/1978) criticized the behavior analysts praxis, questioning the values behind the goals established in the interventions and suggesting that behavior analysts should be responsible for changing the social environment. In this context, Wolf (1978/1976) proposed the concept of social validity, that aims to ask the society whether or not interventions are meaningful. Procedures of social validity were, then, reviewed and organized (e. g. Kazdin, 1977) and their methodological extensions (e.g. Van Houten, 1979) began to be pictured from the literature. Facing the current stage of theoretical and conceptual production on social validity, it is possible to consider the adhesion to social validity by behavior analysts in research and applied areas as inevitable for the thoughts over the issues approached in the concept and application of procedures show values that support the interventions. Methodological issues should be faced as a startingpoint to the development of better systems for evaluating the social relevance of interventions instead of being ignored or put aside for the lack of proper methodology of investigation.

Key words: applied behavior analysis, social validity, social validity procedures.

SUMÁRIO

ORIGEM DO PROBLEMA.....	01
MÉTODO.....	07
Seleção, Registro e Organização de Material Bibliográfico sobre Validade Social.....	07
Seleção de Trabalhos para Leitura Integral e Análise.....	13
UMA CONTEXTUALIZAÇÃO HISTÓRICA DO CONCEITO DE VALIDADE SOCIAL.....	18
A Análise do Comportamento Aplicada e a Questão da Importância Social das Intervenções.....	18
Críticas do Ponto de Vista Ético Social à Prática na Análise do Comportamento Aplicada.....	25
A Proposta do Conceito de Validade Social.....	36
Primeiras Considerações a Respeito da Proposta do Conceito de Validade Social.....	53
Primeiras Extensões da Proposta de Validade Social.....	65
Rumos Posteriores do Conceito de Validade Social.....	78
COMPONENTES DA VALIDADE SOCIAL E DIMENSÕES DO CONCEITO.....	79
Componentes da Validade Social, seus Métodos e seus Objetivos..	79
Objetivo de Tratamento como Componente de Validade Social.....	79
Procedimento de Tratamento como Componente de Validade Social.....	91
Resultado de Tratamento como Componente de Validade Social....	98
DISCUSSÃO.....	103
O Conceito de Validade Social.....	103
Procedimentos de Validade Social: Objetivos e Questões Metodológicas.....	105
A Compatibilidade do conceito de Validade Social com o Behaviorismo Radical e com o paradigma metodológico da ACA: implicações para a Prática do Analista do Comportamento.....	109
REFERÊNCIAS BIBLIOGRÁFICAS.....	113

ORIGEM DO PROBLEMA

O problema da presente pesquisa constitui desdobramento de um projeto anterior aqui descrito brevemente com o intuito de esclarecer como a revisão de literatura em determinada área de aplicação da Análise do Comportamento (AC) veio a ressaltar a relevância da análise de um conceito encontrado nesta literatura.

O projeto inicialmente proposto tinha como objetivo analisar a produção de teses e dissertações em Análise do Comportamento Aplicada (ACA) à área de Educação para a Saúde no que tange às implicações ético-sociais das intervenções nelas descritas.

A leitura de pesquisas históricas sobre a produção científica em Análise do Comportamento no Brasil (e.g., Cesar, 2002; Rodrigues, 2005) chamou atenção para o fato de que, das pesquisas em ACA que podiam ser consideradas como da área de saúde, uma parcela expressiva tratava de intervenções em Educação para a Saúde. A relevância desta área para o alcance de objetivos de saúde e a contribuição da AC para ela, que não restrita à divulgação de informação como em algumas abordagens da Educação para a Saúde, despertaram o interesse para o seu estudo, principalmente no que tange às suas implicações ético-sociais.

Alguns passos foram dados para iniciar tal estudo. Partiu-se de um conjunto de registros com informações de trabalhos produzidos em AC em várias áreas em programas de pós-graduação de todo Brasil para seleção de teses e dissertações em ACA que pudessem estar relacionadas à Educação em Saúde. Concomitantemente à criação de um banco de dados com

informações sobre os trabalhos selecionados e adoção de uma terminologia na área de saúde para classificação destes, foi levantada bibliografia que permitisse iniciar a delimitação de implicações ético-sociais do material, tanto em termos mais gerais quanto em termos relacionados à Análise do Comportamento. A literatura abrangia as áreas de epidemiologia e políticas públicas em saúde. Além dessa bibliografia, foi selecionada literatura que conceituasse, situasse historicamente a área de Educação em Saúde e indicasse tendências em tal área para que fosse possível chegar a um conceito de Educação em Saúde que delimitasse a seleção de trabalhos e guiasse a análise.

Inicialmente, a busca de literatura pertinente a estes assuntos foi feita em áreas mais abrangentes como Educação e Saúde. Posteriormente, estes mesmos descritores foram pesquisados no contexto da ACA, principalmente através de consultas aos artigos publicados no *Journal of Applied Behavior Analysis* (JABA).

Buscas pelas expressões “epidemiologia”, “prevenção de doenças” e “promoção de saúde” levaram ao artigo de Winett, Moore e Anderson (1991). O texto descrevia um programa de prevenção de AIDS, porém tal descrição visava ilustrar seu foco principal: a avaliação da adequação da intervenção. Os autores propunham-se a verificar a importância do problema empregando dados epidemiológicos, a usar diferentes estruturas conceituais para avaliar o valor e a adequação do uso da intervenção e o impacto comportamental da intervenção por meio de estudos de eficácia e eficiência. Com isso, os autores pretendiam expandir o conceito de validade social, como o artigo mencionava em seu título.

O artigo de Winett *et al.* (1991) parecia ir ao encontro dos interesses do projeto que vinha sendo desenvolvido e foi tomado como referência para este. Contudo, despertava atenção a menção ao conceito de validade social¹, que estava relacionado a todas avaliações que o trabalho pretendia fazer envolvendo importância, adequação e impacto respectivamente dos objetivos, procedimentos e resultados. Tal atenção ainda foi ressaltada pelo fato de o artigo ser parte de um volume especial do *JABA* sobre validade social, pois isto inseria o trabalho em uma discussão sobre o conceito que pretendia ser expandido pelos autores, o que poderia possibilitar a melhor compreensão do conceito e avaliação de sua utilidade para a análise dos trabalhos em ACA na área de Educação em Saúde. Então, foi realizada a leitura dos textos deste volume especial que indicavam como origem do conceito o texto de Wolf (1978/1976).

Porém, especial dificuldade foi encontrada em relação à produção teórico/conceitual na área do que vinha sendo chamado de “validade social”. Já na proposição de tal conceito na AC, por Wolf em 1978², se evidencia a natureza subjetiva do objeto ao qual o conceito se direciona: a importância social dos problemas a serem tratados pelas pesquisas em ACA³. Isto ainda

¹ Neste ponto, entende-se por validade social o conceito proposto na Análise do Comportamento relacionado à garantia e/ou avaliação da importância social e/ou adequação da pesquisa aplicada.

² O conceito foi inicialmente proposto por Wolf no encontro de 1976 da American Psychological Association (APA), mas, aqui, se fará referência ao artigo publicado pelo autor a respeito de tal proposta no *Journal of Applied Behavior Analysis* (JABA) em 1978, baseado no manuscrito enviado à Divisão de Análise Experimental do Comportamento da APA e que contou com sugestões de Donald Baer, Curt Braukmann, Steve Fawcett, Dean Fixsen, Bill Hopkins, Frances Horowitz, Kathi Kirigin, Jack Michael, Keith Miller, Todd Risley, Jim Sherman e Sandra Wolf.

³ Wolf (1978) ainda aborda outras questões subjetivas com as quais se defrontou como editor do JABA, ampliando a questão aqui mencionada para os objetivos, procedimentos e resultados da pesquisa. Tais questões serão descritas e avaliadas com mais detalhes no decorrer deste trabalho.

implica a mensuração subjetiva em uma abordagem que pretende objetividade (Wolf, 1978/1976).

A produção a respeito de validade social posterior à proposta de Wolf (1978/1976) não se mostrou menos contraditória. Os textos apresentam um grande dissenso em relação ao conceito. Não se encontra acordo a respeito de sua conceituação: a função da validade social e o procedimento para avaliá-la diferem de uma proposta para outra; a utilidade das medidas de validade social muitas vezes é questionada; áreas com objetivos aparentemente semelhantes raramente são abordadas quando se revisa o tema; pouco se discute sobre as propostas emergentes; questões teóricas, filosóficas e metodológicas dos procedimentos de validação⁴ social além de argumentos sobre seus limites serem recorrentes⁵.

Porém, independentemente da indefinição conceitual encontrada na literatura sobre validade social, as questões ligadas ao tema são importantes para a pesquisa em ACA. Por exemplo, como tais pesquisas podem identificar quais comportamentos devem ser mudados? O que deve ser entendido por um comportamento apropriado? Como lidar com coisas sob rótulo subjetivo como “naturalidade”, “empatia” etc? Sendo as principais questões dos seres humanos descritas por tais rótulos, como ignorá-las? Como avaliar aplicações que tratam tais rótulos? Como desenvolver procedimentos com menores riscos de rejeição, ou seja, mais aceitáveis? (Wolf, 1978/1976)

⁴ O termo “validação social” será usado para se referir aos procedimentos que visam garantir e/ou avaliar a importância social/adequação da pesquisa aplicada, sendo distinto do termo “validade social” que irá se referir ao conceito de forma geral.

⁵ A literatura que apresenta tais questões a respeito da validade social é parte da coleta e discutida posteriormente neste trabalho.

Concluiu-se que a trajetória realizada com o objetivo de analisar os trabalhos em Educação para a Saúde produzidos no contexto da ACA afetou tal problema a ponto de torná-lo secundário diante das questões reveladas a respeito do conceito de validade social.

Decorrente disso, o projeto concentrou seu foco em uma avaliação crítica do conhecimento construído em validade social, suas propostas e suas limitações, procurando responder às seguintes questões:

- 1) Que conceito de validade social pode ser extraído da literatura que aborda o tema?
- 2) Que diferentes procedimentos de validação social encontram-se disponíveis? Que variáveis devem guiar a escolha de um deles e quais as implicações desta escolha? Quais os objetivos dos diferentes procedimentos?
- 3) O constructo “validade social” é compatível com o Behaviorismo Radical e com a Ciência do Comportamento? Como a validação social afeta a prática do analista do comportamento? Qual a importância do seu uso e quais as implicações do seu não uso para a proposição de objetivos, procedimentos e resultados em aplicações da Análise do Comportamento?
- 4) Tais procedimentos de validação social, tal como descritos, contemplam criticamente a mudança social ou a manutenção dos valores em vigor?

5) Que estudos exigem procedimentos de validação social? Todas as aplicações carecem de validação social? Estudos não aplicados necessitam de validação social em alguma circunstância?

Deve-se esclarecer que o objetivo do trabalho não é uma revisão do uso ou não uso de procedimentos de validação social em pesquisas aplicadas, como foi feito por Kennedy (1992) e Kevin *et al.* (1997). O propósito da presente pesquisa é revisar o conceito de validade social e avaliar suas possibilidades, limitações, utilidade e importância.

MÉTODO

Seleção, Registro e Organização de Material Bibliográfico sobre Validade Social

Foram coletados registros de trabalhos que continham os termos “validade social” e “validação social” na base de dados PsycINFO⁶ contida no Portal de Periódicos CAPES.

Duas buscas foram realizadas: uma pelo termo *social validity*, outra pelo termo *social validation*. As duas buscas foram cruzadas pelo conector booleano *or* para que registros que contivessem ambos os termos não fossem duplicados. Foram obtidos 687 registros, gravados em um arquivo pelos mecanismos oferecidos pelo próprio portal.

Foi utilizado um banco de dados construído no *software* Microsoft Access 2002 para a organização dos trabalhos selecionados. Os registros foram incluídos no banco de dados em diferentes campos. Os campos contidos neste aplicativo e suas respectivas contribuições para o objetivo da pesquisa e a seleção de trabalhos para obtenção e consulta⁷ estão descritos a seguir.

⁶ Base de dados de periódicos na área de Psicologia

⁷ Deve-se considerar que a totalidade dos resumos incluídos no banco de dados é material a ser analisado. Este material ainda tem a função de guiar a seleção dos trabalhos que serão obtidos e consultados para uma nova seleção que abrange os trabalhos lidos na íntegra.

I. Título do trabalho

Foi coletado o título do trabalho como este se encontrava no registro da base de dados PsycINFO. O título possibilitou a classificação dos trabalhos nas categorias descritas no tópico VII.

II. Autor(es)

Foi(oram) incluído(s) no banco de dados o(s) nome(s) do(s) autor(es) de cada trabalho como este(s) constava(m) no registro.

III. Tipo de documento

Foram registrados os seguintes tipos de documentos:

- i. Resumo de dissertação
- ii. Artigo de periódico
- iii. Capítulo de livro
- iv. Livro

A classificação dos documentos em diferentes tipos indicou diferenças na possibilidade de acesso ao material. As possibilidades de acesso ao documento foram tomadas como critério para eliminação de textos para obtenção e consulta quando inviabilizavam sua aquisição.

IV. Ano de publicação

O ano de publicação indicado no registro foi incluído no banco de dados e sua função na análise foi a de situar cronologicamente a discussão do tema e evolução do conceito.

V. Fonte

Para documentos do tipo “Capítulo de livro” e “Artigo de periódico”, foi coletado, do registro, o nome do livro ou o periódico de origem para inclusão no banco de dados. Além de ser indispensável para a localização dos trabalhos, esta informação serviu para avaliar quais livros e periódicos direcionavam-se mais ao tema abordado facilitando a seleção de trabalhos para obtenção e consulta no tema e em áreas correlacionadas a este.

VI. Resumo

O resumo do trabalho, como encontrado no registro, foi incluído no banco de dados. Assim como o título, o resumo possibilitou a classificação do trabalho nas categorias descritas no tópico VII e a seleção dos trabalhos para obtenção e consulta após a classificação.

VII. Relação com o conceito de validade social

Os trabalhos foram classificados, a partir do título e/ou resumo, pelo tipo de relação que guardavam com o conceito de validade social. As categorias identificadas na leitura dos registros e consideradas relevantes para a organização do material, seleção de trabalhos para obtenção e consulta e execução do presente trabalho estão descritas a seguir.

- i. Trabalhos que apenas mencionavam o uso de um procedimento de validade social.

Trata-se de trabalhos que relatam a realização de uma intervenção e mencionam que uma mensuração da validade social indica ou não a efetividade do programa. São trabalhos que não entram em detalhes, no resumo, quanto ao tipo de medida usada. Neste ponto, considerou-se que tais trabalhos não tinham muito a acrescentar ao estudo no que diz respeito à produção sobre o conceito de validade social e, portanto, podiam ser desconsiderados na análise, razão porque não foram selecionados para obtenção e consulta. Porém, seus resumos foram mantidos pela sua contribuição potencial na avaliação do uso do conceito de validade social.

- ii. Revisões de literatura de uma área ou problema que abrangia o conceito de validade social.

Tal literatura poderia conter informações importantes a respeito das questões envolvidas na validade social em diferentes áreas e problemas. A seleção de trabalhos para seleção e consulta desta categoria dependeu da ênfase dada à validade social na revisão. Ou seja, caso o trabalho tivesse o objetivo de avaliar poucos tópicos na literatura revisada, e entre estes estivesse a validade social, o trabalho poderia ser selecionado, obtido e consultado.

- iii. Trabalhos que relatavam a aplicação de procedimento de validação social de um programa ou de algum aspecto deste.

Foram avaliados pela possibilidade de conterem alguma proposta e/ou discussão de validade social.

- iv. Revisões de literatura sobre validade social em determinada área ou problema.

Nesta categoria foram incluídos trabalhos que tinham a validade social como questão central, mas especificamente em uma área ou problema. A importância destes trabalhos para o presente estudo é semelhante à dos trabalhos incluídos na categoria ii. Porém, diferentemente destes, os trabalhos aqui incluídos foram marcados para obtenção e consulta mais provável, visto que a validade social era seu principal foco.

- v. Trabalhos que abordavam a validade social como questão central.

Incluíram-se aqui trabalhos que abordavam especificamente o conceito de validade social. Tais trabalhos tiveram importância fundamental para a avaliação de propostas e discussões gerais do tema em todos os aspectos do problema do presente estudo, já que se concentravam especificamente na validade social. Em tal categoria se encontra a maior parte dos trabalhos selecionados para obtenção e consulta.

- vi. Aplicação de procedimento de validação social em pesquisa de caráter não aplicado.

Nesta categoria enquadraram-se trabalhos de caráter exclusivamente experimental incluindo aqueles que empregaram animais para teste de modelos explicativos. Tais trabalhos foram importantes na discussão a respeito da necessidade ou não de validade social em estudos experimentais.

Dos textos que compunham o banco de dados, foram extraídos os trabalhos para obtenção e consulta. Textos não selecionados para leitura e análise integral tiveram seus resumos considerados, entre outras coisas, para análise do uso do conceito de validade social.

Seleção e Busca de Trabalhos sobre Validade Social para Obtenção e Consulta

Os trabalhos incluídos nas categorias iv (28 trabalhos) e v (46 trabalhos) do banco de dados foram considerados de maior relevância para a discussão do conceito de validade social por abrangerem vários aspectos e componentes deste e, portanto, a seleção de trabalhos para obtenção e consulta, no que diz respeito à validade social em si, ficou restringida a elas por razão de racionalidade do presente trabalho.

Os trabalhos selecionados pelos critérios descritos acima foram pesquisados em fontes para aquisição dos mesmos. As primeiras fontes utilizadas foram os acervos *on-line* das bibliotecas da USP, PUC-SP, UFSCar, UNICAMP e UnB. Também foi consultado o acervo do Programa de Estudos

Pós-Graduados em Psicologia Experimental: Análise do Comportamento da PUC-SP. Recorreu-se ao Portal de Periódicos Capes, que possui periódicos com textos completos, para localização de trabalhos não encontrados nos acervos das bibliotecas. A última fonte consultada, para aquisição de trabalhos ainda não localizados, foi o site www.ibict.br, que fornece informações sobre periódicos existentes em diversas bibliotecas do país e que podem ser adquiridos pelo sistema COMUT de envio de material entre bibliotecas. Tal busca resultou em 32 artigos.

Este procedimento de busca não produziu indicação de livros e dissertações. Faz parte dos trabalhos localizados apenas um "capítulo de livro", obtido com um contato pessoal, contabilizando, assim, um total de 33 trabalhos coletados.

Seleção de Trabalhos para Leitura Integral e Análise

Do material incluído no banco de dados, foram derivados descritores que delimitassem o tema e guiassem a seleção, leitura e análise dos textos, além de indicar temas correlacionados dos quais foi selecionada bibliografia pertinente para a análise. A seleção de bibliografia, tanto no que tange à validade social em si como de questões correlacionadas, foi guiada pelo objetivo de responder as perguntas levantadas na atual pesquisa.

Questões consideradas relacionadas ao tema são as que abordam o contexto histórico em que foi proposto o conceito de validade social, questões referentes à validade social em si, como sua proposta, seus objetivos e seus principais procedimentos e decorrências práticas e conceituais deles. Outra questão considerada relacionada à validade social e importante para a

discussão do problema pesquisado diz respeito à vida em grupo/sociedade e sua relação com a ciência do comportamento.⁸

Para a seleção de trabalhos que circunscrevessem o contexto histórico da proposição do conceito de validade social foram consultadas as referências dos trabalhos selecionados do banco de dados e também foram consideradas referências encontradas em ementas de disciplinas dentro da AC que abrangiam a história da AC.

Kazdin (1978) e Baer, Wolf e Risley (1968) foram escolhidos por relatarem o contexto em que foi cunhado o rótulo Análise do Comportamento Aplicada (ACA) e a que esta deveria se dedicar. Descrever tal contexto, permitiu introduzir questões presentes já na proposição de tal rótulo e que guardam relação com as questões envolvidas na validade social.

O texto de Baer, Wolf e Risley (1968) apresenta questões relacionadas à validade social, posteriormente proposta por Wolf (1978/1976), ao definir o domínio do periódico destinado à publicação de trabalhos na área da ACA e ao diferenciar pesquisa básica de pesquisa aplicada. Assim, tal trabalho também foi considerado útil ao possibilitar a apresentação destas questões embutidas na definição dos domínios da ACA.

Também, foram selecionados trabalhos que apresentavam críticas aos rumos da ACA e da prática do analista do comportamento que tinham em seu cerne questões abordadas na discussão da validade social com o intuito de

⁸ Outros temas correlacionados identificados foram os indicados por Kennedy (1992): satisfação do consumidor, aceitabilidade do tratamento, validade ecológica e importância clínica dos resultados do tratamento. Por motivo de racionalização do presente trabalho, tais temas não tiveram sua bibliografia levantada, estando a discussão deles restrita à medida como são abordados na bibliografia selecionada sobre validade social. Diante do interesse de aprofundamento das questões referentes à validade social, considera-se o acesso a tais temas relevante e sugere-se análise de bibliografia pertinente. No entanto, é importante lembrar que a maior parte dos temas descritos não é de uso exclusivo da ACA, estando amplamente difundida por várias áreas de conhecimento.

evidenciar como tais questões já atingiam a ACA e vinham sendo discutidas por membros de sua comunidade antes mesmo que o conceito de validade social fosse proposto. Tais textos também apresentam sugestões que vão ao encontro dos procedimentos propostos para avaliação de validade social.

O artigo de Winett e Winkler (1972) foi escolhido pela relação com a avaliação de importância social de objetivos e resultados, que estão entre as finalidades da validade social.

Holland (1983/1978) apresenta questões relacionadas à importância de objetivos e resultados além de uma análise de como esta é estabelecida, o que também ocorre em Winett e Winkler (1972). O artigo de Holland (1983/1978) ainda aborda a aceitabilidade de procedimentos de tratamento, outro objetivo da validade social.

O procedimento de seleção dos trabalhos sobre validade social inclui os passos descritos nos dois primeiros tópicos do método do presente estudo. Retomando sinteticamente, 687 registros obtidos em uma pesquisa ao PsycINFO foram incluídos em um banco de dados e classificados em seis categorias de acordo com sua relação com o conceito de validade social. Das duas categorias consideradas de maior relevância para o estudo do conceito de validade social, foram adquiridos os trabalhos acessíveis pelos meios considerados viáveis, totalizando, até esta fase, 33 textos.

A seleção de trabalhos para leitura integral e análise consistiu na consulta aos tópicos, às referências bibliográficas e a amostras de texto dos 33 trabalhos adquiridos. A partir deste procedimento foram escolhidos quatro artigos que - por razões elucidadas a seguir - foram considerados mais representativos das questões que deveriam ser abordadas para que as

perguntas da presente pesquisa fossem discutidas. Também foi considerado o ano de publicação por informar o contexto histórico no qual o trabalho se inseria.

O primeiro trabalho selecionado foi o artigo de Wolf (1978/1976), que fazia parte das referências bibliográficas de quase todos os textos consultados. Além disso, tais textos atribuíam a proposta do conceito de validade social a Wolf (1978/1976).

Em seguida, foi selecionado o texto de Kazdin (1977), referido na maioria dos trabalhos consultados, por representar a primeira sistematização de procedimentos de validade social encontrada na amostra consultada.

Outro trabalho incluído nesta seleção foi Van Houten (1979). Este artigo tem seu pioneirismo no tema “validade social” relatado, muitas vezes, junto aos trabalhos de Wolf (1978/1976) e Kazdin (1977). A ele é atribuída a primeira extensão metodológica dos procedimentos de validação social apresentados por Wolf (1978/1976) e Kazdin (1977).

O trabalho de Foster e Mash (1999), apesar de ser uma revisão da validade social na área clínica, destacou-se dos demais autores por descrever diversos componentes da validade social de forma geral, além de explorar várias dimensões do conceito. Seu relato de outros artigos anteriores à sua publicação e que também foram considerados importantes para que o constructo “validade social” fosse remontado foi uma variável decisiva na escolha de tal texto. Foster e Mash (1999) retomam a contribuição de vários trabalhos em descrever ou estender procedimentos de avaliação de validade

social e em apresentar limitações e questões práticas e conceituais envolvidas em tais procedimentos.⁹

Outros trabalhos que foram obtidos e consultados, apesar de não serem analisados integralmente, foram mencionados quando se mostrou ser pertinente.

⁹ A essência do trabalho de Foster e Mash (1999) consiste em uma resenha analítica de procedimentos empregados para a validação social. A princípio, tal trabalho foi selecionado para guiar a análise dos 29 textos restantes da amostra de trabalhos sobre validade social adquiridos. A escolha foi baseada na distinção deste trabalho em descrever pormenores das mais diversas dimensões da validade social e pretendia-se usar os aspectos abordados das dimensões como categorias a serem avaliadas em cada texto e apresentadas no capítulo destinado aos procedimentos de validade social. Dado tal objetivo ter se evidenciado demasiado ousado, levando-se em consideração o tempo disponível, optou-se pelo relato de todos aspectos da validade social presentes em Foster e Mash (1999).

UMA CONTEXTUALIZAÇÃO HISTÓRICA DO CONCEITO DE VALIDADE SOCIAL

A Análise do Comportamento Aplicada e a Questão da Importância Social das Intervenções

Com o crescimento da pesquisa aplicada, ficou clara a necessidade de uma publicação para esta, separada de publicações que visavam à pesquisa básica, como era o caso do *Journal of Experimental Analysis of Behavior* (JEAB) onde foram publicados alguns dos primeiros estudos aplicados com humanos. Tal necessidade decorria das diferenças entre as características da pesquisa básica e da pesquisa aplicada (Kazdin, 1978).

Reconhecida a necessidade de uma publicação à parte para os trabalhos aplicados pela *Society for the Experimental Analysis of Behavior* (SEAB), um de seus membros lutou por um novo periódico para a pesquisa aplicada. Em 1968, foi criado o *Journal of Applied Behavior Analysis* (JABA), que cunhou o rótulo “análise do comportamento aplicada” (ACA) para a nova “área de pesquisa” (*sic*) e começou a definir os domínios de tal área (Kazdin, 1978).

No primeiro volume do JABA, foi publicado o artigo de Baer, Wolf e Risley (1968) que define o domínio de estudo do periódico e distingue pesquisa aplicada de pesquisa básica, tornando-se de grande influência.

Em tal artigo, Baer, Wolf e Risley (1968), considerando o então atual estado de demonstração científica da análise do comportamento individual, ressaltam a possibilidade de sua aplicação a comportamentos-problema.

A respeito de tal possibilidade, os autores consideram que,

*Uma sociedade disposta a considerar uma tecnologia do seu próprio comportamento aparentemente estará propensa a apoiar tal aplicação quando ela lidar com **comportamentos socialmente importantes** (...) (p.91, ênfase acrescentada)*

Com a criação de um periódico onde tais aplicações seriam relatadas, então elas poderiam ser extensamente examinadas, refinadas e, eventualmente, replicadas por aplicações melhores (Baer, Wolf e Risley, 1968).

*Aplicações melhores, espera-se, levarão a um estado de sociedade melhor, a tudo aquilo que estende o comportamento dos membros de forma a contribuir para o **bem da sociedade**. Já que a avaliação do que é uma **sociedade “boa”** é em si um comportamento dos seus membros, **tal esperança torna-se ela própria de certa forma filosoficamente interessante**. Mesmo assim, é justa a suposição de que aplicações comportamentais, quando efetivas, possam às vezes levar à **aprovação social** e a **adoção** (p.91, ênfase acrescentada)¹⁰*

Os autores definem aplicações comportamentais analíticas como sendo o processo de utilização de princípios do comportamento na melhoria de comportamentos específicos e de avaliação do quanto mudanças notadas são atribuíveis ao procedimento aplicado.

Em uma pesquisa não aplicada, é plausível que se olhe para qualquer comportamento e para qualquer variável provável de se relacionar com ele, enquanto em uma pesquisa aplicada obrigatoriamente deve-se olhar para

¹⁰ Considera-se o sentido deste trecho suscetível a diferentes interpretações. Possíveis significados, considerando o contexto em que se inserem no artigo e também relações com outros trabalhos, serão apresentadas na discussão da presente pesquisa.

variáveis que possam efetivamente melhorar o comportamento em estudo. Além disso, a pesquisa aplicada deve necessariamente examinar **comportamentos socialmente importantes**, ao invés de um comportamento conveniente para estudo, como na pesquisa experimental (Baer, Wolf e Risley, 1968).

Isto implica, freqüentemente, a pesquisa aplicada estudar os comportamentos na situação social onde costumam ocorrer e não em situação de laboratório. Conseqüentemente, o controle na pesquisa aplicada se torna mais difícil que em uma pesquisa experimental, já que a última se dá em um ambiente onde as variáveis podem ser melhor controladas, além de a sociedade não permitir que comportamentos importantes sejam manipulados meramente para o conforto de uma audiência cientificamente céptica (Baer, Wolf e Risley, 1968).

Das diferentes características entre pesquisa experimental e aplicada decorrem **diferenças nos critérios para avaliação** de uma e de outra. No segundo caso, para os autores, além de aplicado, um estudo tem ser comportamental e analítico, tecnológico, conceitualmente sistemático, efetivo e ainda demonstrar alguma generalidade. Ou seja,

(...) uma análise aplicada do comportamento deve tornar óbvia a importância da mudança comportamental, suas características quantitativas, as manipulações experimentais que analisam com clareza o que foi responsável pela mudança, a descrição tecnologicamente exata de todos os procedimentos que contribuíram para tal mudança, a efetividade destes

procedimentos em fazer a mudança de valor suficiente e a generalidade desta mudança. (p.97)

Ao discutir pormenores de tais características dos estudos aplicados em relação aos critérios de avaliação de estudos comportamentais analíticos não-aplicados, os autores voltam a evidenciar aspectos pertinentes à **importância social** inerente às pesquisas aplicadas.

No que diz respeito aos critérios para se avaliar se uma pesquisa é **aplicada**, os autores atribuem tal qualificação não aos procedimentos de pesquisa, mas ao interesse que a sociedade mostra em relação ao problema estudado.

*Em aplicações comportamentais, comportamento, estímulo e/ou organismo, são escolhidos em função de sua **importância para a sociedade**, não pela sua importância para a teoria. (p. 92, ênfase acrescentada)*

Outro importante critério indicado pelos autores para considerar uma pesquisa como **aplicada** é a existência de uma relação estreita entre estímulo e comportamento estudados e o sujeito envolvido. Assim como parece haver poucos comportamentos que são intrinsecamente alvo da aplicação, poucos sujeitos devem conferir ou estudo o *status* de aplicação. Esse critério também implica avaliar, em dada pesquisa aplicada, quão imediatamente importante o comportamento ou o estímulo é **para o sujeito**. Por exemplo, uma investigação da detecção de sinal visual em um indivíduo retardado pode ter pouca importância imediata, enquanto um estudo semelhante em observadores de radar pode ter importância considerável.

Quanto aos critérios para que uma pesquisa seja **comportamental**, a posição dos autores reflete o pragmatismo behaviorista, já que, para eles, é necessário que se pergunte como seria possível conseguir que o indivíduo fizesse algo efetivamente. Assim, a pesquisa aplicada estuda o que o sujeito pode fazer, não o que pode dizer, com exceção do caso em que uma resposta verbal é o comportamento de interesse. Neste sentido, a descrição verbal do sujeito a respeito de seu comportamento não-verbal não pode ser aceita como medida de seu comportamento. Para que uma pesquisa possa ser considerada aplicada, deve ser possível dizer **o que o sujeito pode fazer**. Pode-se identificar, na exemplificação dos autores sobre o assunto, que estes assumem que, a menos que a sociedade concorde quanto à importância das respostas verbais do sujeito sobre o próprio comportamento não-verbal, é difícil defender que meramente a mudança do comportamento verbal seja um objetivo socialmente relevante.

Para que um estudo seja considerado **analítico** deve demonstrar os eventos que podem ser responsáveis pela ocorrência ou não-ocorrência do comportamento. A pesquisa é analítica quando o pesquisador exerce **controle sobre o comportamento**. Porém, como já aqui exposto, a demonstração do controle experimental em uma aplicação é mais difícil do que em um ambiente de laboratório. Por isso, um estudo aplicado necessita de um delineamento que demonstre controle, sem prejuízo de uma mudança comportamental importante.

Outra das dimensões de uma pesquisa aplicada diz respeito à **efetividade**. Uma aplicação de técnicas comportamentais terá falhado caso não produza **efeitos de valor prático extensos o suficiente**, ao contrário de

uma pesquisa não-aplicada, que pode ter grande valor teórico mesmo produzindo pequenos efeitos. Na pesquisa aplicada, em geral, a importância teórica não está em questão¹¹. Não se trata de uma questão científica, mas de uma questão prática. O critério essencial em uma aplicação é a **importância prática**, o poder de alterar suficientemente comportamentos para que seja **socialmente importante** (Baer, Wolf e Risley, 1968).

Um problema pertinente à avaliação de determinada pesquisa, quanto à sua produção de mudança comportamental em grau suficiente para merecer o rótulo de pesquisa aplicada, é a determinação de quanto o comportamento deve ser mudado. Um estudo que produza uma pequena mudança pode ter grande importância teórica sem atingir um julgamento positivo de uma audiência. Aqui se revela a questão prática envolvida na avaliação da efetividade ser uma questão de grau, considerando-se que uma mudança de maior grau pode ser julgada um sucesso pela audiência. Para os autores, a resposta a este problema é mais provável de ser fornecida por aqueles que lidam com o comportamento.

A última característica de um estudo aplicado abordada pelos autores (1968) diz respeito à **generalidade**. Diz-se que uma mudança comportamental tem generalidade quando perdurar no decorrer do tempo, aparecer em uma ampla variedade de ambientes ou se propagar para uma variedade de comportamentos relacionados. A implicação da generalidade para a relevância social é a de que, como a aplicação visa à melhoria prática em comportamentos importantes, quanto maior a generalidade, melhor será a

¹¹ Embora posições aqui afirmadas sejam controversas, elas são mencionadas apenas pelo seu valor em relação à contextualização histórica do conceito de validade social. Por esta razão, não serão discutidas neste momento.

aplicação, pois tais **comportamentos serão atingidos de forma mais extensa e duradoura**. Por isso, a generalidade é uma característica valiosa de uma aplicação.

Baer, Wolf e Risley (1968) estabeleceram, então, as características de um estudo aplicado e quais aspectos deveriam ser avaliados para que um estudo pudesse ser rotulado como tal. Muitos deles dizem respeito à importância social, ou seja, é necessário avaliar se os comportamentos visados são socialmente importantes; se a mudança destes levam o sujeito a fazer algo eficientemente e com suficiente valor prático; se a manipulação dos comportamentos respeita o fato de estes serem de valor social e se há extensão destes comportamentos nas dimensões espacial, temporal e comportamental, o suficiente para que a mudança possa ser considerada socialmente importante.

Porém, a intenção dos autores foi a de estabelecer os critérios para caracterizar o domínio da pesquisa aplicada e a diferenciação entre esta e a pesquisa básica, não a de propor procedimentos para a avaliação destes critérios, apesar de ser possível identificar algumas passagens que mostram posições dos autores a respeito de tal avaliação. Dentre estas, algumas dizem respeito à avaliação dos aspectos relacionados à importância social e vêm, posteriormente, coincidir com posições tomadas em propostas de avaliação destes aspectos (e.g., Wolf, 1978/1976). Mesmo assim, os autores não fornecem uma proposta sistemática de avaliação dos critérios que apresentam e, especialmente no que concerne ao presente trabalho, dos critérios relacionados acima por sua ligação com a importância social.

Críticas do Ponto de Vista Ético Social à Prática na Análise do Comportamento Aplicada

Delimitados critérios para a caracterização da pesquisa aplicada, e na ausência de propostas sistemáticas para a sua avaliação, a ACA permanece sem suporte metodológico específico ou construção de conhecimento que o valha para tomadas de decisão em relação à pesquisa no que tange à contribuição para o bem da sociedade.

A questão da ausência de propostas para avaliação das pesquisas em ACA no que diz respeito a questões relacionadas à importância social foi abordada por Hayes, Rincover e Solnick (1980), que selecionaram as dimensões da ACA descritas por Baer, Wolf e Risley (1968) para avaliar os artigos publicados nos primeiros 10 anos do JABA. Os autores não apresentaram dados a respeito da dimensão final, efetividade, porque não se julgaram capazes de medir com segurança se as mudanças no comportamento nos estudos analisados são amplas o suficiente e/ou de suficiente significância social para serem de valor prático.

Em tal conjuntura, o surgimento de críticas à atuação do analista do comportamento aplicado é compreensível e justificável, além de necessária. Um caso ilustrativo aparece no artigo de Winett e Winkler (1972).

Estes autores apresentaram uma revisão do JABA, mesmo periódico em que, no número inaugural, Baer, Wolf e Risley (1968) definiram as particularidades da pesquisa aplicada. Os autores tiveram o intuito de investigar os tipos de comportamento-alvo que vinham sendo reforçados e de diversas formas prescritos em intervenções analítico-comportamentais em classes do ensino infantil não-especial. Notaram que comportamentos

inapropriados eram freqüentemente definidos como aqueles que interferiam na ordem, silêncio e tranqüilidade. Seus resultados indicaram que vários autores esforçavam-se para classificar comportamentos como “adequados” e “inadequados”, o que resultava em relações de comportamentos que contribuíam ou não para a ordem na sala de aula. A obsessão por tais tipos de comportamento passou a ser questionada pelos autores, tanto no que diz respeito à sua origem quanto no que concerne à sua necessidade.

Na origem do problema, os autores destacam o legado do inicial desejo dos modificadores do comportamento de demonstração de controle. Segundo Winett e Winkler (1972), tradicionalmente, na pesquisa operante, sempre foi atribuída uma importância menor à topografia da resposta, sendo a natureza do comportamento uma questão secundária, o que encorajava ainda mais o foco no controle. Como consequência, os analistas do comportamento não questionavam os valores e objetivos das escolas para as quais pretendiam trabalhar. Desta forma, eles usavam seus procedimentos para servir aos valores e objetivos das escolas.

Quanto à necessidade de enfatizar comportamentos que colaboram com a ordem, Winett e Winkler (1972) argumentam que, no caso de ordem, tranqüilidade e silêncio serem imprescindíveis em determinada ocasião, tais comportamentos seriam indispensáveis também. Mas, se o silêncio for considerado necessário para todo tipo de atividade, incluindo soletração e atividades sociais, então as crianças serão forçadas a desperdiçar a maior parte do dia sendo *jovens adultos* quietos, obedientes e tranqüilos.

Em suma, para os autores, os analistas do comportamento têm sido instrumentos do *status quo*, acrílicos em relação a um sistema de valores

vigente, bem sucedido em manter a lei e a ordem em detrimento do processo educacional em si, já que os procedimentos costumam funcionar tornando o sistema ainda mais eficiente¹².

A crítica dos autores ainda se estende a outras instituições como hospitais psiquiátricos, centros correccionais e centros para delinqüentes. Eles questionam até que ponto os analistas do comportamento em tais áreas têm ajudado os sistemas institucionais existentes a alcançarem seus presentes objetivos prevenindo, ao invés de produzir, as mudanças necessárias. Concluem que a Análise do Comportamento conhece o papel do ambiente em produzir o comportamento, mas tem se concentrado principalmente em mudar pessoas, de forma que se ajustem melhor às instituições ou sub-sistemas sociais em que vivem.

Como sugestão, indicam que o analista do comportamento pode exercer outro papel: mudar o ambiente social que mantém o comportamento, criando novos ambientes e que

Há uma forte necessidade de diálogos extensivos em nossa comunidade sobre que tipo de seres humanos nós queremos que nossas crianças sejam e venham a ser. Estes diálogos determinariam que valores e comportamento nós queremos que nossas escolas transmitam e recompensem (p. 502).

As sugestões de Winett e Winkler (1972) são complementadas com uma discussão sobre como estes objetivos poderiam ser alcançados tornando a

¹² Não cabe aqui a discussão da atualidade da crítica, mas deve ser lembrado que o artigo foi escrito em 1972 e que é justamente o momento em que foi escrito que o torna relevante para o presente trabalho.

escola uma experiência realizadora para a criança. Apesar de mostrarem algumas alternativas, os autores reconhecem que

A complexidade destas questões em uma época de rápida mudança social e moral é evidente, e o peso de sua significância política social e econômica é enorme. Suspeitamos que não haja uma solução para os problemas da educação, mas que diferentes alternativas podem ser desenvolvidas por diferentes comunidades ou grupos na comunidade (p. 502).

Holland (1983/1978) também forneceu uma crítica a respeito do objetivo das aplicações dos analistas do comportamento, mas, diferentemente de questionar os comportamentos-alvo da aplicação, como o fizeram Winett e Winkler (1972), este autor destaca a forma como estes objetivos são estabelecidos. Mais especificamente, Holland (1983/1978) questiona quem são as pessoas que estão por trás da delimitação dos objetivos, definição das contingências e liberação dos reforços, que contingências controlam estas pessoas e ainda discute o papel do analista do comportamento.

A questão central do texto de Holland (1983/1978) é a crítica proferida ao comportamentalismo e ao controle comportamental como instrumento de manipulação e exploração por aqueles que objetam contra sistemas políticos nos quais o controle é estratificado.

Para o autor, o temor de um controle manipulativo passa a ser justificável quando os agentes do controle do sistema não são os próprios clientes, mas uma instância administrativa. Um exemplo seria o do trabalho em prisões nas quais o analista do comportamento deve satisfação à administração penitenciária e não ao prisioneiro.

Apesar de admitir que, de certa forma, o controle exercido pelas autoridades que empregam o analista do comportamento possa ter intenção benevolente, como elas próprias julgam, Holland (1983/1978) afirma que a atribuição de causas ao comportamento realizada por tais autoridades imputa a culpa àquele que seria, na realidade, uma vítima das contingências. Por isso, a frequência com que os analistas do comportamento aceitam, como causas do comportamento-alvo de intervenção, traços internos apontados por seus empregadores, é tomada como um fato surpreendente pelo autor. Tal surpresa se deve à desconsideração dos analistas do comportamento dos princípios da Análise Experimental do Comportamento, enquanto ciência. Estes princípios, que afirmam a adaptação do comportamento a quaisquer contingências, são incompatíveis com a prática que se tornou tradicional de arranjar contingências particulares de reforçamento em ambientes institucionais especiais, esperando que o comportamento se mantenha em outros ambientes. Assim sendo, a prática do analista do comportamento, ao visar a mudança do indivíduo, ao invés da mudança ambiental, parece mais inclinada a atender aos interesses das autoridades do que mostrarem compatibilidade com princípios científicos nos quais deveria estar baseado seu trabalho.

Como ressalta o autor, *“mudanças duradouras no comportamento exigem que se modifiquem as contingências que produzem e mantêm o comportamento original.”* (p.61)

Com o intuito de ilustrar o problema, o autor arrola algumas “vítimas” para examinar a situação descrita acima: o alcoólatra, o criminoso e o próprio analista do comportamento.

No que diz respeito aos alcoólatras, Holland (1983/1978) aborda a tradição de uso de terapia de controle aversivo baseada em procedimentos de condicionamento clássico no tratamento de pacientes institucionalizados. A negligência de dados relativos à aprendizagem discriminativa levaria os analistas do comportamento responsáveis por aplicações nesta área a ignorar que, fora da instituição onde foi aplicada a terapia, vigoram as condições que mantêm o comportamento original e este se ajusta a essas condições.

Sobre o uso da terapia aversiva, o autor ainda sugere a existência de um elemento de retaliação, visto que tais técnicas são objeto de severas sanções, mas muitos toleram, ou exigem, seu emprego em problemas que costumam causar repulsa. Assim, apesar da presumida preocupação dos terapeutas em auxiliar o paciente, os problemas do indivíduo vêm sendo definidos por outros, que não são os diretamente envolvidos, de forma a justificar o uso de meios aversivos para mudança do comportamento.

Outra postura inapropriada dos profissionais em relação aos problemas por eles abordados pode ser identificada na crítica de Holland (1983/1978) ao trabalho dos analistas do comportamento com alcoólatras. Ela é dirigida à negligência com variáveis importantes no controle do comportamento-problema, em favor da preservação da “pureza” dos dados da pesquisa. No estudo de Gallant (1970, citado por Holland, 1983/1978) é ressaltada a preocupação com a ocorrência de prisões dos participantes do tratamento, efetuadas pela polícia, já que taxas de detenção eram um dos critérios de eficácia do estudo.

Desta forma, o alcoolismo estaria sendo tratado como o problema, sendo desconsideradas as variáveis que fornecem a base operante do beber, as contingências ambientais comuns que mantêm o beber.

A sugestão de solução para o problema incluiria mudanças nas contingências ambientais que constituem as práticas culturais e econômicas, mas tais mudanças seriam difíceis de executar devido ao interesse contrário de grandes corporações, já que estas obtêm vantagens com as práticas em vigor (Holland, 1983/1978).

Em uma situação comparável à do alcoólatra, estaria o criminoso. O analista do comportamento é contratado para “corrigir” os “desvios” do detento em presídios, embora seus conhecimentos a respeito de princípios de controle do comportamento indiquem a busca de variáveis relevantes no ambiente do indivíduo (Holland, 1983/1978).

O autor sugere um melhor conhecimento da amplitude do problema denominado crime, para a identificação das contingências que produzem o comportamento criminoso, e o abandono de causas mentais e fictícias para sua explicação.

Para a compreensão da dificuldade dos analistas do comportamento em focalizar as contingências que atuam na manutenção do comportamento do alcoólatra e do criminoso, o autor discute as fontes de controle e as contingências planejadas para modelar o comportamento do analista do comportamento.

Holland (1983/1978) argumenta que os objetivos são formulados nos níveis mais altos da hierarquia social e que as metas e valores são estabelecidos de acordo com a perspectiva da classe dominante. Verbas para

o desenvolvimento de pesquisa são originárias de organizações que definem os problemas que desejam ver resolvidos, políticas de atuação são definidas pelos que usufruem o máximo do sistema sócio-político-econômico.

Deriva-se de tal situação a atribuição de causas internas ao comportamento, que, nas palavras do autor, “culpam a vítima”.

O mito das causas internas é alimentado devido ao reforçamento fornecido à elite e também devido ao papel que ele desempenha na manutenção do presente sistema. (...) As causas internas servem como justificativa para aqueles que tiram proveito da desigualdade. (p. 69)

Com isso, são atribuídas, aos desprivilegiados neste sistema, causas internas para explicar seu desfavorecimento, enquanto os que ocupam altas posições na hierarquia podem imputar ao mérito pessoal sua situação, já que considerariam punitivo atribuí-la a um sistema que explora as pessoas, admitindo sua participação nesta exploração (Holland, 1983/1978).

O analista do comportamento, como ocupante de uma posição confortável dentro da hierarquia social, vive uma contingência de esQUIVA na qual adere aos objetivos determinados no sistema de forma a não perder suas regalias. O autor lamenta que até aqueles que estão a serviço da modificação do comportamento venham a aceitar causas internas que culpam a vítima e sobre as quais repousam as políticas financeiras ao invés de focalizar o ambiente.

A despeito do que sabem, os modificadores do comportamento freqüentemente tratam os prisioneiros como se eles fossem os “mal adaptados” que deveriam ser submetidos a um programa

de modificação do comportamento e de certa forma esperam que, ao devolvê-lo ao sistema de reforçamento que os conduziu à prisão, aquelas contingências que modelaram, na sua origem, o comportamento criminoso, agora não mais exerçam controle sobre ele. (p. 69)

Para Holland (1983/1978), se os membros de uma sociedade têm problemas, as contingências envolvidas nas instituições, no sistema econômico e no governo é que devem mudar, já que estas são amplamente programadas e determinam nosso comportamento. É necessário mudar as contingências para mudar o comportamento e os fundamentos da análise experimental podem servir àqueles que planejam alterar o sistema.

Um motivo de otimismo, para o autor, está na ironia do fato de que nossos programas de modificação do comportamento, mesmo servindo ao poder e seguindo objetivos de sistemas estratificados, comecem a favorecer o militante social por modelarem o processo de controle social e o tornarem mais claro, nos dando oportunidade de realizar uma análise comportamental do controle existente na sociedade.

O autor destaca alguns problemas intrínsecos a um sistema de controle estratificado evidenciados pela análise comportamental. Um deles ocorre quando o interesse do controlador difere do interesse do controlado. Neste caso, os objetivos estabelecidos “de cima para baixo” fornecem base para o contra-controle, gerando luta e resistência do controlado.

O segundo problema mencionado é a dificuldade de planejar sistemas de exploração que empreguem exclusivamente reforçamento positivo, sendo necessário o uso de coerção ou de restrição para manter o controlado no

sistema e permitir que a maior parte da riqueza permaneça sob poder da classe privilegiada.

Em tais problemas, Holland (1983/1978) vê a perspectiva de extinção do sistema vigente.

A desigualdade de riqueza e de poder requer proteção através de forças de coerção. Isto gera o contra-controle, a luta de classes e uma eventual substituição do sistema. E aí está a base para a seleção natural das práticas culturais que iriam a favor de um sistema igualitário. (p. 72)

Como direção para a solução destes problemas, o autor sugere que o comportamentalismo e sua aplicação evoluam para formas coletivas, baseadas em cooperação, que aumentem ao máximo o reforçamento por auxiliar o próximo, ao invés de fazê-lo às expensas de outros. Trabalhos interessantes nesse sentido basearam-se no modelo extraído da utopia *Walden II* escrita por Skinner em 1948. Holland (1983/1978) exemplifica a referência a *Walden II* relatando aplicações nos quais os participantes definiam os critérios de comportamento e de avaliação de desempenho ou trabalhavam juntos para analisar comportamentalmente seus problemas.

Porém, não se pode esperar que a atividade de analisar as condições de controle do sistema atual que gera segmentação e exploração seja financiada dentro deste mesmo sistema. Para se livrar de controles desse sistema e atuar de forma a não atribuir a culpa à vítima, o analista do comportamento deveria estar fora do sistema, não tendo “nada a perder”. Considerando uma situação hipotética na qual o analista do comportamento não faz parte do sistema, Holland discorre sobre qual seria seu papel.

No meu entender, precisamos trabalhar com as pessoas em todas as nossas instituições sociais, analisando as contingências que as oprimem. Nós trazemos para este empreendimento algum conhecimento especializado e habilidade no uso de dados, as outras pessoas trazem experiência direta, especializada, com as contingências do dia-a-dia. Elas também constituem a audiência para os resultados da análise; nós e elas, como colaboradores, podemos desenvolver soluções experimentais para os problemas da sociedade. (p. 73)

Assim, o papel do analista do comportamento deve ser o de um catalisador, que ajuda a identificar e planejar alternativas de solução e participa delas.

A possível direção que tal exploração conjunta das contingências controladoras tomaria, na opinião do autor, é a de uma sociedade igualitária, guiada pelos princípios de servir aos outros e ser responsável pelos outros. Em uma comunidade de iguais, seria possível um controle mais adequado do comportamento. Um grupo de iguais que discutem entre si a significação de sua vida pode detectar os menores progressos na direção de um objetivo, assim como sinais de “pensamentos incorretos [sic]” (p.74), e fornecer, imediatamente as conseqüências apropriadas. As mesmas pessoas avaliariam mutuamente o desempenho uma das outras em direção ao objetivo comum. Por serem comuns, pouco haveria de se ganhar com a fraude ou subversão dos objetivos. O controle pela aprovação dos membros seria mais eficiente por ser imediata e por seu maior poder em relação aos reforçadores arbitrários.

Assim, o controle das contingências nesse grupo seria mais preciso e detalhado.

Em suma, Holland (1983/1978) procura mostrar que as contingências contidas no sistema estratificado podem ser a base para a luta dos oprimidos, apoiado nos fundamentos da análise dos sistemas sociais através dos princípios de uma ciência fundamentada em resultados empíricos de laboratório. Explicações que culpam a vítima também têm base comportamental, na medida em que fortalecem o *status quo*. A criação de uma sociedade não opressiva, assim como a eliminação dos problemas sociais, seria acelerada com a disseminação da análise das contingências de controle.

Contemporaneamente, como sugerido por Winett e Winkler (1972), aparecem discussões a respeito da necessidade e utilidade das mudanças no comportamento na Análise do Comportamento Aplicada, que também vão ao encontro das preocupações de Holland (1983/1978) a respeito de como devem ser estabelecidos objetivos para intervenções.

Tais discussões acontecem em áreas como satisfação do consumidor, aceitabilidade do tratamento, validade ecológica e a importância clínica dos resultados do tratamento, que visavam avaliar a adequação e utilidade de serviços e aplicações. Estas áreas podem ser consideradas como relacionadas ao conceito de validade social e abrangidas por ele (Kennedy, 1992).

A Proposta do Conceito de Validade Social

O conceito de validade social foi formalmente proposto por Montrose Wolf em uma discussão ocorrida em 1976, durante o encontro da *American Psychological Association* (APA), e é um dentre outros tantos conceitos de

validade absorvidos pela Análise do Comportamento Aplicada, sendo quase exclusivamente usado por esta disciplina (Geller, 1991).¹³

Envolvido na organização do que viria a ser o periódico para publicação de pesquisas no campo da ACA, o JABA, Wolf deparou-se com várias questões, entre elas a de qual seria a frase impressa na parte interna da capa do periódico que expressaria o propósito da publicação. Em uma conversa informal sobre o assunto, Donald Baer respondeu a Wolf de forma eloquente à questão:

“Ele é para a publicação de aplicações de análise do comportamento a problemas de importância social.” (p.203)

Para Wolf, a afirmação soava razoável, por isso julgou que deveria ser verdadeira. Então, a frase foi colocada no periódico para impressão (Wolf, 1978/1976).

Mas, Wolf ainda via um problema na afirmação: ele não tinha certeza do que significava “importância social”, muito menos de como medi-la. Uma consulta ao dicionário só aumentou sua inquietação: como significado de **importância** constava “o que tem valor”, e de **social**, “pertencente à sociedade”. Com isso, Wolf concluiu que algo de importância social deveria ser julgado por alguém como sendo de valor para a sociedade.

Wolf percebeu que, infelizmente (*sic*), isso soava um tanto subjetivo, e critérios subjetivos não são muito respeitáveis em nosso campo¹⁴. Nós consideramos nossa ciência como natural, preocupada com a medida de

¹³ A afirmação de Geller (1991) pode ser verdadeira se considerada dentro do campo da Psicologia, como foi confirmado pelo levantamento bibliográfico do presente trabalho. Mas, em outras áreas de conhecimento, é muito provável que o conceito de validade social seja comum na literatura. Porém, não com o mesmo sentido que ele aparece na ACA.

¹⁴ Certamente, referindo-se ÀC.

eventos naturais e que, como outras ciências naturais, está interessada nos aspectos objetivos da natureza e, por isso, abandonou a dimensão subjetiva dos eventos naturais, algumas vezes em seu passado primordial. Assim, nós a consideramos como a mais pura e objetiva das ciências sociais, porque as demais freqüentemente misturam a seus dados porções de medida subjetiva difíceis de digerir (Wolf, 1978/1976).

O autor lembra que as medidas subjetivas nem sempre foram tão suspeitas. Ele retoma a história da Psicologia, na qual, durante algum tempo, a introspecção foi o método básico. Em oposição a tal método, surgiu o positivismo, que tentava voltar para os dados básicos e, então, aumentar a concordância e diminuir os equívocos que ocorriam das ambigüidades de sentido. No livro *Behaviorism*, Watson (1930, citado por Wolf, 1978/1976) descreve que havia dois pontos de vista opostos no pensamento psicológico americano: o introspectivo ou psicologia subjetiva e o behaviorismo ou psicologia objetiva. Skinner (1953, citado por Wolf, 1978/1976), em *Ciência e Comportamento Humano*, também argumenta contra medidas subjetivas e eventos privados e, quando da definição de uma análise funcional, recomenda que nos concentremos nos dados comportamentais objetivos.

Bem treinados nessa tradição, todos os analistas do comportamento aplicados concordavam que, no JABA, tudo deveria ser medido de forma objetiva e, como primeira prioridade, medidas subjetivas seriam evitadas. Já Wolf, em um momento de precipitação, ao colocar na parte interior da capa da publicação o propósito sugerido por Baer, havia comprometido o periódico com um objetivo, um critério final, uma razão de ser, claramente subjetivos e sem um meio adequado de mensuração.

Esperando o escárnio de seus colegas ao propósito atribuído ao periódico, Wolf surpreendeu-se ao se deparar com a inclinação das pessoas a aceitá-lo, além de se referirem a ele em suas revisões e até tomá-lo como critério para recomendação ou não de manuscritos para publicação. Os editores freqüentemente relatavam que determinado manuscrito não tinha importância social o suficiente, enquanto outros poucos tinham-na em quantidade moderada. Apesar de não ter certeza do que seria a importância social ou de como medi-la, Wolf notava que muitos de seus colegas não tinham problemas em reconhecê-la ou medi-la, visto que se referiam aos montantes dela encontrados nos trabalhos submetidos para publicação.

Mesmo assim, Wolf continuava receoso de críticas vindas da audiência de leitores. Ocasionalmente, recebeu reclamações a respeito da importância social, mas elas tratavam de um questionamento sobre o motivo das pesquisas publicadas não terem maior importância social. Desta crítica, Wolf podia culpar os autores dos trabalhos. Porém, como editor, caso viesse a ser repreendido pelo uso de um critério subjetivo desfocado, como “importância social”, sabia que não teria desculpas.

A questão da medida subjetiva continuou tornando complicada a vida de Wolf. Uma de suas funções, como editor chefe, era preservar os padrões do periódico, e todos no campo “sugeriam fortemente” que estes fossem rigorosamente mantidos com exceção dos próprios manuscritos que, por sua significância não usual, mereciam consideração especial. Em todos os casos, os padrões cuja preservação lhe foi confiada eram aqueles que requeriam dados objetivos e fidedignos. Neste contexto, Wolf passou a se sentir preocupado quando seus colegas mais estimados começaram a submeter

artigos para o JABA que incluíam, sem disfarces, dados ostensivamente subjetivos.

Wolf menciona um dos primeiros trabalhos que chegou com esta característica, vindo de Bob Jones e Nate Azrin (1969, citado por Wolf, 1978/1976). Os autores conduziram uma série de experimentos e mostraram que podiam eliminar gagueira sincronizando a fala com um batimento regular. Mas eles encontraram um problema: os ouvintes julgavam que a fala dos sujeitos soava “artificial”, apesar da eliminação da gagueira. Então, Jones e Azrin exploraram várias durações do batimento. A parte considerada difícil para Wolf referia-se ao procedimento adotado pelos autores: tomar índices de “naturalidade” da fala em diferentes durações de batimentos, com base em afirmações de juízes.

Wolf quis perguntar a Jones e Azrin se eles sabiam que tipo de precedentes ele levaria para o periódico publicando tal procedimento para medir “naturalidade”; se as pessoas mais influenciáveis do nosso campo não iriam começar a pensar ser possível medir o que as pessoas sentem a respeito de todo tipo de coisas subjetivas e o que aconteceria ao campo da ACA.

Ao mesmo tempo, Wolf imaginava quais seriam as respostas dos autores. Eles deveriam concordar que isto complicaria nossa ciência um pouco, mas argumentariam que, se tais coisas descritas sob rótulos subjetivos eram as mais importantes para as pessoas, seria destas que deveríamos nos ocupar. Além disso, como postulantes de uma ciência aplicada do comportamento humano, deveríamos nos dedicar a ajudar pessoas a se tornarem mais habilidosas em obter reforçadores.

Na mesma época, Wolf recebeu o artigo de McMichael e Corey (1969, citado por, 1978), que relatava que estudantes de um curso tipo PSI (*Personalized System of Instruction*) foram melhor sucedidos nos exames que aqueles que participaram de um curso tradicional de leitura. O problema do artigo, para Wolf, era que eles perguntaram para os alunos quanto eles gostaram de seus cursos e os estudantes do PSI classificaram seu curso de forma muito mais positiva que os do curso tradicional.

Wolf teceu algumas considerações a respeito desta particularidade do artigo. Para ele, os estudantes foram questionados sobre quanto gostaram do programa porque, obviamente, deveriam gostar dele, até mesmo porque isso foi feito para o bem deles. Mas, mesmo que eles dissessem que não gostaram, continuaríamos sabendo o que é melhor para eles. Se o procedimento é efetivo, não importaria se alguém gosta ou não dele. Novamente, Wolf mostrava-se preocupado com os precedentes que a publicação de tal artigo traria para o periódico e para o campo da ACA. Com o tempo, aqueles que não apreciam os riscos dos dados subjetivos começariam a pedir *feedback* dos participantes em seus programas de tratamento.

Também neste caso, Wolf supõe quais seriam os argumentos de McMichael e Corey fossem questionados a respeito do procedimento. Eles diriam que o *feedback* dos participantes não era uma questão trivial. Se os participantes não gostassem do tratamento, então poderiam evitá-lo, interrompê-lo ou reclamar dele. Assim, a sociedade estaria menos propensa a usar tal tecnologia, independentemente do quanto pudesse ser efetiva e eficiente.

Enquanto Wolf se deparava com problemas de medida subjetiva nos artigos submetidos ao JABA, ele e seus colegas de projeto também tiveram problemas com *feedbacks* subjetivos voluntários em questões similares. Colegas, editores e membros da comunidade perguntavam sobre a escolha dos objetivos comportamentais do programa. Tratava-se de perguntas como: “Como sabem que competências ensinar?”, “O que vocês entendem por competências apropriadas?” O grupo de Wolf justificava que, como psicólogos, eram os mais qualificados para julgar o que seria melhor para as pessoas. Porém, não foram capazes de convencê-los com essa lógica.

Mais problemas ocorreram quando o programa foi replicado em outra comunidade. O *feedback* da comunidade foi dado de maneira drástica. Antes de saberem que existiam queixas a respeito do programa, Wolf e seus colegas foram “demitidos”. Ainda havia aqueles que contestavam a importância de alguns dos resultados relatados dizendo que ocorreram mudanças, mas nada indicava que estas fossem importantes.

Assim, Wolf descreve o que concluiu de sua experiência:

A mensagem que parecia estar nos sendo dada era a de que ‘importância social’ era um julgamento subjetivo de valor que apenas a sociedade era qualificada a fazer. Se nosso objetivo era, como descrito no JABA, fazer alguma coisa de importância social, então teríamos que desenvolver sistemas e medidas para perguntar à sociedade se nós estávamos cumprindo este objetivo. A sugestão parecia ser de que a sociedade necessitaria validar nosso trabalho em três níveis:

1. A *significância*¹⁵ social dos objetivos. Os objetivos comportamentais específicos são realmente o que a sociedade quer?
2. A *adequação social* dos procedimentos. Os fins justificam os meios? Isto é, os participantes, os cuidadores e outros consumidores consideram os procedimentos de tratamento aceitáveis?
3. A *importância social* dos efeitos. Os consumidores estão satisfeitos com os resultados? Todos os resultados, incluindo aqueles não previstos? (pp. 206-207)

A partir de então, Wolf e seus colegas passam a tratar estes julgamentos como elementos de análise da *validade social*. A eles parecia que, dar à validade social o mesmo *status* dado à medida objetiva e sua fidedignidade, traria o consumidor, que é a sociedade, para a sua ciência, abrandaria a imagem desta e tornaria mais sólida a busca da relevância social.

A partir de então, Wolf descreve os passos tomados por ele e sua equipe para resolver questões que envolvem valores subjetivos, elaborando métodos para validá-los nos três níveis sugeridos.

Uma destas questões do seu programa, colocada como exemplo por Wolf, diz respeito à demanda dos participantes por “aprender a se relacionar com os jovens”, demanda esta que, a princípio, foi desconsiderada e, posteriormente, os resultados da pesquisa demonstraram ser uma habilidade importante, corroborando o senso comum da comunidade. O desafio

¹⁵ Wolf (1978) refere-se às expressões “significância social”, “importância social” e “relevância social” que, neste trabalho, são entendidas como tendo o mesmo sentido. Parte da bibliografia emprega apenas uma das expressões tanto para objetivos como para resultados (e.g., Foster e Mash, 1999), o que corrobora tal interpretação.

encontrava-se em determinar quais comportamentos os participantes necessitavam para “saberem se relacionar com jovens”, de forma que os pesquisadores fossem sensíveis à comunidade e aos dados.

“Relacionar-se” parecia ser um enigma comportamental complexo, composto de comportamentos sociais sutis, e eles não tinham certeza de como começar a analisá-lo comportamentalmente. Então, além do exemplo que já tinham da questão de “naturalidade” da fala dos gogos, de Jones e Azrin, eles recorreram ao método dos psicólogos rogerianos Haase e Tepper (1972, citado por Wolf, 1978/1976).

Haase e Tepper (1972, citado por Wolf, 1978/1976) estavam interessados em “empatia” e queriam encontrar os comportamentos não verbais do conselheiro, que envolviam empatia, para poderem ensiná-lo e avaliá-lo melhor durante seu treinamento. Então, simularam situações de aconselhamento com vários componentes não verbais, gravaram em fitas de vídeo, apresentaram as fitas para conselheiros experientes que classificaram a quantidade de empatia em cada trecho. Um resultado que pareceu surpreender aos autores foi que comportamentos não verbais somavam mais que o dobro de julgamentos de empatia em relação aos comportamentos verbais.

Ocorreu a Wolf e seus colegas que este modelo poderia ser usado para analisar o “significado” de todo tipo de rótulos verbais subjetivos e complexos e parecia ser uma forma de descobrir que comportamentos fizeram de alguns participantes de seu programa mais hábeis em se “relacionar com jovens” que outros. Em um estudo realizado por eles (Willner *et al.* (1977, citado por Wolf, 1978/1976)) foram mostradas aos jovens gravações em fitas de interações entre participantes e jovens para que fossem listadas coisas de que gostavam

ou que lhes desagradavam. Os comentários foram classificados em 5 categorias e os comportamentos melhor classificados pelos jovens foram levados para o treinamento. Seus participantes foram muito melhor classificados pelos jovens após receberem instruções nos comportamentos preferidos por estes.

Uma importante informação obtida por acaso no estudo de Willner *et al.* (1977, citado por Wolf, 1978/1976) é a de que não foi possível prever os comportamentos preferidos pelos jovens. Para testar a hipótese dos jovens não terem sido capazes de verbalizar comportamentos sutis, foram realizadas entrevistas estruturadas que incluíam descrições mais detalhadas dos comportamentos que os pesquisadores pensavam ser importantes. Mesmo assim, estes comportamentos foram classificados pelos jovens como menos importantes que os indicados anteriormente por eles. Jovens que não participaram da primeira entrevista também mostraram preferências semelhantes aos outros nesta fase. Isto tornou claro para os pesquisadores que eles não podiam prever muito bem o que muitos rótulos subjetivos de fenômenos comportamentais complexos podiam significar para os juízes. Assim, eles estavam mais próximos de revelar os comportamentos sociais envolvidos em se “relacionar com jovens”, apesar de a tarefa não estar completa (Wolf, 1978/1976).

Wolf (1978/1976) voltou a participar de outros estudos que usavam este método para examinar a validade social de objetivos comportamentais. Então, concluiu que, em vez de a validade dos objetivos comportamentais de um programa de tratamento ser decidida pelos profissionais,

(...) poderíamos acessar o consumidor específico ou representantes da comunidade relevante através de entrevistas ou categorias, e determinar mais precisamente quais seriam os problemas socialmente significativos. Baseados nos exemplos de Jones e Azrin (1969) e Haaper e Tepper (1972), descobrimos que podíamos estabelecer a importância ou validade social de classes complexas de comportamento que têm rótulos subjetivos. Suplementando nossas tradicionais medidas objetivas, poderíamos determinar a relação entre os comportamentos medidos objetivamente e os rótulos subjetivos. Este procedimento abriu oportunidades para explorar todos os objetivos importantes que são descritos por rótulos subjetivos.

(p. 209)

Para sumariar o método para determinação de comportamentos-alvo, Wolf (1978/1976) recorre a Minkin *et al.* (1976, citado por Wolf, 1978/1976), um dos trabalhos dos quais participou. Minkin *et al.* (1976, citado por Wolf, 1978/1976) consideraram “afeição” como exemplo de um comportamento complexo. Se for tomado como objetivo de um analista do comportamento tornar pais mais afetivos, os comportamentos importantes componentes de afeição devem ser especificados. Então, devem ser seguidos quatro passos para validar a importância social destes comportamentos. Primeiro, coletar amostras de interações pai-filho. Segundo, desenvolver definições fidedignas e registrar comportamentos específicos. Terceiro, empregar juízes relevantes da comunidade para classificar as amostras de interação e avaliar a quantidade de afeição mostrada na interação do par, o que pode ser feito com uma escala de

classificação bi-polar com pólos rotulados com a quantidade de afeição. O quarto passo envolve correlacionar as classificações dos juízes com a escala composta de comportamentos medidos objetivamente dos pais. O coeficiente de correlação resultante pode indicar o nível de relação dos componentes específicos de afeição objetivamente medidos com o “significado” de afeição como classificado pelos juízes.

Minkin *et al.* (1976, citado por Wolf, 1978/1976) julgaram que alguns dos componentes comportamentais importantes de outras classes de comportamentos sociais complexos poderiam, provavelmente, ser identificados com o uso deste procedimento de validação social.

Para Wolf (1978/1976), estava claro que uma parcela dos mais importantes conceitos da nossa cultura era subjetiva; portanto, também os mais importantes, o que o levou a perguntar porque deveríamos deixar para os outros os melhores objetivos humanos e problemas sociais.

Depois de abordar a validade social dos objetivos, Wolf (1978/1976) mostra um segundo tipo de validade social que imprime sua importância: a adequação social (em termos éticos, de custo e viabilidade) dos **procedimentos** de tratamento. Novamente, o autor descreve um estudo que evidencia analistas do comportamento, como Azrin, Risley, além dele próprio e seus colegas de projeto, pedindo *feedbacks* de clientes e cuidadores, desta vez sobre a aceitabilidade de seus procedimentos.

O autor julga que determinar a aceitabilidade do tratamento pode não ser importante só por razões éticas, mas também porque ela pode estar relacionada à efetividade do programa, assim como à probabilidade de o programa ser adotado e apoiado pelos outros.

A terceira dimensão da validade social discutida por Wolf (1978/1976) é a importância social dos **efeitos** do tratamento comportamental. Trata-se de saber se os consumidores estão satisfeitos com os resultados do programa, todos eles, incluindo os que não foram planejados. Como programas comportamentais são elaborados para ajudar alguém com problemas, a avaliação de quanto o programa é útil ou não cabe somente ao consumidor. Isto não significa que o analista do comportamento não deva dar suas opiniões e que estas opiniões não devam ser apoiadas por dados comportamentais empíricos e objetivos, porém a decisão final¹⁶ sobre a utilidade do programa para a solução dos problemas deve ser dada pelos participantes e outros consumidores.

Wolf (1978/1976) conta que muitos analistas do comportamento estavam começando a validar seus dados objetivos através de medidas subjetivas sistemáticas de satisfação do consumidor. A questão pode ser particularmente bem ilustrada pelos estudos de Brigham *et al.* (1973, citado por Wolf, 1978/1976) e Braukmann *et al.* (1975, citado por Wolf, 1978/1976), entre outros

Brigham *et al.* (1973, citado por Wolf, 1978/1976), estudando a melhora da qualidade de composições de crianças, descobriram que algumas contingências provocavam aumentos em dimensões objetivas e pouco efeito nas classificações subjetivas de qualidade, enquanto outras contingências aumentavam ambas as medidas. Tais achados contrastam com conclusões de outros autores que obtiveram o aumento de suas medidas objetivas acompanhadas de aumentos nas medidas subjetivas, e demonstram que o aumento nos dados objetivos não garante necessariamente um efeito

¹⁶ Tal opinião de Wolf (1978) será discutida nas considerações finais do presente trabalho.

significativo sobre os dados subjetivos. Tal achado corrobora a necessidade de medida subjetiva dos resultados, já que, apesar da alta correlação entre a medida subjetiva e a medida objetiva, que depõe a favor dos estudos onde tal correlação é encontrada, ela não é necessariamente verdadeira para qualquer estudo.

Braukmann *et al.* (1975, citado por Wolf, 1978/1976) descreveram como o *feedback* dos consumidores pôde ser usado para fornecer um controle de qualidade contínuo da disseminação do seu programa. Foram realizadas avaliações anuais nas quais membros e agências da comunidade classificavam, entre outras coisas, a efetividade do programa. Os autores relatam que, desde que tais avaliações passaram a ser realizadas, os pesquisadores deixaram de ser “expulsos” das comunidades. Além disso, as classificações de satisfação do consumidor tinham alta correlação com medidas objetivas de efetividade.

Wolf (1978/1976) demonstra que os analistas do comportamento não são os únicos cientistas sociais aplicados que têm perguntado pela validade de suas medidas verificando-as junto à sociedade. A preocupação com a validade social de medidas objetivas parece ser também uma questão em debate em outras ciências sociais. Campbell (1976, citado por Wolf, 1978/1976) e Schumaker (1973, citado por Wolf, 1978/1976) abordaram tal problema na economia, questionando medidas de qualidade de vida, baseadas somente em dados objetivos, sem considerar informações sob rótulos subjetivos. Na medicina, Levi e Anderson (1975, citado por Wolf, 1978/1976) sugeriram que as medidas objetivas de qualidade de vida deveriam ser suplementadas por medidas subjetivas.

Apesar de ser considerado algo bom, a validade social¹⁷ não tem sido amplamente usada, já que dados subjetivos são considerados dados de risco. Dados subjetivos podem não ter relação alguma com os eventos reais. Portanto, existe um sério perigo de tais dados nos conduzirem ao erro. Alguns pesquisadores do comportamento têm relatado exemplos de falta de correspondência entre dados relatados por clientes e dados obtidos de observadores. Wolf (1978/1976) levanta alguns motivos pelos quais estas discrepâncias poderiam estar ocorrendo.

A primeira possibilidade destacada pelo autor é que as contingências da situação criem distorção. Devemos desconfiar do comportamento verbal, pois sabemos que nem sempre entendemos as contingências que operam sobre ele. Necessitamos de cautela quando pedimos uma descrição de um evento privado, como satisfação com o programa, porque não temos meios adequados de verificar a fidedignidade de relatos verbais de modo independente.

Assim, é possível conceber a manipulação de contingências que alterem a avaliação do consumidor tanto para melhor como para pior. Porém, lembra Wolf (1978/1976), a fidedignidade das medidas objetivas também são passíveis de manipulação. Alguns estudos mostram que o registro de comportamentos por observadores pode ser afetado por uma variedade de variáveis, o que deve sempre ser levado em conta quando desenvolvemos um sistema de medida que envolve observadores. Mas, mesmo assim, continuamos a usar estes sistemas porque são o único meio de obtermos

¹⁷ Deve ser lembrado, neste ponto, que Wolf (1978) considera, a essa altura do texto, validade social como a mesma coisa que medida subjetiva, visto ser esta sua proposta de validação social. Porém, como ficará demonstrado na descrição de trabalhos posteriores, a medida subjetiva não é a única proposta encontrada na literatura para responder às questões aqui colocadas por Wolf (1978) sob o rótulo de validade social.

alguns dados importantes; em geral, eles funcionam e nos sentimos confiantes de que estamos adquirindo conhecimento a respeito das condições que os distorcem.

Da mesma forma, sabemos que medidas de validade social são manipuláveis, mas não podemos permitir que isso nos leve a negligenciá-las. Devemos garantir as condições necessárias para que as pessoas sejam as melhores avaliadoras de seus tratamentos. Apesar de sabermos pouco destas condições, podemos ficar atentos a elas e estudar seus efeitos sobre dados subjetivos, assim como pesquisadores têm estudado seus efeitos em sistemas de medida objetivos dependentes de observador. Previamente, podemos esperar que tais condições envolvam educação sobre opções, ausência de coerção e anonimato (Wolf, 1978/1976).

A segunda hipótese, indicada pelo autor, para explicar as discrepâncias, é que não estejamos registrando objetivamente as mudanças nos comportamentos e condições às quais o consumidor responde. Para ele, se este for o caso, nossos sistemas de medida devem ser aprimorados.

A terceira possibilidade, considerada mais séria por Wolf (1978/1976), é que as medidas subjetivas seriam impossíveis, pois seres humanos não seriam capazes de julgar sua própria situação de forma suficientemente acurada. Tal possibilidade é considerada inaceitável para o autor caso nosso objetivo seja desenvolver uma ciência social aplicada orientada e sensível ao consumidor. Para fortalecer seu argumento, Wolf (1978/1976) retoma o caso de Levi e Anderson (1975, citado por Wolf, 1978/1976) que acreditavam que cada indivíduo deveria ser assumido como o melhor juiz do seu próprio bem estar e

lembravam que muitas opiniões de *experts* e elitistas divergiram do esperado pelo homem comum em alguns momentos da história do mundo.

Deveríamos, segundo Wolf (1978/1976), desenvolver meios de ensinar as pessoas a observar melhor seus comportamentos e condições de forma a fazerem decisões mais acuradas para seu aprimoramento, já que a sondagem de opiniões delas demonstrou ser capaz de excelentes previsões sobre eleição de comportamentos baseada em relatos verbais. Também não deveríamos nos desesperar com as falhas apontadas por estudos nas medidas de validade social, assim como foram apontadas em estudos sobre medidas objetivas dependentes de observadores, pois as medidas têm sido nosso objeto e, assim como desenvolvemos muitos sistemas de medida ingênuos, com certeza também seremos capazes de desenvolver sistemas de medida que possam nos dizer melhor se nossos clientes estão ou não felizes com nossos esforços e resultados.

Wolf (1978/1976) conclui sua proposta de estudar a validade social por meio de medidas subjetivas relembrando como estas foram tratadas nos primórdios da história de nossa ciência. Ele considera que a conclusão de muitos sobre a inadequação de medidas subjetivas foi resultado da preocupação expressa por autores como Skinner e Watson com papéis causais inapropriados que variáveis internas hipotéticas, relatadas subjetivamente, estavam desempenhando nas ciências sociais. Para o autor,

Um novo consenso deve ser desenvolvido. Parece que, se aspiramos à importância social, então devemos desenvolver sistemas que permitam que nossos consumidores nos forneçam opinião sobre como nossas aplicações se relacionam com seus

valores, com seus reforçadores. Isto não é uma rejeição do nosso legado. Nosso uso de medidas subjetivas não está relacionado a variáveis causais internas. Em vez disso, é uma tentativa de acessar as dimensões de reforçadores complexos de modo socialmente aceitável e prático. Trata-se de um evento evolucionário que está ocorrendo em função das contingências do ambiente da pesquisa aplicada; contingências que nossos fundadores provavelmente diriam apreciar, se tivéssemos coragem de questioná-los a respeito de tais opiniões subjetivas sobre nosso comportamento. (p. 213)

Primeiras Considerações a Respeito da Proposta do Conceito de Validade Social

Kazdin (1977), para abordar a validade social, recém-proposta na ACA, descreveu técnicas que avaliam a relevância clínica ou aplicada da mudança do comportamento, revisou pesquisas que usam tais técnicas e discutiu potenciais problemas e limitações nos correntes métodos de avaliação. O autor enfatizou intervenções em ACA, apesar de reconhecer que esta não seja a única área que lida com a mudança clínica. Contudo, a ACA explicitamente tem tentado estabelecer métodos quantitativos para avaliar a relevância das mudanças.

Entre os vários critérios propostos para avaliar os efeitos de intervenções comportamentais, Kazdin (1977) destaca dois sugeridos por Risley (1970, citado por Kazdin, 1977) como necessários, em conjunto, para a avaliação de intervenções: o experimental e o terapêutico. O critério

experimental diz respeito ao que é necessário para estabelecer que uma intervenção é responsável pela mudança do comportamento e é característico tanto de pesquisas aplicadas quanto de pesquisas experimentais. Já o critério terapêutico é mais um critério que a pesquisa aplicada deve cumprir e exige determinar se a mudança no comportamento é clínica ou socialmente importante ou não.

Para Risley (1970, citado por Kazdin, 1977), o critério terapêutico refere-se a uma comparação entre a mudança do comportamento obtida com a intervenção e o nível de mudança requerido para a melhoria no funcionamento cotidiano do sujeito. O meio pelo qual o critério terapêutico é satisfeito é difícil de especificar, porque quem deve determinar o que é desviante ou aceitável são o sujeito e seus pares. Assim, a ACA vinha se resguardando de especificar critérios para avaliar a mudança (Kazdin, 1977).

O critério de validade social expande a avaliação das intervenções aplicadas para além dos critérios experimental e terapêutico propostos por Risley (1970, citado por Kazdin, 1977). Em termos gerais, a validação social se refere à avaliação da **aceitabilidade** de programas de intervenção. Várias facetas da aceitabilidade podem ser distinguidas. Avaliar a aceitabilidade do *foco* da intervenção significa verificar se os comportamentos selecionados são importantes para o indivíduo no ambiente natural. Pode-se avaliar a aceitabilidade, ou satisfação do consumidor, em relação ao *procedimento* para a escolha entre diferentes técnicas efetivas. Ainda é importante avaliar a aceitabilidade da *mudança comportamental* obtida com o tratamento (Kazdin, 1977).

Kazdin (1977) relata haver duas formas de avaliar a aceitabilidade da mudança comportamental: examinando a mudança em comparação à performance de pares “não desviantes” ou através das avaliações de indivíduos do convívio do sujeito. Assim, a validação dos efeitos da intervenção pode ser realizada de formas identificadas pelo autor como medidas de **comparação social** e medidas de **avaliação subjetiva**.

No método de **comparação social**, o comportamento do indivíduo antes e depois do tratamento é comparado com o comportamento de pares “não desviantes”. A questão, nesta comparação, é se o comportamento do cliente depois do tratamento é distinguível do comportamento dos pares.

A característica principal da comparação social é identificar os pares dos sujeitos, os indivíduos que são semelhantes aos sujeitos em variáveis demográficas, mas diferem na performance do comportamento-alvo. O grupo de pares pode ser usado de duas formas. Em uma das formas, um comportamento-alvo é avaliado em todos os indivíduos em uma dada situação para determinar qual comportamento pode ser considerado extremo. Outra forma de uso do grupo é tomar o nível do comportamento dos pares que não precisaram ou receberam tratamento como critério para a avaliação do sucesso do tratamento ou importância clínica. Desta forma, “Caso o tratamento tenha mudado acentuadamente o comportamento, a performance do sujeito estaria dentro do nível normativo de seus pares¹⁸.” (p. 432)

¹⁸ Foram respeitadas as colocações de Kazdin (1977) a respeito do método de comparação social a título de apresentar a definição deste como ela se originou na Análise do Comportamento. Apesar disso, não há concordância por parte da autora de que o fato de ter havido mudança acentuada do comportamento signifique, necessariamente, que o comportamento do indivíduo tenha entrado no nível normativo dos pares.

A **avaliação subjetiva** como meio de validação dos efeitos do tratamento consiste no julgamento de aspectos qualitativos do comportamento. O comportamento que foi alterado é observado por indivíduos que interagem com o sujeito ou estão em uma posição especial em relação a ele. O comportamento do sujeito é julgado pelos indivíduos, o que constitui uma avaliação global do funcionamento do sujeito e de sua *performance*. Tal tipo de procedimento de validação social vinha sendo usado com frequência na ACA para validar os efeitos dos tratamentos.

Ocasionalmente, como nota Kazdin (1977), estudos têm combinado os métodos de comparação social e avaliação subjetiva para avaliar os efeitos do tratamento.

O autor exalta o importante avanço representado pela validade social na avaliação da terapia, no estreitamento do hiato entre resultados obtidos da experimentação em condições restritas de laboratório e a relevância da mudança do comportamento em situações do dia-a-dia, na ponte entre a experimentação e o trabalho clínico e entre a pesquisa e a prática. Porém, alerta para o fato de obtenção de dados normativos e avaliações subjetivas serem enganosamente simples. A interpretação destes dados pressupõe várias questões destacadas por Kazdin (1977).

Sobre o uso de dados normativos, Kazdin (1977) discute duas questões: a inadequação de padrões normativos e a identificação do grupo normativo.

A respeito da inadequação de padrões normativos, Kazdin (1977) considera que níveis padrões ou normativos podem ser um critério inapropriado de avaliação de mudanças para muitos comportamentos, sendo que, para muitos programas, mudar o nível normativo deveria ser o objetivo. Em alguns

programas, o objetivo não é levar os repertórios dos indivíduos para o nível normativo, mas estabelecer um novo nível normativo que será mais desejável socialmente. Como exemplos, o autor menciona programas de reaproveitamento de material reciclável e de uso de transporte público. Mas, intervenções sociais não são os únicos casos em que níveis normativos são inapropriados para avaliar a importância aplicada dos efeitos do tratamento. Um caso ilustrativo é o de programas que visam o desempenho acadêmico, nos quais o uso de níveis normativos é questionável, já que o rendimento pode ser acelerado reprogramando o comportamento do professor e os currículos.

O ponto julgado mais importante por Kazdin (1977), no uso de níveis normativos de desempenho como critério de avaliação de mudança é a satisfação com estes. O autor considera que fazer deles objetivo do tratamento pode ser o maior suporte para o *status quo*, mas considera que, para algumas aplicações, o desempenho normativo deve ser o objetivo, assim como o critério de avaliação da mudança.

Muitos programas têm como foco indivíduos que poderiam funcionar em seus ambientes naturais caso seus comportamentos estivessem em conformidade com os padrões normativos. Porém, a maioria dos programas tem visado levar o comportamento dos indivíduos selecionados ao nível normal, ao invés de acelerar desempenhos medianos de comportamentos desviantes. Assim, segundo Kazdin (1977), níveis normativos de desempenho apenas seriam um ideal se, quando alcançados, forem motivo para celebração.

No tocante à identificação do grupo normativo, Kazdin (1977) afirma ser esta uma tarefa difícil para alguns clientes. Por exemplo, o nível normativo obtido de crianças não retardadas pode ser um critério muito alto para crianças

profundamente retardadas na mesma faixa etária, dependendo do comportamento-alvo.

Outro problema nesta questão é que, para muitas populações, pode ser arbitrário especificar quais indivíduos devem constituir o grupo normativo, o que seria o caso de uma população carcerária. Nestes casos, um grupo normativo externo poderia fornecer padrões irreais de avaliação em função das diferenças de formação, educação, socioeconômicas entre outras variáveis. Definir os pares como grupo normativo esconderia muitas variáveis que poderiam ser relevantes para avaliar os resultados do tratamento, pois o nível normativo dos pares pode estar relacionado a muitas variáveis individuais e demográficas. Levar em conta estas variáveis na definição dos pares poderia alterar o nível de comportamento considerado aceitável ou clinicamente relevante. Assim, a visão do impacto clínico do tratamento dependerá do padrão normativo usado para comparação.

Para Kazdin (1977), uma questão importante no uso de dados normativos é o reconhecimento da relatividade das normas e das variáveis que contribuem para os padrões normativos, o que leva à necessidade de especificação dos parâmetros usados para definir a população a qual o comportamento do cliente será comparado.

As considerações de Kazdin (1977) a respeito do uso de avaliações subjetivas abordam três questões: a seleção de comportamentos, a interpretação da avaliação subjetiva e avaliação da avaliação subjetiva.

Na seleção de comportamentos, a avaliação subjetiva do comportamento de indivíduos não envolvidos com o tratamento é valiosa. Devido à valorização de respostas claramente definidas na pesquisa

comportamental, existe a possibilidade de focalização em comportamentos selecionados arbitrariamente com pequeno ou nenhum impacto no funcionamento do cliente. A validação social ajudaria a estabelecer se os comportamentos influenciam a avaliação de indivíduos em situações cotidianas. A avaliação dos indivíduos de seu próprio comportamento é relevante na decisão do foco do tratamento, já que, freqüentemente, são responsáveis pela identificação do problema e por reportá-los aos profissionais para tratamento em ambiente natural (Kazdin, 1977).

Porém, o autor indica haver um problema potencial em lidar com as opiniões de não profissionais para determinação do foco do tratamento. Muitos comportamentos focados em intervenções aplicadas parecem ser respostas que são significantes para os indivíduos que interagem com os clientes, porém focos identificados por estes indivíduos podem não ser necessariamente o foco primário do cliente, e isso, por vezes, fica evidenciado nos resultados da pesquisa.

Em geral, as opiniões emitidas por não profissionais que interagem com os clientes podem ser valiosas para a avaliação do foco do tratamento e das mudanças produzidas por este, mas estas opiniões não devem ser vistas acriticamente como prescrições, como alerta Kazdin (1977). Para ele, o foco de tratamento deve ser freqüentemente ditado por considerações de pesquisa além do alcance da maioria de não profissionais e deve em relações empiricamente estabelecidas.

A questão principal que deve ser levantada em relação à interpretação da avaliação subjetiva, segundo Kazdin (1977), é se mudanças clinicamente importantes foram alcançadas. Mudanças qualitativas no comportamento

podem ser indicadas por avaliações subjetivas e globais após a demonstração de mudanças do comportamento, mas uma diferença no comportamento que pode ser percebida por avaliadores que fazem julgamentos globais não é necessariamente uma mudança importante. Classificações globais refletem apenas diferenças que possam ser notadas, mas não são baseadas em unidades quantitativas que definam necessariamente um funcionamento adequado. Então, não há garantias de que o desempenho pós-treinamento e o nível de mudança do comportamento seja importante ou adequado em estudos nos quais o desempenho após a intervenção é classificado como melhor ou superior ao que foi antes do treinamento.

Uma forma de assegurar que a avaliação do comportamento reflita uma mudança importante é fornecer tais avaliações para os clientes-alvo ou para pares que não tenham o problema, o que ampara as escalas de avaliação subjetiva em um critério normativo. Avaliações subjetivas de sujeitos-alvo sem relato de nível normativo de comportamento podem ser inadequadas como critério de avaliação de importância clínica e mudança de comportamento. Assim, avaliações subjetivas deixam o nível de desempenho que é necessário para a mudança clínica sem especificação (Kazdin, 1977).

Quanto à avaliação da avaliação subjetiva, Kazdin (1977) aborda a questão da *face validity* (validade de face)¹⁹. A suposição por trás desta é de que o nome do rótulo da dimensão corresponde ao que a escala realmente mede. A validade de face é o critério para seleção da dimensão através da qual o cliente é classificado. O constructo de interesse do investigador é o nome da

¹⁹ Face validity pertains to whether the test "looks valid" to the examinees who take it. The administrative personnel who decide on its use, and other technically untrained observers (Anastasi, 1988, p.144).

dimensão e assume-se que um ou alguns itens avaliam adequadamente a dimensão que está sendo classificada.

Para o autor, ao se voltar para a avaliação subjetiva, a ACA deve se beneficiar do conhecimento obtido da pesquisa em avaliação psicométrica, da qual pode ser extraída a conclusão fundamental de que o nome dado a uma medida deve ter pouca relação com o que a escala realmente mede. A validade de face, sozinha, como critério, não é adequada para validar uma ferramenta.

A construção de escalas ainda requer o reconhecimento do papel da característica e variância do método, a importância da validação convergente e discriminante, possíveis influências das tendências e estilos de resposta, o viés do observador, entre outras coisas. Kazdin (1977) considera que avaliações básicas desejáveis têm sido omitidas para estabelecer a credibilidade das escalas quando classificações subjetivas são usadas na ACA e que falhas em considerar requerimentos e avanços da avaliação psicométrica podem limitar grosseiramente o sentido que classificações subjetivas podem conceder.

Algumas considerações adicionais são feitas por Kazdin (1977). A validação social procura quantificar se uma mudança significativa foi alcançada com o tratamento, apesar de haver problemas específicos em determinar a importância clínica ou aplicada da mudança do comportamento. A constatação de que uma mudança de importância clínica foi obtida é uma noção relativa que requer a comparação entre os indivíduos visados e os outros (ou seja, uma comparação normativa) e julgamentos a respeito da extensão da mudança (uma avaliação subjetiva). Por isso, o autor julga que há necessidade de incorporar diretamente, às avaliações de tratamento, os critérios normativo e subjetivo.

Kazdin (1977) afirma que as pesquisas em ACA têm privilegiado algumas populações no que diz respeito à validação social, mas, em geral, ela tem sido pouco usada e algumas populações têm sido negligenciadas.

Para o autor, dados normativos são necessários para estabelecer o critério contra o qual o tratamento pode ser avaliado, mas as pesquisas têm abordado tal questão em situações selecionadas. Assim, estudos normativos deveriam ser completados com a diversidade de populações estudadas na modificação do comportamento. Dados normativos ainda são padrões relativos a diversas variáveis e, se eles são usados na avaliação dos efeitos do tratamento, as variáveis que limitam sua generalidade têm que ser explicitadas.

Outra área que também deve ser estudada, segundo Kazdin (1977), diz respeito às variáveis que influenciam avaliações subjetivas. A avaliação subjetiva como critério para julgamento das mudanças comportamentais é apenas um lado do problema. Ela também serve de base para seleção de indivíduos para tratamento e para identificá-los como desviantes. Independentemente dos comportamentos desempenhados pelos clientes após o tratamento, a avaliação do comportamento pelos outros é importante. Em muitas populações desviantes, o problema não está somente em seus comportamentos, mas também em como eles são percebidos pelos outros e por eles próprios. A percepção dos outros pode ser influenciada na medida em que os indivíduos identificados como desviantes acreditam ser desviantes, assim como o desempenho não desviante de indivíduos percebidos como desviantes não é o suficiente para sobrepor tal percepção. Assim, é possível que a mudança do comportamento do cliente não altere necessariamente as avaliações de indivíduos com os quais os clientes tenham interagido.

Algumas avaliações subjetivas de efeitos da intervenção usam indivíduos não familiares com os clientes e sua história pregressa. Tais indivíduos não podem ser influenciados por percepções prévias dos clientes, mas a percepção destes indivíduos não é necessariamente representativa das percepções dos indivíduos que têm contato contínuo com o cliente. Por isso, uma área da pesquisa em validação social é a determinação de quem faz a avaliação subjetiva e julga o sucesso do tratamento. O julgamento difere se os responsáveis pela avaliação são pares, estranhos, parentes ou os próprios clientes (Kazdin, 1977).

Kazdin (1977) ressalta a importância da validação social, tanto por meio de avaliações subjetivas como de comparações normativas, para complementação de dados de comportamentos abertos, já que considera a avaliação cuidadosa de comportamentos apresentados um retrato incompleto do que o tratamento realizou. Para o autor, a questão principal na integração de clientes a ambientes que não sejam de tratamento é quão bem o cliente se desempenha em relação aos outros e como ele é percebido pelos outros, e tal questão não pode ser abordada por simples medidas de comportamento aberto.

Apesar de problemas e limitações, elencadas por Kazdin (1977), relativas tanto ao uso de medidas normativas quanto de julgamentos subjetivos, o autor conclui em defesa de tais medidas e indica suas futuras direções.

A adoção de dados normativos e avaliação subjetiva para as observações comportamentais dos clientes não ocorrerão sem problemas. (...) Apesar de obstáculos metodológicos que

podem surgir na validação social, a avaliação do impacto clínico do tratamento fornece informação essencial. A validação social determina a eficácia do tratamento na solução do problema clínico e social mais diretamente do que faz a medida comportamental discreta sozinha. Para a maioria das técnicas de tratamento, esta questão não começou a ser tratada. O critério para avaliação de tratamento além das usuais medidas de resultado deve ampliar-se nos próximos anos para incluir medidas semelhantes como a aceitabilidade do tratamento pelo consumidor, eficácia do custo, facilidade de implementação e disseminação, entre outras. Em meio a isto está a prioridade na seleção de tratamentos que vão além de obter mudanças comportamentais discretas para afetar o funcionamento do cliente no ambiente natural (pp. 447-448).

Ainda nos primeiros anos da proposta da validade social na ACA, Van Houten (1979) propõe uma expansão de seus procedimentos. Segundo o autor, pouca atenção tem sido dada para o uso de procedimentos de validação social na determinação de níveis ótimos para comportamentos-alvo.

A seleção de comportamentos-alvo tem sido identificada como um problema importante na ACA. Van Houten (1979) menciona dois métodos cujas variações têm se mostrado úteis na solução de tal problema. No método de grupos conhecidos, são comparados os comportamentos dos indivíduos que compõem dois grupos diferentes em alguma dimensão global importante, para identificar quais comportamentos são exclusivamente relacionados com o maior sucesso do grupo. No método de satisfação do consumidor, clientes ou

outros consumidores são requisitados a julgar ou classificar vários comportamentos potencialmente importantes. Ambos os métodos visam confirmação experimental da importância dos comportamentos selecionados através do uso de procedimentos adicionais de validação social seguido do tratamento.

Primeiras Extensões da Proposta de Validade Social

Van Houten (1979) considera a seleção de comportamentos-alvo uma aplicação das técnicas de validação social valiosa e necessária, mas acredita que este seja apenas um passo do processo de avaliação como um todo. Outro passo, que tem recebido pouca atenção, é a seleção de níveis ótimos de desempenho para os comportamentos escolhidos na etapa mencionada anteriormente.

Para o autor, a maioria dos comportamentos exibe variabilidade, mas concebe-se uma “faixa” dentro de cujos limites o desempenho é mais adaptativo. Caso tais limites sejam desconhecidos, é possível que o tratamento termine quando o desempenho estiver antes ou depois dessa “faixa”. Terminar o tratamento abaixo deste limite pode ser um resultado indesejável. Porém, a validação social, tomada em seu modo tradicional, no qual é tomado como critério o nível normal dos indivíduos, não pode fornecer informação sobre a ocorrência ou não deste erro. Como exemplo, Van Houten (1979) lembra que, se um desempenho final for superior ao desempenho da linha de base, provavelmente será declarada uma validade social positiva. Mas tal avaliação não informará aos pesquisadores que níveis mais altos de desempenho poderiam levar a um comportamento mais adaptativo por parte do cliente.

Por outro lado, se o desempenho não for julgado melhor após o tratamento, não estará claro se o comportamento selecionado não era importante ou se ele não mudou o suficiente para alcançar resultados positivos (p. 582).

Caso o tratamento termine com o comportamento aumentado para além do limite superior de desempenho competente, o resultado também pode ser prejudicial para o cliente. Um exemplo considerado óbvio para Van Houten (1979) seria o de uma criança que está sendo ensinada a fazer mais solicitações. Um acréscimo no comportamento, até determinado ponto, deve ser julgado adaptativo, mas, a partir de tal ponto, o comportamento pode ser considerado indesejável por aqueles que interagem com a criança. Neste ponto, a frequência do comportamento da criança ultrapassa o valor ótimo, o valor superior do limite que corresponde ao nível ótimo.

Concluindo sua opinião sobre tais questões, o autor afirma que as técnicas tradicionais de validação social, aquelas descritas por Wolf (1978/1976) e Kazdin (1977), não informam se o comportamento é mais funcional no nível mais baixo nem se resultados negativos são resultantes da seleção dos comportamentos errados ou produzidos por mudanças muito pequenas ou muito acentuadas no nível de comportamento-alvo. Da mesma forma, não existem garantias de que os níveis normativos representem o desempenho ótimo. Alguns possíveis passos subseqüentes à determinação dos comportamentos-alvo, na opinião de Van Houten (1979), seriam a determinação da frequência visada para estes comportamentos, a identificação de situações nas quais estes comportamentos são julgados apropriados e a

identificação de várias interações entre estes comportamentos e sua influência sobre a eficácia do tratamento.

A falha em identificar os limites do comportamento adaptativo torna difícil, ou impossível, determinar quando mudar um comportamento para obter os melhores resultados para o cliente, e os profissionais não poderão dizer quando mudar um comportamento em particular e quando abandoná-lo. Sem esta informação, é difícil determinar se um tratamento que produz uma melhora socialmente válida no desempenho de um indivíduo também a produzirá no desempenho de outro, visto que alguém deve saber se este comportamento está em seu limite ótimo (Van Houten 1979).

Os profissionais necessitam de padrões socialmente válidos que possam ser tomados como meta para saber quando iniciar e terminar um tratamento. A disponibilidade destes padrões é especialmente importante para aqueles que não estão engajados na pesquisa aplicada, na opinião do autor, pois freqüentemente tais profissionais não estão em posição de desenvolver uma validação social complexa da mudança que produzem por conta de questões financeiras ou de tempo.

Van Houten (1979) sugere algumas formas de empregar procedimentos de validação social para determinar estes objetivos: procedimentos de seleção com base normativa e procedimentos de seleção com base experimental.

A seleção de padrões de competência através de procedimentos com base normativa é uma das formas apresentadas por Van Houten (1979) para determinar objetivos socialmente válidos. Consiste em avaliar o desempenho de indivíduos julgados altamente competentes em alguma área de interesse. As normas obtidas a partir do comportamento destes sujeitos serviriam como

objetivos para o desempenho de indivíduos menos competentes. Para isso, ambos os métodos mais tradicionais de validação social seriam usados, já que os indivíduos devem ser selecionados por serem julgados competentes e o nível de seu desempenho deve ser determinado.

Uma decisão a ser tomada neste método é a de escolher entre obter um único objetivo universal para cada tarefa do indivíduo mais competente possível de ser encontrado ou obter objetivos separados de cada população especial. Para o autor, o primeiro método é provavelmente melhor, pois, apesar de parecer excessivamente rigoroso para algumas populações, os objetivos estabelecidos a partir dele podem servir como um padrão em preparação para o qual as mudanças produzidas por qualquer tratamento podem ser avaliadas. Além disso, eles têm a vantagem de servir como um desafio para o desenvolvimento de técnicas mais poderosas.

Para Van Houten (1979), não se conseguir levar o comportamento do indivíduo ao nível competente não implica, necessariamente, uma falha experimental, já que certamente é importante levar os indivíduos o mais próximo possível de um nível ótimo de desempenho. Entretanto, os pesquisadores devem permanecer atentos à possibilidade de produzir níveis de desempenho cada vez mais adaptativos. É possível evitar esta questão obtendo competências separadas baseadas em normas de toda a população, o que provavelmente é imprudente, já que níveis de comportamento que indicam competência não mudam porque alguns indivíduos têm maior dificuldade em obtê-los. Todavia, pode ser prático estabelecer objetivos menores a curto prazo em algumas ocasiões porque podem não existir técnicas à disposição para levar o desempenho do indivíduo ao nível competente.

Uma vantagem indicada por Van Houten (1979) no uso de padrões baseados no desempenho de indivíduos competentes é que esta prática não limita o potencial de melhora. Normas obtidas de pares não tratados podem ser melhoradas de várias formas, o que não ocorre com normas obtidas de indivíduos competentes, que podem ser apropriadas para acelerar o desempenho de um grupo inteiro a seu nível.

Embora normas baseadas no desempenho de indivíduos competentes também terem algumas das desvantagens de normas de grupo, elas não devem ser consideradas inflexíveis, porque estudos podem indicar que os mais altos níveis de competência são possíveis. O autor considera que, tanto a seleção de comportamentos-alvo quanto de padrões de desempenho, podem ser vistos como um processo iterativo no qual repetidas validações sociais aplicadas a determinado problema levam à melhora do delineamento de objetivos de tratamento importantes.

Van Houten (1979) descreve o método para obtenção de padrões de indivíduos competentes e as questões envolvidas nela. Segundo o autor, o método é simples e muito usado em algumas áreas. Primeiro, ele alerta que ser capaz de desempenhar uma tarefa não torna um indivíduo proficiente nela. Proficiente seria apenas aquele que é capaz de executar a tarefa em um nível alto o suficiente para ser útil para os outros ou para si próprio.

Caso selecionemos padrões para proficiência na tarefa examinando o desempenho de adultos escolhidos de forma aleatória, é improvável que os padrões obtidos representem o desempenho maximamente útil. Para ter certeza de que os padrões refletem competência e que podem ser visados por uma tecnologia comportamental poderosa devemos tomar como amostra o

desempenho dos indivíduos classificados como mais competentes. Uma forma de selecionar os indivíduos competentes é perguntar àqueles que requerem o comportamento qual o nível deste que julgam funcional (Van Houten, 1979).

Aplicar padrões de desempenho de adultos a crianças pode parecer extremamente rigoroso, mas, segundo o autor, a competência independe da idade. Deve haver um número reduzido de adultos e crianças competentes em determinada tarefa, o que não implica que o objetivo final para o desempenho de crianças seja de alguma forma diferente do objetivo final para adultos.

Alguns indivíduos podem ter dificuldades em obter estes padrões, mas movimentos na direção dos padrões podem ser bem sucedidos. A aplicação de padrões rigorosos pode nos dar a idéia de como abordar o nível de desempenho de um modelo competente e, assim, saber o quanto nossa tecnologia deve ser aprimorada. Van Houten (1979) ainda defende o uso do comportamento competente como padrão já que, caso derivemos padrões de desempenho de dados normativos baseados em médias da população, correremos o risco de trabalhar abaixo do nível competente, o que provavelmente levaria o tratamento ao término antes que os indivíduos se tornassem competentes. O uso do comportamento competente como padrão nos levaria a desenfaturar parte da relatividade envolvida na seleção de normas porque nos fornece objetivos julgados funcionais.

Para selecionar é necessário decidir quais indivíduos são competentes na tarefa. O primeiro passo a ser dado é perguntar aos mais envolvidos sobre classificações de adequação dos objetivos de desempenho dos indivíduos. Para o autor, pesquisas e validações sociais posteriores devem ajudar a estabelecer quão bons são os padrões obtidos, reafirmando sua posição a

respeito da estratégia de aplicações sucessivas de validação social para refinamento dos padrões.

Outra questão envolvida na escolha de comportamentos competentes é a seleção da medida apropriada. Caso usemos uma medida inadequada é possível que não se estabeleça nenhuma relação com a utilidade social do desempenho. Algumas medidas gerais podem constituir uma simplificação exagerada, como seria o caso da seleção da porcentagem de acertos de problemas matemáticos para avaliação da competência de estudantes na resolução de tais problemas, pois ela não nos permitiria diferenciar a competência de um indivíduo capaz de solucionar 80 problemas por minuto com 100% de acerto da competência de outro capaz de resolver apenas 20 problemas por minuto, também com 100% de acerto (Van Houten, 1979).

Isto não implica, segundo o autor, a impossibilidade da obtenção de padrões objetivos de competência, mas pode significar que devemos especificar nossas variáveis de tal forma que estejam relacionadas às medidas de validação social. No caso da resolução de problemas matemáticos, é clara a necessidade da adição da taxa de problemas solucionados como medida de competência. Se pudermos identificar todos os aspectos relevantes do comportamento em questão, então será possível localizar consequências de cada aspecto e trazê-los para o nível normativo de indivíduos competentes.

Apesar de as dimensões relevantes de muitos comportamentos serem demasiado complexas, para Van Houten (1979), ainda é possível caminhar ao encontro do estabelecimento de padrões de desempenho devido à natureza iterativa do processo de validade social. Cada estudo sucessivo pode levar a uma aproximação maior dos aspectos e padrões relevantes.

A outra forma de estabelecimento de objetivos válidos, mencionada por Van Houten (1979), é a *seleção de padrões de competência através de procedimentos experimentalmente embasados*. O autor descreve dois métodos de avaliação empírica da validade social de objetivos. No primeiro método, o experimentador manipula o comportamento de interesse sobre toda sua “faixa” de ocorrência e determina em que valor o comportamento é maximamente útil ou efetivo em atingir objetivos funcionais importantes. O que diferencia este procedimento de validação social do mais tradicional é que se faz com que o comportamento de interesse assuma diferentes valores para que seja selecionado aquele que representa um desempenho ótimo. Tal abordagem é considerada pelo autor altamente compatível com aquela empregada nos procedimentos com base normativa. Validações sociais que variam experimentalmente a frequência de comportamentos considerados importantes, por serem consistentemente encontrados no repertório de indivíduos julgados competentes em determinada tarefa, podem lidar com a validação empírica destes comportamentos. Como poucos pesquisadores têm desenvolvido validações experimentais de padrões de competência, existem poucos exemplos desta abordagem.

Um estudo que ilustra o uso deste método é o de Warren, Rogers-Warren e Baer (1976, citado por Van Houten, 1979). Os autores queriam aumentar as respostas de oferta de compartilhamento de material entre crianças de pré-escola e concluíram que as ofertas seriam mais apropriadas quando aceitas e haveria uma inadequação funcional quando um grande número de ofertas fosse rejeitado. Eles suspeitavam que um fator que poderia influenciar a probabilidade de aceitação da oferta era a taxa da resposta. Em

dois experimentos foram aumentadas as taxas de oferta e examinada a porcentagem de ofertas aceitas. Como resultado, descobriram que a porcentagem de ofertas aceitas diminuía com o aumento da frequência de ofertas, correspondendo a uma função em forma de “U” invertido, na qual a aceitação aumenta até um ponto em que começa a declinar. A conclusão dos autores é de que, se a frequência de dado comportamento determina quanto a reação dos indivíduos será positiva ou negativa, a taxa-alvo ou os objetivos devem ser escolhidos de acordo com estes resultados. Para Van Houten (1979), tal estudo ilustra como os objetivos de tratamento devem ser precisamente especificados, de forma a não se correr o risco de aumentar a taxa de um comportamento inicialmente efetivo a uma taxa menos útil socialmente que a da linha de base.

Da mesma forma, comportamentos relacionados à comunicação podem ser avaliados pelo seu efeito sobre a audiência, como sugerido por Van Houten (1979). Por exemplo, pesquisadores podem variar o contato visual e manter todas as outras variáveis constantes caso queira examinar os efeitos desta variável sobre a audiência. Os clientes podem ser instruídos a manterem-se dentro de determinado limite de duração de contato visual caso seja demonstrado que este atue em um nível ótimo. Resultados comparáveis podem ser facilmente obtidos sobre a altura da voz e outros fatores que possam influenciar a eficiência das pessoas em se comunicar em público. A análise de outras variáveis pode ajudar a identificar muitos efeitos e interações importantes. Não se espera que apenas um pesquisador possa variar todas as condições suspeitas. No entanto, decorrido o tempo, mais conhecimento deverá ser acumulado por séries de replicações sistemáticas. O resultado final

desta operação deve ser a obtenção de objetivos de desempenho empiricamente determinados. A dificuldade em selecionar objetivos desta forma reside no custo de tempo e de esforços associados à variação de comportamentos ao longo de todo seu limite de ocorrência. Mesmo assim, no fim das contas, este problema é compensado pelo grau de confiança dos padrões obtidos pelo método.

Outra forma de avaliar a adequação dos padrões de desempenho baseando-se em experimentação é examinar os efeitos do treinamento de vários padrões na aquisição subsequente de competências mais complexas (Van Houten, 1979). Por exemplo, Ayllon e Azrin (1968, citado por Van Houten, 1979) sugeriram que os comportamentos ensinados deveriam ser funcionais para os indivíduos em seu ambiente quando reintegrados à comunidade. Fixsen, Phillips, Phillips e Wolf (1976, citado por Van Houten, 1979), realizaram um estudo em que avaliavam a validade a longo prazo dos comportamentos-alvo que selecionaram com jovens pré-delinquentes examinando efeitos posteriores no desempenho escolar e recidivas. Eles encontraram, como evidência da validade dos comportamentos escolhidos, acentuadas reduções nos contatos com a polícia além de os resultados positivos em vários aspectos do desempenho escolar.

Van Houten (1979) usa um exemplo de como o problema da seleção do objetivo pode influenciar um programa de pesquisa desenvolvido para ensinar crianças a escrever histórias melhores e que também ilustra como o processo de validação social desenvolve padrões através de aproximações sucessivas.

O autor menciona que alguns artigos publicados anteriormente ao seu sugerem que as classificações de melhora dos juízes são maiores quando há

um aumento no uso de palavras que denotam ação [action words]. Porém, Van Houten não foi capaz de replicar tal resultado em alguns estudos, pois o aumento obtido no uso de palavras que denotam ação não mudou as classificações dos juízes (Van Houten citado por Van Houten, 1979; Van Houten e McLellan citado por Van Houten, 1979). Em ambos os estudos, optou-se por trocar a taxa de palavras que denotam ação pela porcentagem destas, já que a última medida não se confundiria com o número de palavras escritas. Isto tornou claro que o tratamento aumentava a porcentagem de palavras que denotam ação, mas tal aumento não provocava mudança na qualidade atribuída pelos juízes às histórias. Então, julgou-se interessante comparar os níveis de palavras que denotam ação das histórias com níveis normativos baseados em padrões de competência obtidos. Comparando as porcentagens de uso de palavras que denotam ação pelos estudantes com o padrão de competência obtido de contos de escritores renomados, constatou-se que os níveis dos estudantes estavam dentro dos níveis de competência na linha de base. Isto sugeriu que o aumento no uso de palavras que denotam ação não produziu mudança na qualidade da história pois ele já estava em seu nível ótimo antes do tratamento e que, nos estudos anteriores, os alunos deviam estar abaixo do nível ótimo na linha de base e entraram no nível ótimo após o tratamento, o que melhorou a classificação dos juízes sobre seu texto.

Com isso, Van Houten (1979) pretende ilustrar como padrões de desempenho podem evoluir de forma a ajudar a prever se um estudante é ou não beneficiado por um tratamento desenvolvido para aumentar o nível de um comportamento. No entanto, o aumento da porcentagem de uso de diferentes palavras que denotam ação em uma história pode lidar com o aumento de uma

medida socialmente válida se os estudantes estiverem abaixo do nível competente e provavelmente não beneficia aqueles que já respondem dentro das normas. Por isso, o conhecimento destas normas permite determinar com mais cuidado quando um tratamento terá um efeito socialmente válido.

Van Houten (1979) conclui seu artigo defendendo a utilidade, ilustrada nos exemplos anteriores, da validação social vista como um processo iterativo que, quando empregada apropriadamente, ajuda pesquisadores a focarem nos melhores objetivos de tratamento juntamente com metas ou níveis do comportamento selecionado apropriados. Este processo também limita as situações específicas em que este comportamento é mais apropriado. Tal processo tem como resultado a seleção dos comportamentos-alvo com competências visadas expressadas em termos de frequência ótima, duração, porcentagem de acuracidade ou latência e o delineamento do *setting* mais apropriado para que estes comportamentos ocorram.

O autor descreve uma possível seqüência de passos deste processo que compreende avaliações normativas, julgamentos subjetivos e validações a curto e longo prazo.

O processo freqüentemente começa com a seleção de comportamentos e estimativa de níveis competentes destes comportamentos através da reunião de dados normativos dos indivíduos mais competentes ou comparando o comportamento normativo de grupos conhecidos de indivíduos competentes ou incompetentes (Mc Fall, 1976). Depois, os pesquisadores obtêm aproximações cada vez mais acercadas dos melhores comportamentos-alvo e padrões através dos dois tipos de

procedimento de validação social experimental. Primeiro, comportamentos que se acredita serem importantes são variados de um extremo a outro de seu limite natural de ocorrência e o efeito no ambiente e as classificações de juízes são anotadas. Segundo, efeitos globais ou a longo prazo da mudança dos comportamentos-alvo específicos são anotados (p. 590).

Van Houten (1979) ainda ressalta que os aspectos experimentais do processo de validação social permitem que os pesquisadores convirjam a aproximações mais rigorosas de uma guia que especifique quais comportamentos são bons candidatos à mudança assim como quando e quanto estes comportamentos devem ser mudados.

Posteriormente, outros autores propuseram extensões metodológicas dos procedimentos de avaliação de validade social. Winett *et al.* (1991) descrevem um método de avaliação da importância do problema empregando dados epidemiológicos e do impacto comportamental da intervenção por meio de estudos de eficácia e eficiência. Hawkins (1991) argumenta que o termo “validade social” não seria o melhor designativo das questões abordadas sob tal rótulo e propõe o termo “validade habilitativa” que consiste em examinar a contribuição dos objetivos de tratamento para o desempenho funcional ou prejuízo na vida do cliente, ou seja, consiste na seleção de objetivos de tratamento pela sua maximização de resultados positivos a curto e longo prazos e minimização dos custos para o cliente e outros membros do seu ambiente, o que é avaliado através do estabelecimento de relações entre objetivos de tratamento e resultados adaptativos ou mal adaptados. Kennedy

(2002) ainda sugere que o grau em que os ganhos obtidos com o tratamento são mantidos através do tempo em contextos naturais seja considerado o principal indicador de validade social.

Rumos Posteriores do Conceito de Validade Social

Como proposto por Kazdin (1977), as definições e usos de validade social estão em processo de expansão de suas definições iniciais. Contudo, as definições originais fornecidas por Kazdin (1977) e Wolf (1978/1976) continuaram sendo investigadas. Como Van Houten (1979), vários autores ainda têm apelado por avaliações ampliadas e/ou revisões metodológicas das medidas de validação social (Kennedy, 1992). Para o autor, “A tática de desenvolver medidas que reflitam mais sensivelmente as mudanças comportamentais em contextos sociais pode propiciar uma direção alternativa importante para a expansão ou entendimento da validade social.” (p. 154).

Geller (1991) notou que os autores que discutem a questão da validade social o fazem de diferentes perspectivas. No entanto, “(...) é evidente que o conceito de validade social é criticamente importante para qualquer tentativa de influenciar os comportamentos dos indivíduos, grupos e comunidades inteiras.” (p. 180).

COMPONENTES DA VALIDADE SOCIAL E DIMENSÕES DO CONCEITO

Componentes da Validade Social, seus Métodos e seus Objetivos

Foster e Mash (1999) destacam 3 dos que nomeiam **componentes** de validade social, ou seja, 3 aspectos da pesquisa aplicada que podem ser avaliados por procedimentos de validação social: *objetivos*, *procedimentos* e *resultados* de tratamento.

Segundo os autores, os dois **métodos** mais tradicionais de avaliação de validade social destes componentes são a *comparação normativa* e a *avaliação subjetiva*. Alguns trabalhos ainda propõem *extensões metodológicas* dos métodos convencionais para a avaliação da validade social.

A validação social, de cada um destes componentes, pode ter dois **objetivos**²⁰ diferentes, porém correlacionados. O primeiro objetivo é a demonstração da *importância* e da significância dos objetivos ou resultados na vida do cliente. O outro objetivo é tratar a *aceitabilidade* dos objetivos, procedimentos ou resultados para os consumidores relevantes da intervenção.

Objetivo de Tratamento como Componente de Validade Social

Para Foster e Mash (1999),

Objetivos de tratamento são cruciais porque definem a natureza do âmago das dificuldades que o tratamento planeja abordar, ditam as medidas de avaliação dos resultados e fornecem a fundamentação conceitual e empírica da intervenção (p. 309).

²⁰ “Objetivos”, aqui, não devem ser confundidos com os objetivos da pesquisa aplicada como componente de validade social. Neste contexto, “objetivos” refere-se aos objetivos da avaliação de validade social.

Dois aspectos dos objetivos do tratamento podem ser examinados pela abordagem da validade social: a **aceitabilidade social** dos objetivos do tratamento por vários consumidores da intervenção e a **importância social** destes objetivos (Foster e Mash, 1999).

Para discutir a aceitabilidade e a importância dos objetivos do tratamento, Foster e Mash (1999) recorrem ao trabalho de Rosem e Proctor (1981, citado por Foster e Mash, 1999) que diferencia dois tipos de resultados de tratamento. Segundo Foster e Mash (1999), tal diferenciação é relevante para os objetivos do tratamento e tem implicações para o tipo de avaliações de validade social que são apropriadas.

Como descrevem Foster e Mash (1999), *objetivos terminais* referem-se aos problemas que levam o cliente a procurar tratamento. Eles são determinados com a colaboração do cliente. Sua avaliação pode ser feita pelo próprio cliente ou por indivíduos importantes em sua vida.

Em contraste, *objetivos instrumentais* podem ser definidos como resultados desejados que conduzem a resultados terminais sem intervenção ulterior, ou seja, resultados obtidos através da intervenção que levam ao resultado final desejado quando o cliente responde fora do ambiente terapêutico. Eles são determinados pelo terapeuta, que avança a hipótese de que a realização destes objetivos é necessária para atingir os objetivos terminais. São guiados por teoria ou dados e devem ser definidos por sua relação causal ou preditiva com os objetivos terminais

Os autores argumentam ser importante examinar a aceitabilidade e a importância dos dois tipos de objetivos, porque todo comportamento indesejável que o cliente pretende diminuir no tratamento tem um

comportamento incompatível ou conjunto de circunstâncias que deveriam ser aumentadas para “tomar o lugar do problema” (p. 310), mas a aceitação e a importância da diminuição do comportamento indesejável não implica, necessariamente, a aceitação e importância do aumento do comportamento incompatível.

Um dos métodos mencionados por Foster e Mash (1999) para avaliação de validade social dos objetivos do tratamento é a **comparação normativa**. Ela consiste na comparação de níveis de um problema-alvo em relação à sua distribuição em uma dada população. Assim, “níveis que saem da escala ‘normal’ são usados para justificar a importância dos objetivos instrumentais e terminais” (p. 310).

A abordagem normativa verdadeira compara o desempenho do indivíduo com uma amostra da comunidade randomicamente selecionada, mas o método mais comum consiste no uso de “micronormas” locais para a comparação.

Normas locais são estabelecidas coletando dados dos indivíduos do ambiente imediato do cliente. O investigador então compara o resultado do cliente com estes dados ou usa estes dados para selecionar participantes para a intervenção (p. 310).

Segundo os autores,

Comparações normativas são mais úteis quando um problema ocorre ao longo de um continuum e o desvio dos níveis normativos se mostrou (através de meios empíricos) problemático (p.310).

Assim, a abordagem da comparação normativa pode fornecer alguma justificativa para que as dificuldades do cliente mereçam tratamento.

Porém, a abordagem normativa tem algumas limitações destacadas pelos autores. Ela não trata diretamente a aceitabilidade social dos objetivos de tratamento nem o sofrimento causado pelos problemas-alvo para o cliente e para outros em seu ambiente. Também não mostra como os níveis normativos prejudicam o funcionamento do dia-a-dia do indivíduo.

Algumas questões práticas e conceituais também limitam o uso de dados normativos para estabelecer a importância dos objetivos terminais. Há dificuldades em encontrar dados normativos apropriados disponíveis e em marcar um ponto de corte para o que constitui a “escala normal” ou limite de desempenho aceitável. Outra questão que deve ser considerada é a de que a seleção do grupo de comparação normativa requer cuidado, pois este deve ser representativo dos membros do ambiente no qual o cliente é solicitado a funcionar. Um cuidado adicional exigido por tal tipo de procedimento é a obtenção de consentimento daquele que fornece os dados, o que pode se tornar proibitivo quando envolve um grande número de clientes. O procedimento também requer do investigador o estabelecimento de critério para a seleção de indivíduos que irão contribuir para as normas locais, a determinação do número de indivíduos necessário para fornecer uma amostra adequada e a decisão de quando estes indivíduos terão desempenhos “normais” ou “exemplares”. Um último contratempo do uso de comparações normativas, listado por Foster e Mash (1999), é que as medidas usadas para selecionar membros do grupo normativo devem ser válidas e confiáveis.

O uso de normas também tem recebido críticas. Pode-se argumentar que alguns objetivos de tratamento podem ser justificados em bases legal e ética, independentemente de seu caráter normativo. Além disso, o fato de o

objetivo terminal ter sido estabelecido com base na norma não garante que o comportamento resultante seja adaptativo, e vice-versa. Finalmente, fato de o desempenho enquadrar-se na norma não é necessariamente uma indicação dos níveis de desempenho sejam “ideais” e “competentes” (Foster e Mash, 1999).

De acordo com os autores, níveis normativos também são insuficientes para estabelecer a importância dos objetivos instrumentais que é independente da realização dos objetivos terminais. A importância daqueles objetivos é melhor determinada empiricamente, ligando-os longitudinalmente e experimentalmente aos objetivos terminais. Quando diferentes objetivos instrumentais podem produzir mudanças nos objetivos terminais, aqueles podem ser selecionados considerando a efetividade, eficiência e aceitabilidade social relativas das alternativas para realização dos objetivos terminais.

A segunda abordagem que examina a validade social de objetivos de tratamento, descrita por Foster e Mash (1999), é a **avaliação subjetiva**. Ela consiste na avaliação, por parte de alguns consumidores dos serviços de tratamento, de um ou mais objetivos quanto à sua importância ou aceitabilidade.

Como recurso alternativo, às vezes é usada na formulação de uma intervenção.

Neste caso, o investigador usa dados fornecidos por clientes ou membros da comunidade relevante para construir os objetivos de tratamento, assim como documentar a importância ou aceitabilidade dos objetivos já estabelecidos. A suposição por trás dessa abordagem é a de que ela confere importância social

ou aceitabilidade nos objetivos de tratamento ao envolver membros relevantes da comunidade social (p. 310).

Exemplos desta abordagem incluem selecionar clientes que não relatam desconforto a respeito do problema investigado ou que não apresentam prejuízo no funcionamento e entrevistar membros de uma equipe hospitalar sobre comportamentos que incomodam os clientes que devem ser alvo de tratamento. Outro exemplo é a abordagem da combinação de modelos de Cone, que consiste na coleta de dados de potenciais consumidores do comportamento do cliente. Tais dados avaliam especificamente as expectativas de desempenho que o cliente será requisitado a atingir na situação ou ambiente no qual este deseja funcionar. Esta informação é usada pelo investigador para criar um modelo de comportamentos desejáveis e indesejáveis para ambientes e situações particulares, que se tornam os alvos de avaliação e tratamento (Foster e Mash, 1999).

Em outros casos, segundo os autores, “a avaliação subjetiva dos objetivos envolve coleta de dados formais e pode envolver requisitar clientes ou outros [membros da família/comunidade] para a avaliação de potenciais objetivos do tratamento” (p. 311). Tal abordagem é comum quando os sujeitos são crianças e indivíduos deficientes, já que estes não costumam estar envolvidos na seleção de objetivos. Por este meio, pode-se avaliar tanto aceitabilidade quanto importância de componentes de validade social, dependendo do conteúdo da avaliação e da população de consumidores que a completam. Questões que perguntam sobre importância pessoal, sofrimento ou distúrbios causados pelos problemas-alvo podem estabelecer a importância dos objetivos terminais para procurar tratamento. Pedir uma classificação da

importância dos objetivos de tratamento podem fornecer informação sobre as crenças leigas depositadas nos mais diretamente relevantes para a aceitabilidade do objetivo para pessoas no ambiente social do cliente.

Avaliações subjetivas, do ponto de vista de Foster e Mash (1999), em geral não são diretamente relevantes para estabelecer a importância de muitos objetivos instrumentais, que deveriam ser determinados através de investigação científica. Ainda é possível argumentar que

(...) perguntar aos clientes ou colocar pessoas para avaliar a importância dos objetivos instrumentais é equivalente a avaliar a extensão em que sua opinião corresponde ao conhecimento científico. Procedimentos que pedem para experts em áreas não psicológicas identificar e operacionalizar objetivos de tratamento, no entanto, podem ser úteis quando dados científicos não estão disponíveis (p. 311).

Entretanto, avaliações subjetivas são importantes para estabelecer a aceitabilidade dos objetivos instrumentais para clientes e pessoas do ambiente do cliente.

Um objetivo terminal pode ser aceitável, mas o objetivo instrumental poderá não sê-lo. Além disso, avaliações subjetivas dos objetivos instrumentais podem revelar circunstâncias nas quais o cliente (ou outros) não acreditam que os objetivos instrumentais irão produzir os objetivos terminais e podem realçar áreas nas quais os clientes precisam ser educados sobre como os passos da terapia são prováveis de levar à

resolução das dificuldades que levaram o cliente ao tratamento
(p. 311).

Uma questão-chave no uso de avaliações subjetivas está na determinação de quem é o consumidor relevante dos serviços de saúde mental, que foi tratada por Schwartz e Baer (1991, citado por Foster e Mash, 1999) e é relatada por Foster e Mash (1999). Segundo os autores, Schwartz e Baer (1991, citado por Foster e Mash, 1999) teriam descrito quatro grupos de consumidores. O primeiro seria o *consumidor direto*, que é aquele que recebe diretamente a intervenção. O segundo grupo é o de *consumidores indiretos*, que arcam com o ônus do tratamento de um cliente ou são afetados de alguma forma importante pelo tratamento. *Membros da comunidade imediata* formam o terceiro grupo e são aqueles que interagem com os consumidores diretos em uma base regular mas é menos envolvido que o consumidor indireto. O último grupo compreende os *membros da comunidade indireta*, que inclui indivíduos não conhecedores do cliente mas que vivem na mesma comunidade.

Alguns grupos são mais úteis para efeitos de uma validação social particular que outros. Por exemplo, na avaliação da importância dos objetivos terminais do tratamento, será melhor identificar os consumidores mais afetados pelas dificuldades do cliente e solicitar deles uma avaliação dos objetivos. Por outro lado, se avaliação subjetiva for usada para aferir a aceitabilidade do tratamento, devem ser identificados aqueles na comunidade que facilitarão ou impedirão a implementação do programa ou tratamento. Assim, de acordo com Foster e Mash (1999), avaliações de potenciais clientes, de outros indivíduos significantes na vida do cliente, de profissionais prováveis de implementar o

tratamento e de interlocutores da comunidade serão importantes na determinação da possibilidade de implementação e aceitação do tratamento.

Algumas objeções têm sido expressas sobre o uso de avaliação subjetiva para estabelecer a importância e a aceitabilidade dos objetivos de tratamento. Uma delas é a de que a avaliação subjetiva tende a ser expressa em termos verbais e esse conteúdo pode não corresponder exatamente à adesão do cliente ao programa ou sua permanência nele.

Outro embargo às avaliações subjetivas reside no argumento de que, mesmo se tratando de uma avaliação subjetiva a determinação da significância do objetivo para a vida do cliente, cabe mais a ele próprio do que aos membros da comunidade; por outro lado, se a questão disser respeito à importância do objetivo para a saúde mental do cliente, então trata-se de um problema de julgamento profissional.

Membros da comunidade podem ser ou não bem informados e podem basear seus julgamentos em muitos fatores que não a informação relevante. Isto é particularmente verdade quando a avaliação subjetiva é usada para julgar a importância de objetivos instrumentais, que são melhor determinados por profissionais experts usando seu conhecimento das bases científicas das dificuldades do cliente e sua expertise profissional (pp. 311-312).

A terceira dificuldade no uso de avaliações subjetivas, relatada por Foster e Mash (1999), está relacionada à necessidade de os investigadores selecionarem cuidadosamente o grupo de consumidores de modo a evitar

vieses a favor de avaliações positivas dos objetivos da intervenção. Por isso, seus procedimentos de amostragem devem ser claramente descritos.

A última crítica descrita pelos autores refere-se ao fato de, apesar de julgamentos da comunidade terem valor devido ao seu embasamento na história e no contexto do problema, algumas vezes podem ocorrer conflitos entre as opiniões da comunidade geral e opiniões da comunidade profissional sobre a importância social.

Nestes casos, argumenta-se que a informação da validade social é capaz de sinalizar uma necessidade de alterar opiniões da comunidade e padrões. Esta informação pode, contudo, ser muito importante para antecipar resistência aos procedimentos de tratamento por pelo menos alguns membros da comunidade (p. 312).

O propósito principal de obter indicadores de aceitabilidade é avaliar se os objetivos do tratamento condizem com os valores da comunidade, mas, admitem os autores, definir a “comunidade” não é fácil em sentido algum.

(...) a sociedade é composta por subgrupos identificáveis e heterogêneos , alguns dos quais têm pontos de vista diametricamente opostos sobre importância social e aceitabilidade de objetivos de tratamento (p. 312).

Avaliações de aceitabilidade de objetivo devem claramente identificar o grupo consumidor a ser avaliado e o porquê de o grupo ser relevante para avaliar a questão. Uma situação em que a validade social é particularmente importante, e na qual os investigadores devem ser especialmente sensíveis a estas questões, é o momento em que tratamentos são usados com populações

diversas daquelas nas quais uma intervenção foi originalmente estudada. O investigador deve também seguir procedimentos planejados para assegurar amostragem representativa do grupo e prover informação descritiva sobre características dele quando objetivos são avaliados por outros membros da comunidade que não seja o cliente ou um membro claramente identificado como importante no ambiente do cliente. Satisfazer os requisitos psicométricos, como validade de conteúdo e fidedignidade, é outra qualidade indispensável dos instrumentos usados para avaliar importância ou aceitabilidade dos objetivos.

Segundo Foster e Mash (1999), por causa das limitações da avaliação subjetiva, Hawkins (1991, citado por Foster e Mash, 1999) sugeriu que investigadores adotassem uma abordagem alternativa para examinar a importância dos objetivos de tratamento, criando, assim, uma **extensão metodológica** para a avaliação de validade social de objetivos de tratamento.

A abordagem que Hawkins (1991, citado por Foster e Mash, 1999) chamou de validade habilitativa, consiste em examinar a contribuição dos objetivos de tratamento para o desempenho funcional ou prejuízo na vida do cliente, ou seja, consiste na seleção de objetivos de tratamento pela sua maximização de resultados positivos a curto e longo prazos e minimização dos custos para o cliente e outros membros do seu ambiente. O principal método de demonstração de validade habilitativa é o estabelecimento de relações entre objetivos de tratamento e resultados adaptativos ou mal adaptados. Mas, ele também tem algumas dificuldades, como a necessidade de operacionalizar e avaliar resultados importantes a curto e longo prazo, associados aos objetivos de tratamento; a necessidade de reconciliar circunstâncias nas quais um

objetivo em particular produz resultados conflitantes, alguns negativos, alguns positivos e a impossibilidade de assumir necessariamente *a priori* se alguns resultados são “adaptativos” ou não, necessitando os próprios resultados de investigação para sua validade social (Foster e Mash, 1999).

Alguns autores recomendaram recentemente atenção para atividades e requisitos do funcionamento do dia-a-dia do cliente. Muitos nomearam implicitamente uma variedade de áreas como universalmente significativas e importantes para o cliente, indiferentemente de constituírem objetivos terminais ou instrumentais do tratamento. Exemplos destas áreas são a restrição ou acesso a vários ambientes que permitem autonomia e desenvolvimento, a qualidade de relação com a família e com pares e o custo monetário dos problemas do cliente. Porém, os autores certificam que nenhum estudo sozinho é capaz de avaliar todas essas áreas e que nenhuma delas é igualmente relevante para todos os tipos de dificuldades do cliente.

Além disso, a validade de conteúdo dos instrumentos usados para avaliar importantes domínios de funcionamento deve ser examinada cuidadosamente para justificar a operacionalização desses conceitos intuitivamente invocados, mas freqüentemente difíceis de definir (p. 312).

Para Foster e Mash (1999), a importância do foco do prejuízo funcional e adaptação

(...)vai além da comparação normativa e da avaliação subjetiva ao ressaltar a importância de examinar objetivos de tratamento que lidam com domínios importantes de funcionamento ou,

alternativamente, ao relacionar empiricamente objetivos e resultados adaptativos ou mal adaptados (p. 312).

Procedimento de Tratamento como Componente de Validade Social

Segundo Foster e Mash (1999), a validade social dos procedimentos de tratamento pode examinar seu aspecto de **aceitabilidade**, que é

(...) parte da questão mais ampla da viabilidade de tratamento e é particularmente importante quando pesquisadores transportam suas intervenções para locais da comunidade e estendem suas aplicações para populações diversas (p. 312).

Os autores relatam que uma recente pauta para desenvolvedores de intervenções psicológicas da *American Psychological Association* (APA) inclui explicitamente questões relacionadas à validade social em seu eixo de utilidade clínica, relacionado a avaliações da capacidade e prontidão do profissional para usar, e do paciente para aceitar, o tratamento em questão, e da aplicabilidade deste tratamento.

As duas principais formas pelas quais os investigadores têm avaliado a aceitabilidade do tratamento são o uso de *estudos experimentais análogos* e de questionários de *satisfação do consumidor* completados pelos clientes (Foster e Mash, 1999).

Estudos experimentais análogos envolvem, em uma de suas formas, apresentar a estudantes, clientes potenciais, membros do público ou profissionais de saúde mental, esboço literário de descrições dos problemas do cliente e possíveis procedimentos de tratamento e então pedir aos avaliadores que classifiquem os planos usando uma escala de classificação.

A maior parte desta pesquisa tem focado componentes de intervenções comportamentais, assim como contratos comportamentais, procedimentos de reforçamento e time-out, embora um pequeno número de investigadores tenha incluído intervenções não comportamentais, assim como medicação e intervenção paradoxal.

Investigadores têm usado inúmeras escalas para avaliar aceitabilidade nestas investigações análogas (...) (pp. 312-313).

Embora estudos de aceitabilidade do tratamento tenham alguma consistência em seus achados, os autores consideram que eles têm inúmeras limitações em sua validade externa. Muitos não usam populações de clientes para avaliar a aceitabilidade do tratamento, então seus achados podem ser mais relevantes para as reações da comunidade em geral do que para as reações dos clientes potenciais. A generalidade do análogo para as situações do mundo real mantém uma questão empírica com estes estudos, pois, embora classificações sejam assumidas como refletindo uma prontidão do cliente para participar de determinado tratamento para um problema particular, relatos verbais da aceitabilidade do tratamento não garantem que o cliente irá realmente seguir ou aceitar o tratamento por necessitarem deste ou, inversamente, que eles farão objeção a estes serem oferecidos em sua comunidade. Tais estudos têm focado quase exclusivamente procedimentos comportamentais usados com crianças ou indivíduos com distúrbios de desenvolvimento, talvez por estas populações não serem consideradas capazes de informar consentimento verdadeiro e, então, têm de ser tratadas com particular cuidado pela necessidade de preservar seu bem estar. Outro

problema é que estas pesquisas têm dedicado muito pouca atenção a intervenções não comportamentais.

Ainda sobre limitações dos estudos experimentais análogos, Foster e Mash (1999) observam que avaliações análogas geralmente apresentam informações sobre o tratamento como este deveria ser apresentado ao cliente em seu princípio. Entretanto, em muitas situações o profissional propõe um plano de tratamento depois de avaliar as dificuldades do cliente e fornece uma racional que relaciona o tratamento à avaliação dos resultados. Na verdade, estudos análogos têm mostrado que o tipo de racional, o meio pelo qual o investigador indica que o tratamento será implementado e quanto o tratamento é considerado colaborativo versus unilateral (para o cliente ou o profissional), pode influenciar a classificação de aceitabilidade do tratamento para crianças.

Completando o relato de limites dos estudos análogos, os autores atentam ao fato de a maioria das intervenções em estudos análogos serem examinadas em isolado, e os achados sobre elas poderem não se generalizar para tratamentos que usam combinações de intervenções.

Um aspecto delicado da aceitabilidade do tratamento é o de que esta pode variar através do tempo, particularmente em como clientes experienciam a natureza e o resultado dos tratamentos. Disto, Foster e Mash (1999) destacam

(...) a importância de olhar a aceitabilidade através do tempo em populações clínicas e de relacionar pontos de vista do cliente a outras dimensões importantes do processo terapêutico e resultados (p. 313).

Outra questão importante nos estudos análogos mostrada pelos autores é que, assim como qualquer avaliação relevante de aceitabilidade social, o investigador deve estabelecer a(s) população(ões)-alvo em estudos de aceitabilidade do tratamento e amostrar esta população adequadamente. “Muitas pessoas diferentes controlam se um tratamento pode e será implementado com sucesso em situações da comunidade” (p. 313).

A última questão, colocada por Foster e Mash (1999) envolvendo tais estudos relaciona-se à determinação de qual aspecto da “aceitabilidade” deverá ser avaliado. Embora investigadores tenham tipicamente operacionalizado aceitabilidade requisitando que avaliadores façam julgamentos globais, aceitabilidade deve ser examinada também de outras formas. Lennox e Miltenberger (1990, citado por Foster e Mash, 1999) sugeriram que avaliações de aceitabilidade de tratamento avaliam a aceitabilidade de efeitos colaterais do tratamento, o potencial de estratégias de tratamento serem abusivas, a compatibilidade do tratamento com a lei, a cooperação dos funcionários, a competência dos funcionários ou profissionais com os procedimentos, a eficiência do tratamento e a avaliação custo-efetividade. Embora seja possível o debate sobre a responsabilidade destas variáveis avaliarem diretamente a aceitabilidade em si,

(...) cada uma é claramente relevante para a questão mais ampla da viabilidade ou praticabilidade do tratamento e indica a natureza multidimensional do constructo da aceitabilidade do tratamento (p. 313).

Outro método ao qual se referem os autores, freqüentemente chamado de medida de **satisfação do consumidor**, que avalia a aceitabilidade do

tratamento, tem contado com questionários auto-aplicados completados pelos clientes, habitualmente após o tratamento.

(...) estes questionários geralmente pedem que os clientes avaliem vários aspectos do tratamento em algumas escalas de classificação específicas ou globais (p.314).

Foster e Mash (1999) pontuam alguns aspectos a serem considerados na avaliação da satisfação do cliente com os procedimentos de tratamento. O primeiro se refere ao fato de que existem elementos conceitualmente distintos da aceitabilidade do tratamento com os quais os

(...) clientes podem estar satisfeitos ou insatisfeitos: (a) o grau de mudança nos objetivos instrumental ou terminal, (b) os procedimentos de tratamento e (c) o terapeuta (p. 314).

Outro aspecto é a variedade de níveis de análise em que a satisfação pode ser analisada. Em um nível molar, o cliente pode responder sobre sua satisfação em relação ao tratamento como um todo, enquanto que, em um nível molecular, ele pode responder sobre diferentes elementos do tratamento.

Como clínicos e investigadores colocam grande ênfase em tratamentos combinados, medidas de satisfação global podem obscurecer quais elementos do tratamento combinado contribui mais para a variação nas classificações de satisfação. Por outro lado, classificações de elementos específicos de tratamentos combinados ou multimodais podem não refletir seu impacto na satisfação ou aceitabilidade quando tratamentos são apresentados como um pacote. Além do que, classificações de elementos específicos do tratamento tomadas muito depois do

fato podem ser influenciadas pelas experiências do cliente ao longo de toda intervenção (p. 314).

A última questão para a qual os autores requisitam atenção é a necessidade de medidas de satisfação do consumidor serem submetidas ao mesmo escrutínio psicométrico cabível às medidas de resultados.

A interpretação de resultados de classificações de aceitabilidade de tratamento de clientes pós-tratamento também requer atenção para algumas questões metodológicas importantes, segundo Foster e Mash (1999). Frequentemente, apenas aqueles que completam o tratamento participam da avaliação. Sendo assim, outros que não chegaram a participar da avaliação podem ter expressado sua opinião não aceitando inicialmente o tratamento ou retirando-se dele. A necessidade de reduzir possíveis características da demanda usando um avaliador que não seja o terapeuta para conduzir a avaliação e para assegurar o anonimato do cliente é outra questão metodológica importante a ser considerada.

Os autores abordam algumas dimensões adicionais da aceitabilidade do tratamento. Embora classificações do consumidor de vários aspectos da aceitabilidade do tratamento forneçam alguma indicação da viabilidade do tratamento, este não é o único dado relevante para a aceitabilidade do tratamento. As orientações da APA embutem a aceitabilidade no domínio mais abrangente da praticabilidade do tratamento, que é o objetivo maior da avaliação de validade social. Elas estipulam que a praticabilidade inclui fatores como aceitabilidade da intervenção, obediência do paciente aos requerimentos da intervenção, facilidade de disseminação da intervenção, facilidade de administração da intervenção e o custo da intervenção. Outros indicadores

diretos de praticabilidade são relevantes nesta questão: número de clientes elegíveis que não se inscrevem no tratamento quando solicitados, os que o abandonam e não mostram classificação e conclusão de dados dos requerimentos do tratamento. Schwartz e Baer (1991, citado por Foster e Mash, 1999) descreveram alguns indicadores de aceitabilidade pobre tais como demandas que o programa não cobre; baixo entusiasmo durante as sessões do programa e queixas a amigos, funcionários e mídia. Gresham e Lopez (1996, citado por Foster e Mash, 1999) destacaram indicadores do terapeuta de aceitabilidade do tratamento como o uso do tratamento e a fidelidade com que este é implementado. Embora outros fatores além da aceitabilidade possam contribuir para todas estas respostas,

(...) estes comportamentos são indicadores relevantes da viabilidade do tratamento e são partes particularmente importantes de informação sobre possíveis dificuldades quando da exportação de uma intervenção da pesquisa para situações da comunidade (p. 314).

Uma característica importante da aceitabilidade do tratamento, ressaltada por Foster e Mash (1999) é ser uma variável dinâmica e não uma variável estável. Isto significa que os pontos de vista dos indivíduos sobre produtos podem ser aperfeiçoados ou alterados com informação. Pesquisas sobre efeitos da educação na aceitabilidade do tratamento suportam tal afirmação.

Resultado de Tratamento como Componente de Validade Social

Resultados de tratamento podem ter avaliada, pela validade social, sua **importância social** (Foster e Mash, 1999).

Do ponto de vista de uma validade social, como mostram os autores, a demonstração de mudança significativa requer que o investigador meça importantes aspectos da vida do cliente e que a quantidade de mudança produzida pelo tratamento seja significativa, ou seja,

Os problemas devem ter declinado ou o funcionamento adaptativo aumentado o suficiente para fazer diferença em algum aspecto importante da vida do cliente (p. 314).

Comparações normativas fornecem uma base para examinar o nível de mudança depois do tratamento, assim como ocorre com outros componentes de validade social.

Outro método para examinar significância da mudança envolve a **avaliação subjetiva** pelo cliente, por outras pessoas importantes para o cliente em seu ambiente ou por *experts* ou juízes representativos e, neste método, a melhora dos julgamentos globais presumivelmente fornece evidência de que mudanças importantes ocorreram. Segundo Foster e Mash (1999),

Esta abordagem é mais apropriada para intervenções que visam habilidades comportamentais altamente específicas (...) quando avaliações globais do desempenho são indicadores importantes de qualidade do desempenho e quando níveis objetivos de desempenho aceitável da habilidade são difíceis de serem estabelecidos. Determinar níveis de aceitabilidade das habilidades é particularmente desafiador quando muitos tipos

diferentes ou combinações de comportamentos podem ser aceitáveis ou quando uma miríade de fatores interage para influenciar uma avaliação de julgamentos (p. 315).

Questões metodológicas críticas, a serem consideradas nessa abordagem, no relato dos autores, incluem características dos registros para observação. Por exemplo, é necessário que os segmentos de registros sejam longos o suficiente para que os juízes façam uma avaliação informada e representativos do desempenho típico do cliente. Também há a limitação da abordagem na documentação da importância dos resultados quando os objetivos da intervenção não são diretamente observáveis, quando os objetivos terminais são tão amplamente definidos que abrangem muitos aspectos do desempenho do cliente em uma variedade de situações, ou quando o cliente recusa o procedimento. Nestes casos, questionários podem ser usados para avaliar o grau em que, segundo o julgamento dos clientes (ou de outras pessoas) o tratamento atinge seus objetivos terminais.

Independentemente do método usado outros aspectos estão envolvidos na avaliação subjetiva para examinar resultados, incluindo questões sobre quem deve fornecer avaliação, o que os avaliadores devem ser requisitados a avaliar e como eles devem ser requisitados a conduzir sua avaliação.

Como na avaliação de objetivos, o(s) indivíduo(s) escolhido(s) para avaliar resultados deve(m) ser selecionado(s) com base na relevância de seu julgamento para o tipo de resultado a ser avaliado, resultados de objetivos instrumentais ou terminais (p. 315).

Segundo os autores, a validade social de objetivos instrumentais é menos bem delineada, pois “(...) objetivos instrumentais devem ser selecionados com base em relações presumidas ou empiricamente estabelecidas com os resultados terminais” (p. 315), o que torna a satisfação do cliente com estes resultados teoricamente irrelevante. Isto não significa que estes não possam ser avaliados. “Na realidade, esta avaliação é crucial para testar as ligações entre resultados terminais e os processos pelos quais se presumiu que o tratamento funcionaria” (p. 315). A demonstração de que objetivos terminais melhoram enquanto os instrumentais não o fazem ou vice-versa deveria levar o investigador a repensar as hipóteses que sublinharam o planeamento da intervenção.

Melhoras em ambos os conjuntos de objetivos, junto com resultados que mostram que melhoras nos objetivos instrumentais estão correlacionadas com melhoras nos objetivos terminais, suportam as supostas ligações entre os dois tipos de resultados (p. 315).

Avaliadores também podem ser requisitados a julgar várias diferentes dimensões de mudança. É comum, por exemplo, que se pergunte aos clientes, ou outras pessoas significativas para ele, sobre classificações do desempenho do cliente antes e depois do tratamento, em geral usando questionários que avaliam percepções de sintomas ou comportamentos relacionados aos objetivos terminais. Isto avaliaria se o comportamento do cliente mudou. Porém, na opinião dos autores, é mais relevante para a validade social informações que revelem se o cliente ou outros de seu ambiente estão satisfeitos ou felizes com os resultados.

Assim como avaliações de aceitabilidade do tratamento, avaliações de resultados devem ser específicas para o foco de avaliação, ou seja, devem avaliar a satisfação com os resultados em si e não a satisfação com os procedimentos. No caso de avaliação por outros indivíduos do ambiente do cliente, deve-se considerar que as percepções de outros podem ficar aquém do comportamento genuíno do cliente. Os autores usam como exemplo o estudo de Oden e Asher (1977, citado por Foster e Mash, 1999) no qual, apesar de ter ocorrido uma melhora nos sujeitos após o tratamento, a avaliação mostrou que seus pares não gostaram do resultado. No entanto, avaliações após 1 e 6 meses indicaram que a opinião dos pares a respeito dos sujeitos tratados havia melhorado, o que, presumivelmente, significa ser necessária uma experiência com as mudanças para que estas resultem em uma alteração da percepção global dos pares.

Assim como na avaliação de outros aspectos do tratamento, a avaliação da satisfação com resultados deve incluir tantos informantes quanto possível. Conceitualmente, cada informante tem um ponto de vista único; então, relatos de diferentes informantes não são formas paralelas da mesma medida. Discordância interinformante não é um problema desta perspectiva, já que diferenças de percepção não são erros de medida mas devem ser examinadas em sua própria verdade (Foster e Mash, 1999).

O emprego de múltiplos informantes também têm vantagens metodológicas, particularmente por permitir que sejam ressaltadas associações indevidas resultantes de flutuações decorrentes do método ou de idiossincrasias dos informantes. Isso pode ser feito quando informantes múltiplos avaliam muitos diferentes aspectos dos resultados do tratamento.

Escalas da satisfação do consumidor com o resultado do tratamento devem ser administradas de maneira que controlem diferenças iniciais da demanda, assim como ocorre com escalas de aceitabilidade de tratamento.

Assim como ocorre em relação à avaliação de outros componentes de validade social, também para a determinação da satisfação do consumidor, “medidas psicometricamente seguras são requeridas para avaliação adequada” (p. 316). Ainda, lembram os autores, a satisfação do consumidor é apenas uma perspectiva pela qual os resultados podem ser avaliados. “Pode-se argumentar que a satisfação do cliente com os resultados pode ser necessária mas não suficiente para mostrar o sucesso do tratamento” (p. 316). Medidas que não são subjetivas em natureza, mas, não obstante, falam sobre resultados importantes, como recidivas ou re-hospitalização, podem abordar esta questão.

A terapia pode produzir melhoras significativas mesmo que ela não atinja completamente “resultados importantes”, segundo Foster e Mash (1999). Um exemplo disto é o caso de uma terapia de casal que termina em divórcio, porém amigável. Um cliente pode não estar completamente satisfeito com o resultado do tratamento mas, mesmo assim, estar melhor em sentidos importantes.

A despeito destes embargos, dados de satisfação do cliente com o tratamento, juntamente com informações sobre variáveis relacionadas com adaptação e prejuízo funcional, (...) são partes importantes de dados que deveriam suplementar medidas de resultados instrumentais e terminais (p. 316).

DISCUSSÃO

O Conceito de Validade Social

As questões envolvidas no conceito de validade social estão presentes no âmbito da AC associadas às implicações ético-sociais da AC, desde o momento em que a ciência do comportamento começa a se consolidar uma tecnologia de intervenção.²¹

Quando do estabelecimento dos critérios para caracterização das dimensões que caracterizavam a ACA, por Baer, Wolf e Risley (1968), foram abordadas questões relacionadas à importância social em diferentes dimensões da pesquisa aplicada. Para os autores, objetivos deveriam ter importância social, já que os comportamentos manipulados na ACA não seriam os mais convenientes para o pesquisador, como na pesquisa básica, mas sim comportamentos **socialmente importantes**. O critério de importância social também atingia os resultados, visto que os autores sustentavam que uma aplicação deveria ter o efeito de alterar os comportamentos visados o suficiente para que a mudança fosse considerada socialmente importante.

Avaliações da prática do analista do comportamento, nos anos seguintes à publicação do artigo de Baer *et al.* (1968), também abordaram questões posteriormente embutidas no conceito de validade social. Winnett e Winkler (1972) alertaram que os valores por trás dos objetivos adotados não vinham sendo questionados pelos analistas do comportamento. Para Holland

²¹ Apesar de o surgimento de discussões envolvidas na questão da validade social serem aqui destacadas no contexto da ACA, deve-se lembrar que tais questões estavam presentes anteriormente na AC, fato evidente no trabalho de Andery (1990), que demonstra a presença da proposta de uma ciência para o homem e para a sociedade no desenvolvimento do sistema skinneriano.

(1983/1978), os objetivos adotados por analistas do comportamento, e a aceitação ou rejeição de procedimentos da parte deles, estavam ficando sob o controle de valores das classes dominantes.

Ao discutirem alternativas para mudanças nos rumos da ACA, os três trabalhos, sem exceção, atribuem à sociedade a competência e legitimidade de avaliar a importância das intervenções. Baer *et al.* (1968) qualificam uma pesquisa como aplicada não pelos procedimentos que esta emprega, mas pelo interesse que o problema desperta na sociedade. Winnett e Winkler (1972) supõem que possíveis alternativas para o problema da educação, principal foco dos autores, podem ser desenvolvidas por diferentes comunidades ou grupos dentro da comunidade. Holland (1983/1978) julga que o papel do analista do comportamento é ajudar os membros da sociedade a identificarem os problemas e planejarem soluções.

Como é ilustrado por esta afirmação de Holland (1983/1978), os autores dos trabalhos mencionados não deixam de atribuir responsabilidades a respeito destas questões ao analista do comportamento. Para Baer *et al.* (1968), é dever de uma pesquisa em ACA e, portanto, dos pesquisadores envolvidos nela, evidenciar a importância da mudança comportamental e a efetividade dos procedimentos em promoverem mudanças de suficiente valor prático. Winnett e Winkler (1972), assim como Holland (1983/1978), imputam ao analista do comportamento o dever de mudar o ambiente social que mantém o comportamento.

Neste contexto, a proposição do conceito de validade social por Wolf (1978/1976) é a atribuição de um rótulo a questões que vinham sendo discutidas na ACA desde a delimitação de seus domínios. Pode-se conceber, a

partir de Wolf (1978/1976), a validade social como um constructo que incorpora ferramentas desenvolvidas pelos analistas do comportamento com o intuito de evidenciar valores sociais nos quais se baseiam suas aplicações e, assim, aprimorar tais aplicações de forma a viabilizar a mudança social.

Apesar das questões envolvidas no conceito de validade social já serem discutidas na AC e ACA anteriormente à sua proposição, tal proposta trás como novidade a viabilidade de se trabalharem tais questões na prática.

Procedimentos de Validade Social: Objetivos e Questões Metodológicas

Em sua proposta original por Wolf (1978/1976), o conceito de validade social está estritamente relacionado à medida tomada dos julgamentos subjetivos coletados dos membros da comunidade, já que a importância social foi considerada pelo autor um julgamento subjetivo de valor que apenas a sociedade estaria qualificada a fazer.

No relato do próprio Wolf (1978/1976), uma vez constatada a inevitabilidade da validação social na prática do analista do comportamento, a questão que se colocava dizia respeito ao desenvolvimento de sistemas de medida que perguntassem à sociedade se as intervenções faziam algo socialmente importante. Os desenvolvimentos metodológicos por ele encaminhados já propunham que a validade social fosse avaliada em três níveis: objetivos, procedimentos e resultados.

Segundo Wolf (1978/1976), objetivos e resultados deveriam ter sua importância social avaliada, enquanto os procedimentos necessitariam ser avaliados quanto à sua adequação, ou seja, sua aceitabilidade. Assim, são apresentadas pelo autor três diferentes dimensões a serem socialmente

validadas: importância dos objetivos, aceitabilidade dos procedimentos e importância dos resultados. Apesar de não ser apresentada pelo autor uma distinção formal entre tais dimensões, fica claro que se tratam de coisas diferentes, porém, relacionadas. A relação das diferentes dimensões, de acordo com o texto, parece residir nos valores presentes no ambiente social que elas evidenciam.

Wolf (1978/1976) descreve procedimentos que visam validar socialmente as três dimensões mencionadas. Tais procedimentos têm seus pormenores apresentados no tópico destinado à proposta inicial do conceito de validade social (ver tópico “A Proposta do Conceito de Validade Social”).

Uma implicação do uso de procedimentos de validade social suposta por Wolf (1978/1976) é a de que ele aproxima a sociedade da ciência do comportamento, abrandando a imagem que ela faz desta ciência e consolida a busca da ACA por relevância social, questão a ser retomada adiante.

Embora Wolf (1978/1976) não faça referência explícita a limitações dos procedimentos que apresenta, referentes a cada uma das dimensões, ele acaba relatando dificuldades encontradas na sua aplicação, que foram resolvidas com novas avaliações de validade social. Exemplo disso é o caso de Willner *et al.* (1977, citado por Wolf, 1978/1976) que notaram que os comportamentos preferidos pelos jovens requisitados a julgar interações de pares, a fim de classificarem habilidades relacionadas ao “relacionar-se com jovens”, não correspondiam às expectativas dos pesquisadores e aplicaram um novo procedimento. Tal prática é coerente com as expectativas de Van Houten (1979) de que o processo de validação social seja iterativo. Ela também revela que a opinião de *experts* tem validade questionável e corrobora o apelo de Wolf

(1978/1976) à necessidade de acessar julgamentos da sociedade para a validação social de intervenções.

O apelo de Wolf (1978/1976) à validade social pode ser interpretado para além de uma sugestão de uso dos procedimentos de validação propostos pelo autor. Ao argumentar que os analistas do comportamento necessitam desenvolver sistemas e medidas melhores para perguntar à sociedade se eles estão fazendo algo socialmente importante, o autor parece invocá-los a construir tais sistemas.

A repercussão do apelo de Wolf (1978/1976) torna-se clara com a imediata resposta de Kazdin (1977), também imediatamente seguida da resposta de Van Houten (1979).

Kazdin (1977) mostra que, como no caso da avaliação subjetiva, alguns procedimentos de comparação normativa vinham sendo usados para obtenção de validade social antes mesmo que o conceito fosse cunhado. O autor ainda discute limitações dos dois métodos, além de outras questões metodológicas envolvidas neste. Apesar da apresentação de limites dos procedimentos de validação social, o autor é enfático na defesa de seu uso e da importância dele.

A extensão metodológica proposta por Van Houten (1979) é embasada em questões pertinentes à validade social, mas que os procedimentos anteriormente propostos não eram capazes de tratar. Sua iniciativa de resolver limitações metodológicas dos procedimentos de validação social e a repercussão de seu trabalho demonstram a possibilidade de o constructo “validade social” abranger diferentes ferramentas de avaliação da importância das intervenções.

Um aspecto interessante encontrado no confronto entre os trabalhos de Wolf (1978/1976) e Kazdin (1977) é que, em Kazdin (1977), procedimentos de validação social usados pelo grupo de Wolf e considerados por este como avaliação subjetiva são tomados como “mistos”, sendo revelados aspectos normativos, enquanto a própria descrição de Kazdin (1977) do método de comparação normativa demonstra aspectos de avaliação subjetiva ao apresentar o argumento de que o desempenho no nível normativo só deve ser considerado socialmente válido caso este represente uma melhora substancial. No método proposto por Van Houten (1979), também são ressaltados aspectos subjetivos e normativos nas diferentes etapas do procedimento. Assim, é possível que tanto aspectos normativos quanto subjetivos sejam elementos inevitáveis em grande parte das propostas de procedimentos de validação social, visto que estes, em geral, trabalham com diversos indivíduos e procuram evidenciar a questão do valor.

A diversidade de procedimentos e seus respectivos aspectos metodológicos presentes no artigo de Foster e Mash (1999) revelam a intenção de manter a funcionalidade e a análise dos elementos das contingências presentes. Com isso, parece ser necessário que o pesquisador componha elementos para seus procedimentos de validação social em cada situação de intervenção, não havendo procedimentos-padrão a serem seguidos.

A Compatibilidade do conceito de Validade Social com o Behaviorismo Radical e com o paradigma metodológico da ACA: implicações para a Prática do Analista do Comportamento

Questões filosóficas da validade social e, portanto, relacionadas ao Behaviorismo Radical, são notadas já na definição dos domínios da ACA. Ao distinguir a pesquisa básica da pesquisa aplicada usando como principal critério a necessidade de esta de examinar comportamentos socialmente importantes, Baer *et al.* (1968) revelam uma questão filosófica na delimitação da ACA. Os autores ainda tornam explícita a relevância da discussão filosófica do que deve ser entendido por “bem da sociedade”.

A discussão apresentada em Wolf (1978/1976) é essencialmente epistemológica, já que seu foco é colocado na questão do valor da aplicação da Ciência do Comportamento para o ser humano. Pelo artigo, é possível entender tal questão como inevitável, principalmente no estado em que a Ciência do Comportamento se encontra: intervindo na sociedade em busca da solução de seus problemas importantes. Para o autor, as contingências do ambiente da pesquisa aplicada levaram a AC ao questionamento sobre o acesso de “*dimensões de reforçadores complexos de modo socialmente aceitável e prático*” (p. 213). Tal questão, conforme aqui reiterado, de acordo com Wolf (1978/1976), é inevitável. As análises decorrentes da presente pesquisa levam a comparar a negligência desta questão, ao tratamento dado pelo Behaviorismo Metodológico à questão dos eventos privados: ignorá-los por serem considerados inacessíveis com a metodologia “então” disponível.

Wolf (1978/1976) também mostra que mesmo a tecnologia mais objetiva está sujeita a erro e nem por isso ela é abandonada. Assim, tecnologias que

acessam dados subjetivos não são, necessariamente, menos precisas que aquelas que avaliam dados objetivos. Há necessidade de desenvolver tecnologias cada vez mais confiáveis para medir dados subjetivos, da mesma forma como foram desenvolvidas tecnologias precisas de avaliação de dados objetivos. A menos que se demonstre que a subjetividade seja intrínseca aos fenômenos abarcados pelo conceito de validade social, a questão pode ser colocada nos mesmos termos em que fenômenos a um tempo dados como qualitativos tornaram-se quantitativos por se ter sido capaz de identificar parâmetros que os definiam.

Da proposição da validade social, de procedimentos que a avaliem e do uso destes procedimentos, decorrem várias implicações para a prática do analista do comportamento.

Wolf (1978/1976) ressalta a relevância do método que propõe para avaliação de importância dos objetivos descritos por rótulos subjetivos, já que tais rótulos são atribuídos a muitos problemas da sociedade. Com isso, procedimentos de validação social têm a implicação de permitir que os analistas do comportamento sejam mais hábeis em explorar estes problemas.

A aceitabilidade dos procedimentos de tratamento, como mencionado em Wolf (1978/1976), pode estar relacionada à efetividade do programa, pois dela depende a adesão a ele. O autor relata que, após a adoção de procedimentos de avaliação de aceitabilidade do tratamento por seu grupo de trabalho, seus membros deixaram de ser expulsos das comunidades às quais levavam seu programa. Considerando que o programa de Wolf e seus colegas visava uma intervenção não institucional, na qual os problemas das comunidades eram tratados nas situações em que eram adquiridos e mantidos

- uma postura mais coerente com a ciência do comportamento tomando as críticas proferidas por Holland (1983/1978) às intervenções institucionais -, é possível que avaliações de validade social tenham o efeito de distinguir os analistas do comportamento de membros da agência governamental em programas que atuam neste *setting*.

O uso de procedimentos de validação social pode ser uma ferramenta para que a ACA contemple a mudança social. Quando Wolf (1978/1976) propõe que a sociedade avalie a importância social, ele tem a oportunidade de mudar o foco de comprometimento do analista do comportamento das classes dominantes e das instituições que o emprega para a sociedade, respondendo às críticas de Winnett e Winkler (1972) e Holland (1983/1978). Caso tal mudança de foco não ocorra, minimamente serão esclarecidos os valores por trás da intervenção, o que, segundo Holland (1983/1978), também favorece a mudança social.

O conceito de validade social surgiu a partir das necessidades práticas do analista do comportamento de avaliar se suas aplicações eram importantes para a sociedade. Neste sentido, a avaliação de validade social apenas caberia às intervenções em ACA. Porém, dada a relação entre pesquisa básica e aplicada no sentido em que esta é teste e fonte de problemas para aquela (Andery, 1990), talvez coubesse a avaliação de validade social dos objetivos da pesquisa básica. Tal avaliação, provavelmente, seria obtida através da consulta a analistas do comportamento envolvidos com aplicações, visto que estes devem ser os mais capacitados a avaliar que processos constituem objetivos instrumentais importantes para que os objetivos terminais sejam alcançados. Assim, ao desenvolverem pesquisas básicas, os analistas do comportamento

poderiam legitimar o despendimento de recursos para tais pesquisas. Evidentemente, comitês de ética também podem se privilegiar do conhecimento construído em validade social a respeito de aceitabilidade de procedimentos e resultados.

Cabe uma repetição que é, na verdade, o conhecimento maior tirado desta pesquisa. Levando em conta o atual estado da produção teórica, conceitual e metodológica sobre validade social, considera-se inevitável a adesão à validade social pelo analista do comportamento como pesquisador e prático no que diz respeito à reflexão das questões embutidas no conceito e aplicação de procedimentos que evidenciem valores que embasam sua prática. Entraves metodológicos devem ser encarados como propulsionadores do desenvolvimento de melhores sistemas de avaliação da importância social das intervenções ao invés de impeditivos do uso de procedimentos de validade social.

REFERÊNCIAS BIBLIOGRÁFICAS

Anastasi, A. (1988). *Psychological testing*. New York: Macmillan Publishing.

Andery, M. A. P. A. (1990). *Uma tentativa de (re)construção do mundo: A ciência do comportamento como ferramenta de intervenção*. Tese de Doutorado. Programa de Estudos Pós-graduados em Psicologia Social. Pontifícia Universidade Católica de São Paulo.

Baer, D. M., Wolf M. M., e Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, 1, 91 - 97.

Cesar, G. (2002). *Análise do Comportamento no Brasil: uma revisão histórica de 1961 a 2001, a partir de publicações*. Dissertação de Mestrado. Psicologia Experimental: Análise do Comportamento. Pontifícia Universidade Católica de São Paulo.

Foster, S. L. E Mash, E. J. (1999). Assessing social validity in clinical treatment research: Issues and procedures. *Journal of Consulting and Clinical Psychology*, 67(3), 308-319.

Geller, E. S. (1991). Where's the validity in social validity? *Journal of Applied Behavior Analysis*, 24, 179 – 184.

- Hawkins, R. P. (1991). Is social validity what we are interested in? Argument for a functional approach. *Journal of Applied Behavior Analysis*, 24(2), 205-213.
- Hayes, S. C., Rincover, A. e Solnick, J. V. (1980). The technical drift of Applied Behavior Analysis. *Journal of Applied Behavior Analysis*, 13, 275 – 285.
- Holland, J. G. (1983/1978). Behaviorism: Part of the problem or part of the solution? [Comportamentalismo - parte do problema ou parte da solução?, Tradução] (D. G. Souza & S. P. Botomé). *Psicologia*, 9(1), 59-75.
- Kazdin, A. E. (1977). Assessing the clinical or applied significance of behavior change through social validation. *Behavior Modification*, 1, 427 – 452.
- Kazdin, A. E. (1978). *History of behavior modification: Experimental foundations of contemporary research*. Baltimore: University Park Press.
- Kennedy, C. H. (1992). Trends in the measurement of social validity. *The Behavior Analyst*, 15, 147 – 156.
- Kennedy, C. H. (2002). The maintenance of behavior change as an indicator of social validity. *Behavior Modification*, 26(5), 594-604

- Rodrigues, M. E. (2005). *A contribuição do Behaviorismo Radical para a formação de professores: uma análise a partir das dissertações e teses no período de 1970 a 2002*. Tese de Doutorado. Programa de Estudos Pós-graduados em Psicologia da Educação. Pontifícia Universidade Católica de São Paulo.
- Van Houten, R. (1979). Social validation: the evolution of standards os competency for target behaviors. *Journal of Applied Behavior Analysis*, 12, 581-592.
- Winett, R. A. e Winkley, R. C. (1972). Current behavior modification in the classroom: be still, be quiet, be docile. *Journal of Applied Behavior Analysis*, 5, 499 – 504.
- Winett, R. A., Moore, J. F. e Anderson, E. S. (1991). Extending the concept of social validity: behavior analysis for disease prevention and health promotion. *Journal of Applied Behavior Analysis*, 24, 215 – 230.
- Wolf, M. (1978/1976). Social validity: The case for subjective measurement or How apllied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis*, 11, 203 – 214.