

# A Validação do Teste de Juízo Moral (MJT) para Diferentes Culturas: O Caso Brasileiro

## *Moral Judgment Test (MJT) Adaptation for Different Cultures: Brazilian Case*

Patricia Unger Raphael Bataglia\*  
Universidade Estadual Paulista UNESP, Marília

### **Resumo**

O presente artigo trata da adaptação do Teste de Juízo Moral (MJT) para a língua portuguesa. São apresentados os princípios do teste, seu formato e o processo de validação de conteúdo, constructo e de critério. Da mesma forma, a confiabilidade é discutida utilizando os resultados das pesquisas brasileiras que inclusive geraram a necessidade de elaboração de outro dilema para compor o MJT-xt (estendido).

**Palavras-chave:** Competência moral; Processo de validação de escalas; Desenvolvimento moral.

### **Abstract**

The present article deals with the adaptation of the Moral Judgment Test (MJT) for the Portuguese language. The principles of the test, its format and the process of content, construct and criterion validation are presented. In the same way, reliability is discussed using Brazilian research results that produced the necessity of elaborating another dilemma to compose the MJT-xt (extended).

**Keywords:** Moral competence; Scale validation process; Moral development.

Desde que Lawrence Kohlberg publicou “*The measurement of moral judgment*” em 1987 (Colby et al., 1987) propondo uma forma de aplicação e avaliação de estágio de desenvolvimento moral, com sua entrevista de juízo moral (*Moral Judgment Interview – MJJ*) vários outros instrumentos foram elaborados na tentativa de abreviar o tempo de aplicação para avaliação do nível de juízo moral, ou seja, avaliação da capacidade de ajuizar a respeito de conteúdos morais, ou para buscar captar outras dimensões do desenvolvimento da moralidade. Alguns desses instrumentos foram traduzidos e validados para a língua portuguesa.

Em busca de uma escala mais objetiva que permitisse a aplicação em maior escala sem perder o vínculo com os estágios de desenvolvimento moral elaborados por Kohlberg, James Rest e colaboradores (Rest & Narvaez, 1998) elaboraram o DIT (*Defining Issues Test*). Do mesmo modo que o MJJ, o DIT apresenta dilemas que devem ser avaliados pelo sujeito. Porém, ao invés de responder a questões, o sujeito é convidado a avaliar doze respostas em uma escala de cinco graus e em seguida deve escolher as quatro respostas mais importantes. O

DIT fornece um escore chamado P, que indica a porcentagem de respostas pós-convencionais. Posteriormente, o DIT, agora chamado DIT2, foi reformulado pelos autores (Rest & Narvaez, 1998) e revalidado para a realidade brasileira (Biaggio, Shimizu, & Martinez, 1998).

O SROM (*Socialmoral Reflection Objective Measure*) é também constituído por questões a respeito de temas como manutenção de uma promessa, dizer a verdade, ajudar a alguém, salvar a vida de alguém, valor da vida humana, roubo e obediência à lei. As respostas são classificadas de acordo com uma chave de correção que se baseia nos estágios kohlberguianos. O questionário apresenta a vantagem de poder ser administrado coletivamente, mas sua avaliação é tão trabalhosa quanto a do MJJ, uma vez que o pesquisador precisa confrontar a resposta dada a uma lista de possibilidades, podendo inclusive concluir que a resposta do sujeito não se parece com nenhuma das presentes na lista. O teste foi adaptado para a realidade brasileira por Biaggio e Brandão (1990). Como citado por Gibbs, Basinger e Fuller (1992), os argumentos dos autores para não utilizarem dilemas e sim questões abertas sobre temas é o de que as situações oferecidas são frequentemente criticadas como artificiais, inapropriadas ou irrelevantes para crianças (Damon, 1977; Stein, Trabasso, & Garfin, 1979), mulheres (Gilligan, 1982), certas culturas (Boyes & Walker, 1988), e situações práticas (Hann, Aerts, & Cooper, 1985). O escore fornecido pelo SROM indica o nível de juízo moral, semelhante ao MJJ.

\* Endereço para correspondência: Rua Turiassu, 152, Apto. 21, Perdizes, São Paulo, SP, Brasil, CEP 05005-000. Tel.: +(11) 3667 8466, +(11) 9202 5327. E-mail: [patriciabat@terra.com.br](mailto:patriciabat@terra.com.br)

Agradeço o apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelo financiamento das pesquisas que possibilitaram o trabalho de adaptação do MJT para a língua portuguesa.

O presente artigo tratará de outro instrumento, o MJT (*Moral Judgment Test*). Originalmente elaborado em alemão (*Moralisches Urteil Test* - MUT) na década de 70, por Georg Lind, professor da Universidade de Konstanz, o MJT tem estudos de adaptação para 27 línguas e tem se mostrado muito útil para mensurar os efeitos de programas de desenvolvimento moral.

Durante a adaptação no Brasil e no México notou-se algumas particularidades que merecem ser discutidas e que levaram ao desenvolvimento de outro dilema que passou a integrar o teste, agora chamado de MJT<sub>xt</sub> (*extended*).

Ainda que tenhamos mantido na língua portuguesa a denominação teste, o MJT é na realidade um experimento realizado com o objetivo de avaliar em que medida o grupo (e não o sujeito individualmente) é capaz de avaliar a qualidade de argumentos morais ou fica preso à sua própria opinião.

É importante salientar que a adaptação do MJT<sub>xt</sub> foi feita para a realidade brasileira. O instrumento exige uma adaptação cultural para cada situação.

Há alguns aspectos desse instrumento que são muito interessantes para a continuidade das pesquisas sobre moralidade. Nesse artigo procuramos organizar as idéias da seguinte forma: apresentação do Teste de Juízo Moral, processo de adaptação do MJT para o Brasil e pesquisas recentes usando o MJT<sub>xt</sub> no Brasil.

### **O Teste de Juízo Moral: Avaliação da Competência Moral**

O conceito de competência moral foi originalmente cunhado por Lawrence Kohlberg em 1964. Foi definido como “a capacidade de tomar decisões e emitir juízos morais (baseados em princípios internos) e agir de acordo com tais juízos”. Esse conceito é muito importante para os estudiosos do desenvolvimento moral em geral, e para os estudiosos da formação ética do profissional em particular, quando se dedicam ao estudo de como transformar teoria e técnica em uma práxis que seja socialmente comprometida.

Embora Kohlberg tenha, ele próprio, se dedicado durante toda sua obra à capacidade de ajuizar, está na sua definição de competência moral a possibilidade de pensarmos as implicações de sua teoria na esfera da conação, entendida como a tendência consciente para a ação, nesse caso em situações sociais que envolvam problemas morais.

Quando tratamos da construção da capacidade de ajuizar, referimo-nos à esfera cognitiva, condição necessária, mas não suficiente para a competência moral.

Além de reconhecer regras, aceitá-las e conseguir emitir juízos sobre tais regras, é necessário também considerar a disposição do sujeito para agir de acordo com tais regras. Disposição para a ação é o elemento que permite ligarmos o juízo à ação moral propriamente dita. Não é

possível avaliarmos a ação em si como moral, imoral ou amoral, pois não temos acesso às intenções do sujeito, mas podemos estudar o grau de coerência com que o sujeito diferencia e integra princípios morais e os aplica nas decisões do dia-a-dia.

A competência nesse caso é do juízo moral e exprime a habilidade que o sujeito possui para avaliar situações e juízos de acordo com princípios e não simplesmente com interesses ou disposições momentâneas. Lind (2006) resume bem a importância da construção dessa competência na sociedade atual:

Um dos princípios centrais das democracias modernas é resolver conflitos ou dilemas por meio de negociação e discussão preferencialmente ao uso do poder, força ou violência. Obviamente, um pré-requisito dos mais importantes para a negociação pacífica é a habilidade dos participantes para ouvir-se uns aos outros independentemente do fato de serem oponentes ou até inimigos. Se quisermos encontrar a base moral para a solução de um conflito, devemos apreciar argumentos não apenas advindos de pessoas que suportam nossa posição, mas também daquelas que são nossas oponentes. Tal competência é crucial para a participação em uma sociedade pluralística e democrática (Lind, 2006, Summary, para. 2).

Diferentemente dos instrumentos anteriormente mencionados, o MJT não foi elaborado com o propósito de avaliar o nível de juízo moral, mas sim, a competência moral. Georg Lind chama a atenção para um aspecto fundamental já anunciado por Piaget (1976): os componentes afetivos e cognitivos são distintos, mas inseparáveis na emissão do juízo moral. Sendo assim, não é possível atribuímos um escore para a construção da estrutura cognitiva sem considerarmos a atitude do sujeito frente ao dilema usado para a mensuração.

Dito de outra forma, Lind argumenta que é fácil para alguém que tenha habilidade para argumentação e conhecimento dos estágios ou de filosofia, responder a um dilema em níveis superiores de juízo moral, falsificando o seu próprio estágio de desenvolvimento. A retórica não garante a competência na ação. Sendo assim, Lind se preocupou antes do que avaliar o nível de juízo moral do sujeito, um tipo de experimento que permitisse ao sujeito mostrar sua competência em aplicar sua estrutura em situações adversas, isto é, situações com as quais ele não compartilhe a opinião ou ajuíze de modo contrário.

O MJT é um questionário que confronta o sujeito com dilemas morais e solicita que sejam avaliados argumentos a favor e contra a atitude do personagem principal do dilema e inclui originalmente duas estórias.

A primeira estória é sobre um operário que decide arrombar a empresa em que trabalha e roubar provas de atividades irregulares da gerência para denunciá-la junto a instâncias superiores. A segunda é o dilema clássico da eutanásia em uma situação em que o paciente termi-

nal pede para que o médico realize algum procedimento que abrevie seu sofrimento.

Em cada uma delas, o personagem principal é colocado em uma situação de dilema. Qualquer que seja sua decisão estará entrando em conflito com algumas regras de conduta social e talvez pessoal. O personagem toma uma decisão e o sujeito é solicitado em primeiro lugar a avaliar a decisão do protagonista. Em seguida, é convidado a avaliar seis argumentos a favor da decisão do protagonista e seis contrários. Tais argumentos representam os diferentes níveis de raciocínio moral de acordo com os seis estágios descritos por Kohlberg (1964).

Uma coisa particularmente interessante nas histórias é que o personagem toma uma decisão e age. Não cabe ao sujeito decidir o que fazer. Este procedimento simples é muito importante porque não permite a elaboração de outras alternativas. A ação já é estabelecida e cabe ao sujeito concordar ou não com ela.

O escore do MJT leva em conta o padrão de respostas do sujeito no teste como um todo e não em respostas isoladas de modo a medir a competência moral. Por exemplo, se alguém julga um argumento que se baseia na consciência moral como altamente aceitável para uma situação de eutanásia e ela própria é favorável à eutanásia, não sabemos ao certo se esse juízo reflete uma alta consideração pelo valor da consciência moral ou simplesmente seu compromisso a favor da eutanásia. Somente analisando-se o padrão de consistência de respostas do sujeito em situações diversas com relação à atitude, é que podemos dizer algo a respeito de sua competência.

O MJT propõe uma tarefa moral e não simplesmente mede a atitude do sujeito. Muitas tarefas morais podem ser imaginadas para avaliarmos a competência moral do sujeito, mas poucas são factíveis e/ou válidas. Medir a resistência a tentação do sujeito ao roubo, induzindo-o a roubar, por exemplo, parece pouco ético. Verificar se um sujeito ajuda a alguém em dificuldade, não garante o valor moral do ato (ou decisão) uma vez que as razões poderiam ir da busca de autopromoção até um princípio universal, passando por obediência a um código externo, busca de aprovação dos pares, etc.

A tarefa moral proposta pelo MJT é a de confrontar o sujeito com contra-argumentos. Enquanto a reação do sujeito em relação a argumentos favoráveis a sua própria decisão mostram a preferência do sujeito por um determinado nível, as reações a contra-argumentos evidenciam sua habilidade para usar um nível particular de modo consistente quando ajuizando a respeito do comportamento de alguém.

É interessante notar que as duas dimensões do comportamento moral: afetivo e cognitivo não se apresentam necessariamente conectadas, embora ocorram de modo integrado, isto é, apesar de muitos indivíduos preferirem argumentos de estágios morais superiores, apenas aqueles com estruturas mentais reversíveis podem ser também competentes moralmente, isto é, apresentar uma preferência pelos mesmos estágios quando avaliam

contra-argumentos, ou argumentos rivais a sua opinião. Lembrando que a reversibilidade cognitiva por si só também não garante a competência moral, o que a reduziria a uma operação intelectual.

O MJT serve segundo Lind, a dois propósitos: permite-nos testar as modernas teorias de desenvolvimento moral e educação e permite-nos avaliar métodos educacionais em relação ao seu poder de elevar a competência moral dos sujeitos submetidos a eles. Em busca de satisfazer a esses dois propósitos, o MJT deve ter validade teórica e ser útil do ponto de vista educacional.

Em psicologia assim como na maior parte das outras ciências, testes e instrumentos de medida são construídos para prover dados. Com tais dados busca-se testar a verdade empírica de tais teorias (ou de hipóteses derivadas de tais teorias), ou avaliar os efeitos de certos métodos de intervenção, ou ambos ao mesmo tempo. Se usarmos os instrumentos de medida para testar teorias, esses instrumentos devem ter validade teórica, isto é, deve realmente medir aquilo que pretende medir. De outro modo, os dados produzidos pelo instrumento são irrelevantes para a teoria ser avaliada e assim sendo é inútil (Cronbach & Meehl, 1955; Popper, 1979). Se usarmos um instrumento de medida para avaliar métodos de educação ou psicoterapia, o instrumento deve ser educacionalmente útil, isto é, deve medir exatamente os aspectos do comportamento humano que desejamos educar ou aprimorar. (Lind, 2006, *Criteria for evaluating the MJT: Validity and Utility*, para. 2).

### Validação do MJT para o Brasil

A adaptação do MJT para a língua portuguesa do Brasil incluiu várias etapas desde a sua tradução em 1998 até suas mais recentes modificações em 2006. Para comparação dos resultados obtidos no Brasil com o de outras culturas e mesmo dos resultados obtidos em diferentes estudos no Brasil, procedeu-se a um estudo de validação.

Quando abordamos o tema da validação é importante diferenciar duas situações: a construção de um instrumento que relacione indicadores ao conceito que se pretende mensurar e a adaptação de um instrumento já construído para as necessidades de uma cultura e línguas diferentes.

A construção do MJT seguiu parâmetros de validade que devem ser seguidos nas validações feitas em outras línguas.

O tema validade é complexo, controverso, porém particularmente importante em pesquisas que envolvam o comportamento humano. Quando perguntamos a respeito da validade de um instrumento de medida, estamos no fundo querendo saber se estamos mensurando aquilo que dizemos e pretendemos medir.

Tradicionalmente, as escalas de mensuração devem passar por basicamente dois crivos para serem conside-

radas precisas e coerentes. A precisão está relacionada à validade e a coerência à confiabilidade (Breakwell, Hammond, Fife-Shaw, & Smith, 2006; Carmines & Zeller, 1979). A seguir discutem-se essas duas dimensões.

### Validade

A validade se refere ao quanto a escala mede o que de fato desejamos medir. Esta escala é precisa ou válida se por um lado pode obter resultados generalizáveis (validade externa) e por outro se tiver sucesso em mensurar na amostra em questão algo definido pela teoria (validade interna).

A validade externa depende de fatores pertinentes à definição da amostra. Diz respeito à possibilidade de generalização para (e entre) épocas, contextos ou pessoas (Cooper & Schindler, 2003). Portanto, faz sentido pensarmos que a validade externa do MJT é ampliada a cada processo de validação desenvolvido (replicação do procedimento de mensuração original) para uma amostra independente ou mesmo para uma população diferente da que foi utilizada originalmente por Lind em 1980.

Quanto à validade interna tem sido classificada em três principais formas: validade de conteúdo, de critério e de constructo. Veremos a seguir o que caracteriza cada uma das formas e como o MJT tem respondido a elas.

A validade de conteúdo é “o quanto esse instrumento fornece cobertura adequada das questões investigativas que orientam o estudo” (Carmines & Zeller, 1979, p. 20). A determinação da validade de conteúdo depende de uma análise cuidadosa do examinador a respeito dos itens que compõem seu instrumento de mensuração (Breakwell et al., 2006). O pesquisador deve ser capaz de especificar o conjunto completo dos fatores relevantes para mensurar uma situação específica. Em ciências não exatas, nas quais mensuramos não dados ou eventos, mas indicadores de conceitos abstratos, a validade de conteúdo é especialmente complexa (Carmines & Zeller, 1979). Uma forma aceita de se garantir essa validade é a consulta a especialistas.

Lind (2005b) esclarece em relação à versão original do MJT (alemã) que como validade de conteúdo, pode-se contar com uma construção do teste estritamente baseada na teoria do desenvolvimento moral de Kohlberg e uma avaliação bastante minuciosa dos itens do teste. Seis peritos no modelo de estágios de desenvolvimento moral de Kohlberg analisaram a adequação teórica dos argumentos no MJT na versão alemã e inglesa. No processo de validação para a língua portuguesa, após a *backtranslation*, o MJT foi submetido a uma avaliação do autor e de um perito bilíngüe para assegurar a correspondência dos argumentos aos estágios de Kohlberg.

A validade de critério ou validade preditiva, “avalia se um constructo tem o desempenho esperado em relação a outras variáveis identificadas como critérios significativos” (Carmines & Zeller, 1979, p. 17). No caso da com-

petência do juízo moral, há uma sensível progressão de acordo com a quantidade e qualidade de educação formal. O MJT tem se mostrado altamente sensível a este índice, então poderíamos dizer que este seja um critério concorrente, ou seja, a variação da quantidade e qualidade de educação formal acompanha a variação crescente do índice chamado *score C* (de competência).

Vejam as figuras abaixo resultados de diferentes pesquisas realizadas no Brasil utilizando-se formas variadas do MJT (dois dilemas ou três).

A Figura 1 se refere aos resultados encontrados na primeira validação do teste para a língua portuguesa em 1998.

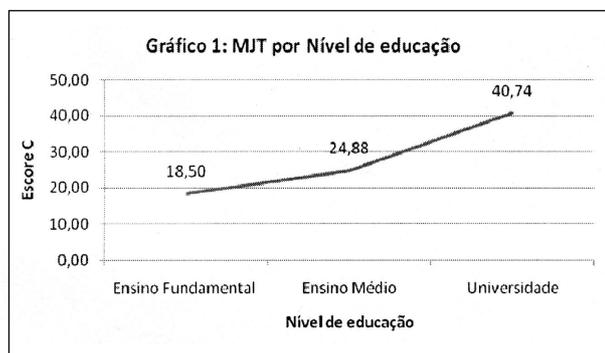


Figura 1. MJT por nível de educação (Bataglia, 2001)

Observe-se que o índice do *score C* aumenta progressivamente conforme a quantidade de educação. Outro exemplo é a validação de outro dilema para o MJT, realizada em 2003:

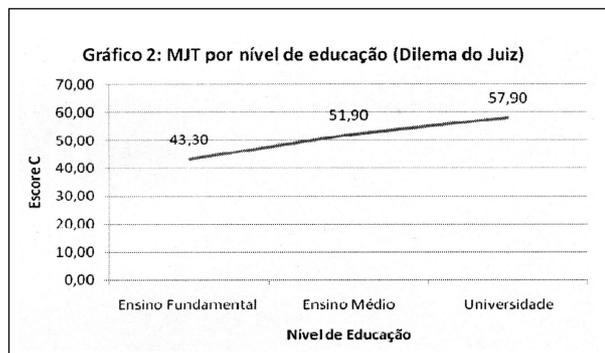


Figura 2. MJT por nível de educação - Dilema do Juiz (Bataglia, Schillinger-Agati, Lind, & Quevedo, 2003)

Os índices superiores nesse trabalho serão comentados mais adiante.

Em todos os trabalhos de validação realizados no Brasil (Bataglia, 2001; Bataglia, Agati, et al., 2002; Bataglia et al., 2003; Bataglia et al., 2006; Schillinger, 2006) esta progressão se manteve o que garante a validade de critério. Algumas variações interessantes ocorrem no que diz respeito à qualidade de educação. Quando comparadas,

por exemplo, universidades competitivas e não competitivas o índice do escore C é bem diferente, conforme o trabalho de Schillinger (2006) que compara primeiros e últimos anos de três cursos universitários em três países diferentes. As universidades competitivas são aquelas que por serem mais procuradas acabam selecionando alunos egressos de experiências educacionais mais enriquecedoras do ponto de vista formal. Embora o conteúdo escolar não influencie nos resultados do teste, acompanha a educação formal oportunidades de formação que são importantes para a construção da competência moral, tais como oportunidades de assunção de responsabilidades e reflexão dirigida.

A validade de constructo “avalia o que o conceito ou escala está, de fato, medindo” (Hair, Babin, Money, & Samouel, 2005, p. 202). Devem-se identificar hipóteses derivadas teoricamente dos constructos que estão sendo medidos e sua consistência com as medidas obtidas (Carmines & Zeller, 1979).

Derivou-se da teoria cognitivo-desenvolvimentalista quatro hipóteses básicas sobre a competência moral que devem estar coerentes com as mensurações obtidas pelo MJT: preferência por argumentos de estágios superiores, preferência por estágios vizinhos, maior do que por estágios distantes, paralelismo afetivo-cognitivo e resistência ao falseamento por instruções.

Com relação à primeira hipótese, Rest (1969) postula que um dilema deve suscitar no sujeito a preferência por argumentos de estágios superiores em termos de raciocínio moral, com o maior índice de preferência para estágio seis e o menor para o estágio um. Os seis estágios podem ser ordenados em 720 diferentes modos e a probabilidade de ordenação ao acaso é de 1/720 ou 0,001, ou seja, muito pouco provável (Lind, 2006).

A respeito disso, as amostras no Brasil responderam muito bem. A Figura 3 apresenta a preferência pelos estágios, por meio da aceitabilidade dos argumentos. Algumas pequenas inversões de preferência de estágio (especialmente entre estágios 1 e 2, bem como entre estágios 5 e 6) podem ocorrer, e isso não invalida a nova versão do teste. Pesquisas trans-culturais suportam essa hipótese bastante bem (Gross, 1997; Lind, 1986).

Com relação à segunda hipótese, de acordo com a Kohlberg (1958, segundo Lind, 2006), a preferência por estágios vizinhos deve ser maior do que a preferência por estágios distantes. A essa estrutura ele chama de quase-simplex. A seguir é apresentado a Figura 4 sobre a estrutura quase-simplex dos resultados da validação do MJT em 1998.

Note-se que os estágios se encontram distribuídos em ordem crescente, o que caracteriza a correlação em estágios próximos

Com relação à terceira hipótese, de acordo com Piaget (1976) há um paralelismo afetivo-cognitivo no desenvolvimento moral, isto é, “os mecanismos afetivos e cognitivos são inseparáveis, embora distintos: o primeiro depende da energia, e o último depende da estrutura” (p. 71). Por conseguinte, Kohlberg descreveu seu modelo de estágios do desenvolvimento moral de modo que fosse uma descrição de ambos os aspectos (Kohlberg, 1958, segundo Lind, 2006). Há uma discussão a respeito de se a proposta de Kohlberg corresponderia ou não às idéias de Piaget com relação ao desenvolvimento moral (Biaggio, 2002; Hala, 1997), contudo, abordar esse tema escaparia do escopo do presente artigo. Assumir-se-á aqui a posição de Lind que prevê como forma de validação de constructo a correlação do Escore C com a preferência por estágios. De acordo com a teoria, os índices mais

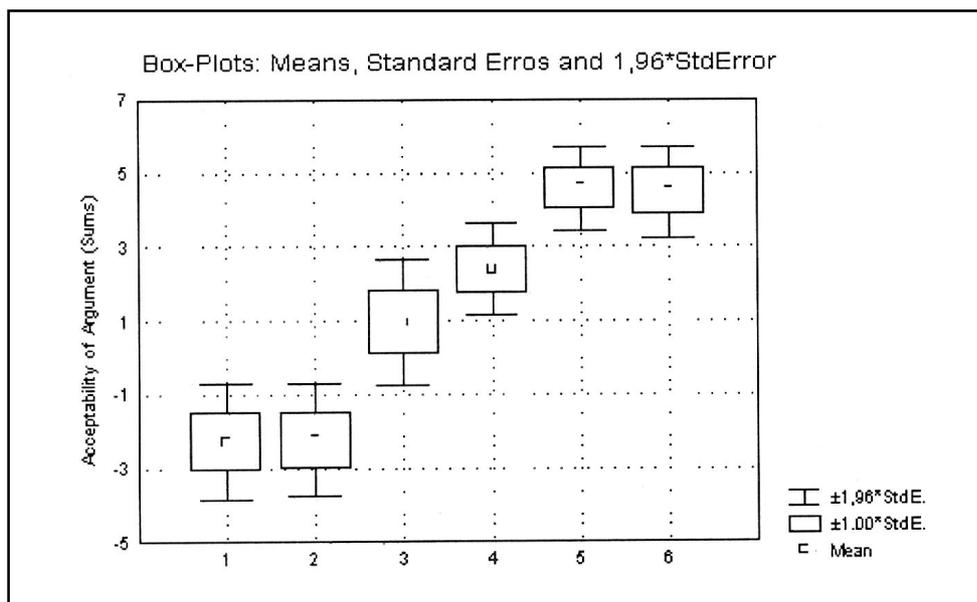


Figura 3. Aceitabilidade dos argumentos – Médias e desvios padrão (Lind, 2006)

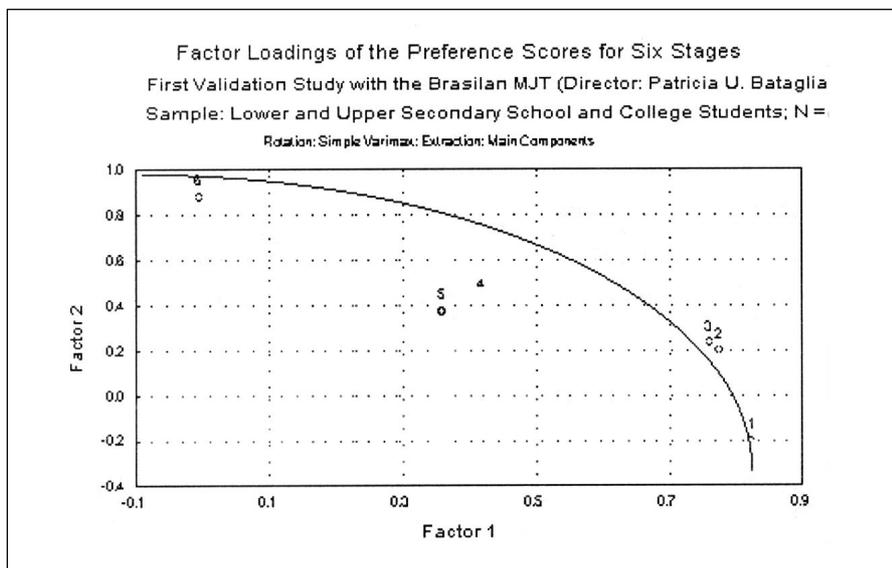


Figura 4. Estrutura quase-simplex – Validação do MJT em 1998 (Lind, 2006)

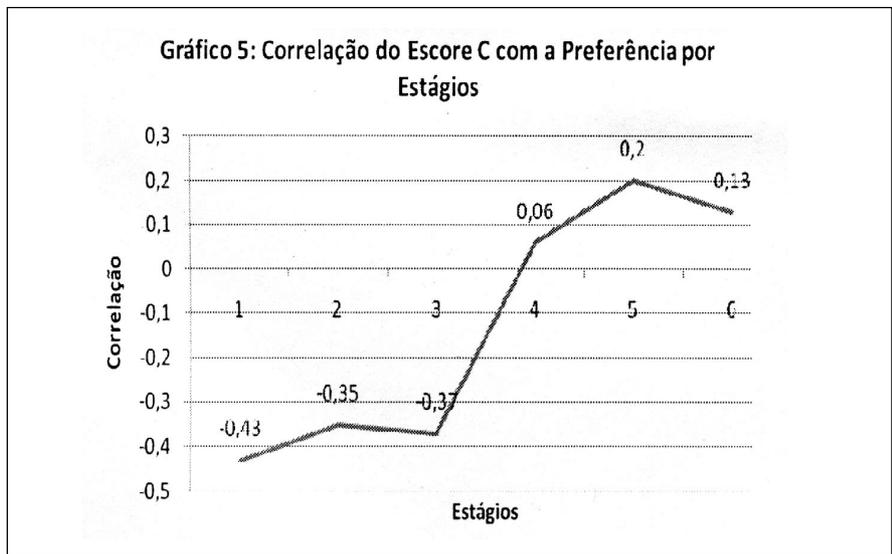


Figura 5. Correlação do escore C com a preferência por estágios (Bataglia, 2001)

altos de escore C devem correlacionar positivamente com a preferência pelos estágios mais altos, de modo neutro com os estágios intermediários e negativamente com os estágios inferiores.

A Figura 5 mostra a correlação entre estágio e preferência de um dos trabalhos de validação do MJT.

Por fim, com relação à quarta hipótese, o MJT deve ser resistente a falseamento por instruções. Por exemplo, os participantes não devem ser capazes de aumentar seu índice de escore C por receberem algum tipo de instrução particular (Lind, 2000). Sobre isso, Lind cita uma pesquisa interessante na qual dois grupos de sujeitos com posturas políticas divergentes (conservadores e liberais) são testados e convidados a responder novamente ao teste desta vez simulando ser parte do outro grupo. A

conclusão é que o grupo com escore inferior não conseguiu falsear seu escore de competência, elevando seu resultado.

*Confiabilidade*

Conforme, dito acima, a confiabilidade é um critério clássico de coerência.

A confiabilidade está “relacionada à estimativa do grau em que uma mensuração é livre de erro aleatório ou instável” (Cooper & Schindler, 2003, p. 186). São estimativas de confiabilidade: equivalência, estabilidade e consistência interna.

A equivalência é o grau em que as formas alternativas da mesma medida produzem resultados iguais ou similares. Não há formas equivalentes ao MJT. Lind (2006)

relata alguns estudos que comparam os índices de vários instrumentos, mas não com objetivo de evidenciar ou não equivalência.

A estabilidade se refere à consistência no tempo, isto é, a aplicação do teste na mesma amostra com um intervalo de tempo que pode variar de alguns dias a alguns anos deve apresentar resultados com um índice pequeno de variação, já que o mesmo constructo está sendo mensurado. Lerkiatbundit, Utaipan, Laohawiriyanon e Teo (2006) relatam resultados com índice de correlação no pré e pós teste na ordem de  $r=0,90$ .

O cuidado a se tomar nesse tipo de avaliação é realizar o pós teste em intervalo de tempo que não seja demasiadamente curto para não provocar a fadiga no sujeito, baixando os níveis de competência por variáveis intervenientes; e também não muito longo, podendo ocorrer interferência de processos de intervenção. Deve-se lembrar ainda que a avaliação do MJT nunca é feita para diagnósticos individuais, sempre trabalhamos com as médias de grupos.

A consistência interna se baseia no princípio de que cada parte do teste deve ser consistente com as outras partes (Cooper & Schindler, 2003). Para realizar essa verificação Spearman (1907, apud Cooper & Schindler, 2003) propõe a técnica do *split-half*, ou seja, divide-se os itens do teste ao meio de modo que os resultados de ambas as partes sejam correlacionadas. Tendo em vista que há diferentes formas de dividir o teste ao meio e as correlações podem ser diferentes, Kuder e Richardson (1937, apud Cooper & Schindler, 2003) propuseram uma outra fórmula de cálculo da correlação chamada KR20, desenvolvida para uso em escalas dicotômicas. Cronbach (1951) generalizou o conceito e a fórmula para escalas contínuas e propôs o cálculo que ficou conhecido como  $\alpha$  de Cronbach.

A análise da consistência interna do MJT pode ser feita usando-se o *split-half* e não o  $\alpha$  de Cronbach ou o KR20. Isso se deve ao fato de que os itens do teste podem ser consistentes ou não de acordo com nível de competência moral do grupo e isso não é um atributo do teste.

Comentaremos a seguir os resultados das pesquisas no Brasil e as peculiaridades com relação ao *split half* que geraram o desenvolvimento do MJT-xt.

### **Pesquisas no Brasil**

Houve uma situação muito interessante nas avaliações que ocorreram no Brasil e México. Isso será considerado a seguir.

O processo de validação de novas versões ou versões traduzidas deve seguir passos similares ao da validação alemã. Bataglia (2001) relata os passos de validação para a língua portuguesa. Os resultados foram considerados equivalentes ao da versão alemã. O processo para a validação para a língua portuguesa do Brasil foi:

*Tradução para a Língua Portuguesa e Retro-Tradução por um Tradutor Juramentado para a Língua Ingle-*

*sa.* A retro-tradução foi encaminhada ao autor do teste para que ele verificasse se a precisão do sentido teórico foi mantida (validade de conteúdo).

*Aplicação em Três Amostras de Vinte Sujeitos Cada (N=60), sendo Cada Amostra formada por Indivíduos de Diferentes Níveis de Educação Escolar.* Lind (2000) sugere que a testagem seja feita em diferentes níveis de educação dado que os resultados de pesquisas têm mostrado que a competência moral é sensível a educação formal e se correlaciona negativamente com a idade quando os indivíduos não participaram de processos educacionais formais. A aplicação foi realizada em uma classe de 6ª série, outra de 2º colegial e numa classe de terceiro ano de faculdade.

Os Resultados foram Enviados Novamente ao Autor, que ofereceu sua “Expertise” para Proceder à Avaliação, verificando que as Correlações estavam de Acordo com o Pretendido para Considerá-la Válida (Validade de Critério).

*Das Quatro Hipóteses, Três são Usualmente Testadas na Validação do MJT em Outras Línguas: Preferência por Estágios, Correlação Maior entre Estágios Vizinhos e Paralelismo Afetivo-Cognitivo.* Essas três hipóteses foram testadas na versão brasileira e os resultados foram positivos, isto é, a validade de constructo foi assegurada. As Figuras 1, 4 e 5 anteriormente apresentados mostram os resultados da avaliação de 2001.

Ocorreu que a comparação dos resultados do Brasil e México com os resultados dos Estados Unidos e Europa mostrou índices muito mais baixos nas pesquisas da América Latina do que nas demais pesquisas (Moreno, 2005). Realizou-se então uma segmentação da avaliação e passou-se a avaliar os dois dilemas que compõem o teste separadamente (*split half*). O resultado foi que para o dilema da eutanásia os índices eram muito baixos e para o dilema do operário, os índices eram equivalentes aos da Europa, evidenciando assim, uma falta de equivalência entre os componentes do teste.

Os pesquisadores têm investigado este fato de várias formas (Bataglia, Schillinger-Agati, et al., 2002; Bataglia et al., 2003; Bataglia, Schillinger, et al., 2006; Moreno, 2005; Schillinger, 2006) e algumas considerações são possíveis, ainda que não haja uma resposta plenamente satisfatória.

Lind (2005a) afirma que a segmentação do juízo moral mostra que o escore C não reflete apenas a competência individual e a dificuldade em executar a tarefa, mas também a poderosa influência dos agentes sociais como a igreja, o exército e outras instituições. Sobre esta influência no próprio juízo moral Bataglia (1996) já havia suspeitado, quando utilizando os dilemas de Kohlberg, especialmente um que trata de uma situação de guerra, os sujeitos tenderam a responder pela lei e ordem, ao passo que em outros dilemas respondiam segundo outros níveis de juízo moral.

Frente a estes resultados, Bataglia, Agati, et al. (2002), Bataglia, Schillinger-Agati, et al. (2002) e Schillinger

(2006) testaram a religiosidade como fator importante para a segmentação e de fato aparece no segundo estudo uma correlação significativa entre baixos escores C no dilema da eutanásia e uma filiação religiosa importante. Não se trata de atribuir à religiosidade a responsabilidade pelos baixos resultados no desenvolvimento da competência moral, mas cabe refletir a respeito do tipo de relação que se estabelece com a autoridade. Hinojosa (2000, segundo Moreno, 2005) propõe uma comparação entre sociedades que protestantes e católicas. A autora propõe que: sociedades que adotaram a reforma luterana tiveram acesso aos textos sagrados sem mediação, de modo autônomo, enquanto que sociedades contra-reformistas mantiveram tradições baseadas na autoridade do Papa como a única fonte autorizada para ler e interpretar a bíblia.

Santos (2000) trata da falência do paradigma sócio – cultural atual e propõe uma alternativa teórica e analítica do modelo presente. Quando inicia sua crítica, o autor apresenta uma imagem do espelho social que da mesma forma que os espelhos pessoais, em relação ao indivíduo, refletem o que a sociedade é. Os espelhos sociais são instituições, as regras, ideologias, a ciência, o direito, a educação, a informação, a religião e a tradição, por exemplo. Quanto mais intenso o uso desses espelhos, maiores eles se tornam, até tornarem-se super-espelhos, cuja importância ultrapassa a do próprio refletido. Tornam-se não só mais potentes, mas também mais rígidos, tornam-se estátuas. A estátua não reflete mais nada, ela é a própria sociedade. A realidade irrepreensível. Tome-mos essa imagem e apliquemo-la ao que parece ter ocorrido na segmentação observada nos dilemas do MJT. O que deveria ser um espelho da subjetividade e como espelho refletir a imagem dinâmica, alterando-se na medida da evolução do próprio conhecimento, transformou-se em estátua impedindo a visualização do objeto de estudo original. Na medida em que os sujeitos entram em contato com a estátua, perdem de vista o refletido: o Ser. Nesse momento, as respostas do sujeito perdem em competência moral porque perdem a dimensão do outro.

Sendo fato que os dilemas do operário e do médico não se mostravam equivalentes, passou-se a desenvolver outro dilema que trouxesse a mesma questão à tona, ou seja, o valor da vida, mas que não tocasse diretamente em temas religiosos típicos da cultura latino-americana. Bataglia et al. (2006) apresenta a padronização final do dilema do juiz que passou a integrar o MJT, que passou a ser chamado MJTtxt (MJT estendido). No dilema do juiz, está presente a questão do valor da vida de uma pessoa em contraposição a vida de muitas pessoas.

Trata-se de uma situação de atentado em que uma pessoa suspeita de conhecer os detalhes da operação é detida e se recusa a falar. A questão que se coloca é de autorizar a tortura dessa mulher para salvar a vida de pessoas inocentes ou não.

Os escores alcançados para esse dilema mostram-se em média muito próximos dos escores obtidos no dilema do operário.

As Figuras 1 e 2 mostram a diferença de escores quando trabalhamos com o MJT sem segmentação (Figura 1) e usando só o do juiz (Figura 2).

Recomenda-se atualmente que no Brasil os pesquisadores utilizem o MJTtxt composto por três dilemas e utilizem a avaliação do escore C total e segmentado.

### Questões para Pesquisas Futuras

Há basicamente duas linhas de pesquisas desenvolvidas no Brasil atualmente: a primeira que investiga a influência da educação formal na construção da competência moral, mais especificamente, a influência do ambiente acadêmico, e a segunda que investiga a influência da religião nos resultados apresentados na América Latina.

Cito Silva, Bataglia e Aprile (2008) que apresentaram um estudo a respeito da construção da competência moral em estudantes de pedagogia. O estudo levantou o escore C no primeiro, segundo e terceiro anos do curso e constatou que não houve alteração significativa. Isso leva a pensar em como o ambiente acadêmico, entendido em seus aspectos curriculares, semi-curriculares e extra-curriculares exerceram pouca ou nenhuma influência na formação ética dos profissionais.

A respeito das pesquisas que envolvem a relação entre religião e a competência moral, realizou-se este ano a validação de uma escala de crenças, o PCBS ([*Post-Critical Belief Scale*], Duriez, Soenens, & Hutsebaut, 2005). Após a publicação do estudo de validação serão iniciados os estudos a respeito da relação entre o tipo de vinculação dos sujeitos a crenças religiosas e a competência moral.

### Referências

- Bataglia, P. (1996). *Um estudo sobre o juízo moral e a questão ética na prática da Psicologia*. Dissertação de Mestrado não-publicada, Instituto de Psicologia, Universidade de São Paulo, SP.
- Bataglia, P. (2001). *A construção da competência moral e a formação do psicólogo*. Tese de Doutorado não-publicada, Instituto de Psicologia, Universidade de São Paulo, SP.
- Bataglia, P. (2006). *Moral segmentation in MJT studies: Cultural influences*. Paper presented at the meeting of the Association for Moral Education, Fribourg, Switzerland.
- Bataglia, P., Agati, M. M., Torres, S. S., Crivelaro, D. B. Z., Oliveira, D. D., & Quevedo, T. L. (2002). *The development of moral competence and religious commitment in Brazil*. Paper presented at the meeting of the Association for Moral Education, Chicago.
- Bataglia, P., Schillinger-Agati, M., Lind, G., & Quevedo, T. L. (2003). *Testing the segmentation hypothesis with an extended version of the MJT*. Poster presented at the meeting of the Association for Moral Education, Krakow, Poland.

- Bataglia, P., Schillinger-Agati, M., Torres, S. S., Crivelaro, D. B. Z., Oliveira, D. D., & Quevedo, T. L. (2002, November). *The development of moral competence and religious commitment in Brazil*. Paper presented at the meeting of the Association for Moral Education, Chicago.
- Bataglia, P. U. R., Schillinger, M., & Lind, G. (2006). *Moral segmentation in MJT studies: Cultural influences*. Paper presented at the meeting of the Association for Moral Education, Fribourg, Switzerland.
- Biaggio, A. M. B. (2002). *Lawrence Kohlberg: Ética e educação moral*. São Paulo, SP: Moderna.
- Biaggio, A. M. B., & Brandão, M. S. B. (1990). Adaptação brasileira de uma medida objetiva de julgamento moral. *Arquivos Brasileiros de Psicologia* (Rio de Janeiro), 42(1), 107-119.
- Biaggio, A. M. B., Shimizu, A. M., & Martinez, T. M. (1998). *DIT-2 Opiniões sobre problemas sociais: Tradução e adaptação do Defining Issues Test - 2*. Minneapolis, MN, USA: University of Minnesota.
- Boyes, M. C., & Walker, L. J. (1988). Implications of cultural diversity for the universality claims of Kohlberg's theory of moral reasoning. *Human Development*, 31, 44-59.
- Breakwell, G. M., Hammond, S., Fife-Schaw, C., & Smith, J. A. (2006). *Research methods in psychology* (3rd ed.). London: Sage.
- Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment*. Thousand Oaks, CA: Sage.
- Colby, A., Kohlberg, L., Abraham, A., Gibbs, J., Higgins, A., & Speicher, B. (1987). *The measurement of moral judgment* (Vols. 1-2). New York: Columbia University Press.
- Cooper, D., & Schindler, P. (2003). *Métodos de pesquisa em administração*. Porto Alegre, RS: Bookman.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.
- Damon, W. (1977). *The social world of the child*. San Francisco: Jossey-Bass.
- Duriez, B., Soenens, B., & Hutsebaut, D. (2005). Introducing the shortened Post-Critical Belief Scale. *Personality and Individual Differences*, 38, 851-857.
- Gibbs, J., Basinger, K., & Fuller, D. (1992). *Moral maturity: Measuring the development of sociomoral reflection*. Trenton, NJ: Lawrence Erlbaum.
- Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Cambridge, MA: Harvard University Press.
- Gross, M. (1997). *Ethics and activism: The theory and practice of political morality*. Cambridge, MA: Cambridge University Press.
- Hair, J. F., Jr., Babin, B., Money, A. H., & Samouel, P. (2005). *Fundamentos de métodos de pesquisa em administração*. Porto Alegre, RS: Bookman.
- Hala, S. (1997). *Development of social cognition*. East Sussex, UK: Psychology Press.
- Hann, N., Aerts, E., & Cooper, B. (1985). *On moral grounds: The search for practical morality*. New York: New York University Press.
- Kohlberg, L. (1964). Development of moral character and moral ideology. In M. L. Hoffman & L. W. Hoffman (Eds.), *Review of child development research* (Vol. 1, pp. 381-431). New York: Russel Sage Foundation.
- Lerkiatbundit, S., Utaipan, P., Laohawiriyanon, C., & Teo, A. (2006). Randomized controlled study of the impact of the Konstanz method of dilemma discussion on moral judgement. *Journal of Allied Health*, 35(2), 101-108.
- Lind, G. (1980). *Reliability of test or structure of personality? Empirical findings on the problem of interpreting judgment consistency*. Paper presented at the meeting FORM-Document Nr. 88 of the European Coordination Center for Research and Documentation in Social Sciences, Portoroz, Yugoslavia.
- Lind, G. (1986). Cultural differences in moral judgment? A study of West and East European University Students. *Behavioral Science Research*, 20, 208-225.
- Lind, G. (2000). O significado e medida da competência moral revisitada: Um modelo do duplo aspecto da competência moral. *Psicologia: Reflexão e Crítica*, 13(3), 399-416.
- Lind, G. (2005a). *The cross-cultural validity of the Moral Judgment Test (MJT). Confirmation of 17 Cross-Cultural Adaptations*. Paper presented at the MOSAIC 2005, Konstanz, Germany.
- Lind, G. (2005b). *Validation and certification procedure for the Moral Judgment Test*. Retrieved March 20, 2006, from <http://www.uni-konstanz.de/ag-moral/mut/mjt-certification.htm>
- Lind, G. (2006). *Introduction to the Moral Judgment Test (MJT)*. Retrieved March 11, 2007, from <http://www.uni-konstanz.de/ag-moral/mut/mjt-intro.htm>
- Moreno, C. (2005). *Moral education in higher education and the transformation of a concern: A historical account*. Paper presented at the annual meeting of the Association for Moral Education, Cambridge, MA
- Piaget, J. (1976). The affective unconscious and the cognitive unconscious. In B. Inhelder & H. H. Chipman (Eds.), *Piaget and his school* (pp. 63-71). New York: Springer.
- Rest, J. R. (1969). Level of moral development as a determinant of preference and comprehension of moral judgments made by others. *Journal of Personality*, 37, 220-228.
- Rest, J. R., & Narvaez, D. (1998). *Guide for DIT - 2*. Minneapolis, MN: University of Minnesota.
- Santos, B. S. (2000). A crítica da razão indolente: Contra o desperdício da experiência. In *Para um novo senso comum: A ciência, o direito e a política na transição paradigmática* (2. ed., Vol. 1). São Paulo, SP: Cortez.
- Schillinger, M. (2006). *Learning environment and moral development: How university education fosters moral judgment competence in Brazil and two German-speaking countries*. Aachen, Germany: Shaker Verlag.
- Silva, M. C., Bataglia, P. U. R., & Aprile, M. R. (2008). *A ética na formação do professor: Uma investigação a respeito da construção da competência moral*. Poster presented at the meeting of the Sociedade Brasileira de Psicologia Clínica, Campinas, SP.
- Stein, N. L., Trabasso, T., & Garfin, D. (1979, September). Comprehending and remembering moral dilemmas. In S. Goldmans, *Understanding discourse: Interactions between knowledge and process*. Symposium presented at the meeting of the American Psychological Association, New York.