

UM EXERCÍCIO DE AVALIAÇÃO DA QUALIDADE DA PESQUISA AGROPECUÁRIA NA EMBRAPA¹

TARCÍZIO REGO QUIRINO² e ERYCSON PIRES COQUEIRO³

RESUMO - Medir qualidade da pesquisa é um empreendimento polêmico e difícil, que raramente tem sido enfrentado. O trabalho propõe um método globalizante para medir a qualidade da pesquisa e o aplica às unidades da Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA). Baseia-se teoricamente na concepção de status social aplicando-a às unidades da organização. Um painel de 44 juízes, formados por especialistas, foi usado como a fonte de julgamentos, dos quais foram gerados os Índices de Qualidade, de Concordância e de Visibilidade. A medida de qualidade foi validada, apresentando o coeficiente de fidedignidade de $r = 0,890$ através do processo teste-reteste com apenas seis juízes, após um intervalo médio de aplicação de seis meses. Foram estudadas a validade lógica e do construto, chegando-se à conclusão de que o Índice de Qualidade discrimina de modo consistente, previsível e coerente os objetos submetidos à mensuração. Com a finalidade de estudar as discrepâncias entre os julgamentos, o painel de juízes foi submetido a análise no que se refere a quatro características: localização do juiz na estrutura organizacional, nível educacional, experiência internacional e conhecimento do sistema. A educação revelou-se como o maior correlato com as diferenças de julgamento. Conclui-se que os Índices de Qualidade, Concordância e Visibilidade podem servir como instrumento para a análise e melhoria das organizações de pesquisa agropecuária e para aperfeiçoamento das técnicas de mensuração.

Termos para indexação: pesquisadores, qualidade da pesquisa, EMBRAPA, pesquisa agropecuária.

AN EXERCISE OF AGRICULTURAL RESEARCH QUALITY EVALUATION IN EMBRAPA

ABSTRACT - Measuring research quality is a polemic and difficult task, which has been seldom undertaken. The present paper proposes a method of measuring research quality and applies it to the Brazilian Agricultural Research Corporation (EMBRAPA) research units. It is theoretically based on the conception of social status as applied to organizational units. Judgements have been provided by a pannel of 44 judges, all of them specialists of agricultural research, and Indices of Quality, Agreement and Visibility have been computed thereupon. The quality measure was validated, allowing a test-retest reliability coefficient $r = 0,890$, with only six judges after a mean application interval of six months. Examination of the logic and construct validity yielded the conclusion that the Quality Index consistently, predictably and coerently discriminates between objects submitted to measurement. The judges' pannel was submitted to closer scrutiny, aiming at understanding the discrepancies of judgements. Among the four aspects taken into account, namely, judge's location in the organizational structure, educa-

¹ Recebido em 15 de janeiro de 1985.
Aceito para publicação em 24 de julho de 1985.

² Ph.D. em Sociologia, Assessor do Departamento de Recursos Humanos (DRH/EMBRAPA) - Ed. Venâncio 2000 - 6º andar - CEP 70333 - Brasília, DF.

³ Pesquisador do CENARGEN/EMBRAPA - Ex-chefe do Departamento de Recursos Humanos (DRH/EMBRAPA) - Av. W/5 Norte - Parque Rural SAIN - Caixa Postal 102372 - CEP 70770 - Brasília, DF.

tional level, international experience and acquaintance with the system, education was found to be the closest one to correlate with judgement differences. It is concluded that the Indices of Quality, Agreement and Visibility are useful tools for analysis and improvement of agricultural research organizations as well as for improving the measurement techniques.

Index terms: research workers, research quality, EMBRAPA, agricultural research.

MEDIR QUALIDADE DA PESQUISA ?

Medir qualidade da pesquisa tem-se tornado cada vez mais importante, à medida que cresce o número de pesquisadores e de organizações de pesquisa, que diminui a quantidade relativa dos meios disponíveis a dividir entre eles, e que aumenta o interesse e o controle da sociedade sobre os resultados do trabalho científico.

A idéia de medir qualidade assusta muita gente, especialmente a quem não está familiarizado com as modernas técnicas de mensuração das ciências sociais. Apesar disso, as pessoas são capazes de dizer se um pesquisador é melhor do que outro, se uma organização de pesquisa produz trabalhos de qualidade superior, se uma universidade contribuiu mais do que as demais para o progresso da ciência. Medir qualidade de pesquisa significa, basicamente, repetir de modo sistemático, explícito e preciso, o processo que as pessoas usam para gerar os seus julgamentos individuais, atendendo, ao mesmo tempo, as exigências técnicas das ciências sociais a respeito de mensuração.

A Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA), por ser uma organização cuja finalidade é produzir tecnologia aplicável à agropecuária brasileira, necessita monitorar de modo constante e sistemático a qualidade de seu produto. Por isso, qualquer progresso que se fizer na mensuração desta, potencialmente trará benefícios à melhoria da organização. A integração entre a EMBRAPA e os demais órgãos do Sistema Cooperativo de Pesquisa Agropecuária (SCPA), os quais são apoiados por ela em seu trabalho de geração de tecnologia, torna ainda mais importante esse monitoramento.

Em data recente, os economistas Evenson & Wright (1983) afirmaram que, ao examinarem os métodos para avaliar a qualidade da pesquisa agropecuária, não conseguiram localizar nenhum método formal de avaliar e, até mesmo, de defini-la.

Duas ordens de dificuldades devem ser enfrentadas para que se possa medir a qualidade da pesquisa. Em primeiro lugar, há o problema do entendimento do que seja qualidade, quando se trata de pesquisa. Depois, deve-se enfrentar o problema metodológico da escolha de indicadores e coleta dos dados apropriados.

O que se vai expor a seguir é um exercício que explora um dos possíveis caminhos para resolver o problema da mensuração da qualidade da pesquisa, no contexto específico da pesquisa agropecuária brasileira. Contudo, o caráter do trabalho sugere que ele seja partilhado com o público interessado, de modo a permitir seu aprimoramento e a incentivar sua aplicação.

“Qualidade” é um conceito que implica em comparação. Uma coisa ou evento

tem boa ou má qualidade se está ou não de acordo com critérios pertinentes, — implícitos ou explícitos. Determinar como chegar a critérios apropriados para basear neles o julgamento é, pois, o ponto essencial da questão.

Evenson & Wright (1983) chamam a atenção para o fato de que o termo "qualidade da pesquisa" significa coisas diferentes em contextos diferentes. Essa assertiva pode ser ilustrada em outra área, através da abundante literatura sobre a qualidade da educação. Os conceitos de qualidade se abrem em leque, focalizando, desde as instalações físicas da sala de aula, o modo de produzir a instrução, a quantidade de aprendizagem decorrente, até os efeitos causados pela educação na alocação do seu produto na estratificação social. No caso específico da pesquisa agropecuária, qualidade pode significar, para o economista, o resultado altamente favorável da análise custo benefício, para o agricultor, a possibilidade de aumentar seus lucros ou sua produção, para o cientista acadêmico, "a conformidade com os padrões estabelecidos pelo trabalho científico na fronteira da disciplina".

O caráter aplicado da pesquisa agropecuária exige que a ocupação do pesquisador seja exercida em um contexto social em que os padrões de julgamento acadêmicos se sobrepõem e, às vezes, conflitam com os padrões de julgamento utilitários (Quirino, 1981). Os estudos até aqui existentes procuram explorar um ou outro desses padrões. Com efeito, as avaliações *ex post* têm procurado medir o retorno da pesquisa (Norton & Davis, 1981; Anderson, 1983) ou então algum aspecto do desempenho do sistema de pesquisa (Wright, 1983). Ao se medir o retorno da pesquisa, o interesse central é determinar o valor econômico conseguido pelo avanço de conhecimento originário da pesquisa e compará-lo com o que foi gasto em sua geração. Se a ênfase é em algum aspecto específico do desempenho do sistema, em geral são escolhidos indicadores mensuráveis de qualidade científica entre os quais os principais são a quantidade de publicações, de patentes conseguidas e de citações em outros trabalhos científicos. Quando se trata de publicações, o aspecto qualitativo pode ser ressaltado através da consideração seletiva da qualidade do veículo de publicação.

A presente abordagem sugere que o objeto de referência, ou seja, a qualidade da pesquisa, comporta um julgamento final por parte dos indivíduos, o qual é composto dos julgamentos parciais feitos por estes sobre os aspectos específicos da pesquisa. Os produtores de pesquisa (pesquisadores individuais, grupos ou organizações) recebem julgamentos que são compartilhados socialmente pelas pessoas que os fazem, terminando por cristalizar-se em forma de nível de prestígio e deferência, de modo semelhante ao que acontece com o prestígio ocupacional ou dos indivíduos (status). Assim sendo, propomos um método globalizante de aferição da qualidade da pesquisa, através do qual é possível captar os julgamentos de indivíduos que sejam familiares com a mesma e compartilhem dos critérios de cientificidade e aplicabilidade. Pelo tratamento sistemático desses julgamentos, chega-se à construção de indicadores do status que cada unidade produtora de pesquisa desfruta na parcela pertinente da comunidade científica. Pelz & Andrews (1976) empregaram

método semelhante para medir a variável dependente (desempenho) no seu trabalho clássico sobre cientistas e organizações, referindo-se, porém, aos pesquisadores e não às suas instituições.

O julgamento da qualidade da pesquisa será tanto mais fidedigno quanto mais os indivíduos que a julgam tiverem conhecimento das pesquisas. Será tanto mais homogêneo, quanto mais os juízes compartilharem do conjunto de valores determinantes dos critérios a serem julgados. Será tanto mais geral, quanto mais aspectos da qualidade da pesquisa forem levados em consideração para o julgamento. A metodologia adotada procurou maximizar tais relações, como veremos a seguir.

É possível tomar uma, dentre diferentes unidades de análise, para medir qualidade da pesquisa. Podemos concentrar a atenção no projeto, no pesquisador, na unidade organizacional que produz as pesquisas ou no total de pesquisas produzidas em, digamos, um País, Região ou Estado. No presente caso, o interesse está focalizado para o nível organizacional. Por isso definimos como unidade de análise aquela representada por cada unidade de pesquisa com independência administrativa relativa dentro da EMBRAPA. Isto permite aplicar os resultados à análise de aspectos organizacionais e administrativos.

A abordagem de qualidade aqui adotada assemelha-se ao conceito de efetividade adotado nos estudos de organização complexa. É o grau em que um sistema social realiza suas metas. É também chamado de desempenho e de sucesso organizacional (Price, 1972). Esta abordagem exige que se distinga com precisão as metas da organização e que essas sejam levadas em consideração como ponto central de referência no processo de mensuração. Ambas as exigências foram amplamente atendidas, tanto pela existência de farta bibliografia que testemunha as metas e propósitos da EMBRAPA (Alves, Pastore, Pastore, 1980; Gastal, 1980; Blumenschein, 1979; Quirino, 1981), como pela escolha da metodologia de aplicação do Índice.

O método usado para este exercício baseia-se na sistematização das opiniões de pessoas capacitadas, envolvidas com a administração da pesquisa na EMBRAPA e profundas conhecedoras da pesquisa agropecuária no Brasil, e é uma adaptação da técnica conhecida como "painel de juízes" (Goode & Hatt, 1960). As pessoas que formaram o painel de juízes, são pesquisadores exercendo funções na Empresa, as quais permitem obter uma visibilidade do que se passa em toda ela ou, pelo menos, em um grande segmento da mesma. A circunscrição da escolha dos juízes aos membros da Empresa visa aumentar a probabilidade de que compartilhem dos valores básicos de cientificidade e de aplicabilidade adotados por esta, julgando, pois, a qualidade da pesquisa de cada Unidade nos termos que a Empresa adota para si própria como um todo.

Os dados foram coletados durante o segundo semestre de 1983. Pediu-se que os juízes classificassem "cada unidade de pesquisa de acordo com seu julgamento sobre a qualidade da pesquisa agropecuária que se desenvolve nela. Levar em consideração todos os aspectos da qualidade da pesquisa". Através deste estímulo, pro-

curou-se evitar que um aspecto particular, como a contribuição para o progresso da ciência ou para o desenvolvimento da agricultura, fosse privilegiado em detrimento dos demais aspectos. Aos que hesitaram em fornecer suas opiniões, por considerá-las pouco fundamentadas em evidência, foi explicado que se tratava de medir o status relativo das diversas unidades, ou seja, o prestígio e a fama que são atribuídos a cada uma por causa da qualidade da pesquisa. Os resultados, pois, devem ser considerados uma medida de percepção coletiva dos status das unidades de pesquisa, no aspecto específico da qualidade. O elevado nível do grupo de pessoas que detêm essa percepção, determina a importância e utilidade dos índices obtidos a partir da mesma.

Foram excluídas da classificação as seguintes Unidades: CNP - Defensivos Agrícolas, UEPAEs de Cáceres, Boa Vista e Macapá, Unidade de Apoio ao Programa Nacional em Biologia do Solo, Serviço de Produção de Sementes Básicas e as Representações. Estes dois últimos foram excluídos por não se dedicarem à pesquisa. Os demais, porque foram organizados a partir de 1982.

O painel de juízes foi formado por 30 chefes das unidades de pesquisa e por 14 pessoas da Sede da EMBRAPA. Todos os chefes das unidades estudadas participaram no painel, com exceção, de dois que não se renderam aos argumentos de que, com o conhecimento que possuíam, seriam capazes de classificar as demais unidades. As pessoas da sede fazem parte da Diretoria ou são Chefes de Departamentos e Técnicos que estão diretamente envolvidos com o acompanhamento da pesquisa na Empresa e nos demais órgãos que compõem o Sistema Cooperativo de Pesquisa Agropecuária - SCPA. Assim, todos os participantes do painel pertencem a um reduzido grupo de pessoas que possuem visão muito próxima do que vem sendo produzido na pesquisa agropecuária do país, assim como das metas da EMBRAPA. Quase todos têm formação agrônoma e grande experiência como pesquisadores. Os cargos que ocupam demonstram o reconhecimento público da liderança e do valor como cientistas. Seus julgamentos, como um bloco, certamente refletem de modo bastante aproximado a verdadeira qualidade da pesquisa e seguramente influenciam poderosamente a percepção que outras pessoas menos informadas têm sobre o padrão qualitativo da pesquisa que vem sendo produzida nas diferentes unidades da EMBRAPA.

Os resultados apresentados a seguir sintetizam e discutem esses julgamentos.

A ESTRATIFICAÇÃO DAS UNIDADES SEGUNDO A QUALIDADE DA PESQUISA

Os juízes classificaram as unidades em três níveis de qualidade, definidos como: 1 = Superior, 2 = Média e 3 = Inferior. Absteram-se de julgar suas próprias unidades, assim como aquelas sobre as quais não se consideraram possuidores de informações suficientes para fundamentar de modo aceitável seus julgamentos. Foram então criados o Índice de Qualidade, o Índice de Concordância e o Índice de Visibilidade referentes a cada uma das unidades. Com base em cada um desses Índices

foram construídas as colunas "quartil", o qual agrupa as unidades em estratos de acordo com sua posição relativa na distribuição. A partir desta, foi organizada a coluna "ordem".⁴

A estratificação das unidades segundo a qualidade da pesquisa

A média representa a medida mais direta de qualidade da pesquisa agropecuária das diversas unidades de pesquisa, pois foi obtida a partir do julgamento de todos os juízes, exceto daqueles que se abstiveram. Por isso chamamo-la de Índice de Qualidade. A relação das unidades na ordem decrescente segundo o Índice de Qualidade representa em si o mais substantivo resultado da pesquisa. Nela serão baseadas muitas das observações que se seguem.

Pode-se observar que as médias próximas de 1 (qualidade superior) chegam mais perto desse número do que as médias próximas de 3 (qualidade inferior). Se dividirmos a escala em partes iguais, poderemos considerar as médias localizadas entre 1,0 e 1,6 como correspondentes à qualidade superior; as médias entre 1,7 e 2,4 como a qualidade média e aquelas entre 2,5 e 3,0, como de qualidade inferior. Os resultados mostram que, de acordo com tal critério, apenas uma unidade de pesquisa foi considerada como inferior, sendo que nove se classificaram como superior.

Os quartis podem ser interpretados como julgamentos relativos, pois demonstram a posição relativa da unidade na distribuição empírica dos dados. O primeiro quartil corresponde a "alto" na ordem do atributo, os segundo e terceiro, a "moderado" e o quarto a "baixo". A média do Índice de Qualidade no primeiro quartil é de 1,354, no segundo, de 1,676, no terceiro, de 1,911 e no quarto, de 2,279.

A unidade melhor classificada alcançou o índice de 1,129, enquanto a mais baixa chegou a 2,556. Isso demonstra uma propensão dos juízes para ver as unidades de modo positivo. A diferença entre a mediana (1,762) e a média (1,810) pode ser considerada uma medida dessa tendência para classificações positivas (0,048), visto que em uma distribuição normal tal diferença não deveria existir. Mesmo assim, há tendência para a distribuição normal das médias agrupadas, apesar do número de casos ser relativamente pequeno para isso. A média do Índice de Qualidade (1,810) está contida nos limites do conceito "qualidade média" tendendo, porém, para a qualidade superior.

No primeiro quartil, só aparecem Centros de Pesquisa e todos estão localizados na região Centro-Sul. No último quartil há predominância de unidades da região amazônica.

Os três centros que foram considerados os melhores, são os de Fruticultura de Clima Temperado, Soja e Trigo. Um deles se localiza no Paraná (Soja) e dois no

⁴ Como é usual em pesquisas semelhantes, a Tabela básica a que o texto se refere deixa de ser publicada para evitar a individualização dos resultados referentes a cada unidade. No texto serão feitas algumas excessões a este uso.

Rio Grande do Sul.

Como é sabido, a estrutura organizacional da EMBRAPA se compõe de unidades que têm diferentes objetos de pesquisa: produtos, recursos, serviços e problemas específicos ao âmbito de estados ou territórios. Quais são consideradas as melhores unidades, no que diz respeito ao objeto de pesquisa a que se dedicam? Esta resposta pode ser obtida através de cálculo da média do índice de qualidade em cada tipo de unidade. Os resultados nos mostram que os centros de produtos são considerados os melhores (1,650), seguidos dos de serviços (1,790) e dos de recursos (1,778), deixando por último as UEPAEs (1,952). Aplicando-se, porém, o teste de diferença de médias, as unidades podem ser reduzidas a três grupos, vistos que as médias dos centros de recursos e de serviços não se mostram significativamente diferentes a um nível de probabilidade de 0,05. Por sua vez, os centros de serviços não diferem significativamente, ao mesmo nível, das UEPAEs, porém se aproximam mais dos de recursos.

Como se comparam os centros nacionais de produtos agrícolas, com aqueles que pesquisam produtos animais? As médias apontam significativa superioridade dos primeiros (1,563), que são um total de dez, sobre os quatro dedicados à pecuária (1,971).

Quais são as melhores unidades em cada categoria, de acordo com o objeto de pesquisa a que se dedicam? Dentre os quatorze centros de produtos, os três primeiros colocados são os de Fruticultura de Clima Temperado, de Soja e de Trigo. É bom lembrar que o Centro Nacional de Fruticultura de Clima Temperado só recentemente assumiu esse papel tendo, pois, conquistado seu prestígio ao tempo que era denominado UEPAE Cascata. Das três unidades cujo objeto de pesquisa são os recursos, a dos Cerrados foi tida como a melhor. Entre as 12 Unidades Especiais de Pesquisa de Âmbito Estadual ou Territorial (UEPAEs) as de Bento Gonçalves, Teresina e Bagé são as três que sobressaíram. Dos três centros de serviços, o de Recursos Genéticos foi considerado o primeiro em qualidade. Além do CNP Fruticultura de Clima Temperado, Soja e Trigo, figuram ainda no primeiro quartil, os Centros dos Cerrados, Arroz e Feijão, Hortaliças, Milho e Sorgo e CENARGEN.

Não é nossa intenção oferecer aqui explicações para as diferenças de status, mas é bom lembrar que os resultados não podem ser atribuídos aprioristicamente ao mérito ou demérito dos pesquisadores e das unidades. Certamente que esses elementos os influenciam. Mas a tradição na pesquisa, a existência de conhecimentos transferíveis de outros países a respeito dos assuntos pesquisados, a qualidade educacional e procedência acadêmica do pessoal e a idade da unidade, são fatores a que é possível atribuir parte dos resultados.

Assim, as causas para explicar tais resultados nem sempre são desabonadoras. Além disso, em qualquer classificação ordinal, forçosamente existem os últimos, assim como existem os primeiros.

O nível de concordância

Às vezes existe acordo quase geral de que uma unidade produz pesquisa de um nível determinado, enquanto outras vezes há desacordo. O Índice de Concordância mede o nível de acordo entre os juízes, usando para isso o cálculo do desvio padrão.

As três unidades consideradas as melhores, (Fruticultura de Clima Temperado, Soja e Trigo) receberam também maior índice de concordância, fortalecendo, assim, seu lugar privilegiado. Das demais unidades que figuram no primeiro quartil em qualidade, dois Centros aparecem também no primeiro quartil em concordância. Um outro recebeu uma concordância moderada, havendo discordância acentuada a respeito das classificações dos dois restantes.

Os dez centros nacionais que fazem pesquisa vegetal reuniram em torno de si mais concordância (0,584) do que os quatro da pesquisa animal (0,607). A média do Índice de Concordância é de 0,604.

Levado em consideração o objeto da pesquisa, há mais concordância quanto ao julgamento dos centros de produtos (média do Índice de Concordância = 0,670) que são seguidos pelas UEPAEs e centros de recursos (respectivamente 0,690 e 0,723), havendo maior discordância sobre o julgamento da qualidade dos serviços (0,743). Esses resultados são coerentes com o que se observa ao estudar a estruturação da EMBRAPA, e com o sentimento corrente nas diversas Unidades.

O problema da visibilidade

As unidades de pesquisa diferem quanto à oportunidade que as pessoas têm de conhecerem o que nelas está sendo feito e de terem, portanto, uma opinião sobre a qualidade da pesquisa ali produzida. O Índice de Visibilidade mede tal diferença. Todas as unidades foram submetidas ao julgamento de 44 juízes, mas não se permitiu que os chefes julgassem suas próprias unidades.

Diversos juízes não atribuíram classificação às unidades que julgaram não lhes serem suficientemente visíveis. O Índice de Visibilidade da unidade foi calculado como sendo a proporção em que ela foi julgada. Alcançou a média de 71,7.

A mais visível de todas as Unidades, de acordo com os resultados da pesquisa, é o Centro Nacional de Recursos Genéticos (CENARGEN, Índice de Visibilidade: 90,7), seguida de perto pelo SNLC de Solos (88,4) e CNP Soja (86,0). Em geral as Unidades de Serviços são mais visíveis (média do Índice de Visibilidade = 81,4), seguidos pelos Centros de Recursos (80,6) e Centros de Produtos (73,4). As UEPAEs são as menos visíveis (64,9) como seria de esperar por sua função no sistema e pela localização geográfica quase sempre afastada dos pontos de acesso mais fácil no país. No que se refere ao produto, os dez centros nacionais de produtos agrícolas são ligeiramente mais visíveis (73,7) do que os quatro de produtos animais (72,7).

Qualidade, concordância, visibilidade

As três variáveis medidas neste exercício são características de todas as organizações, embora não tenham sido exploradas em conjunto pela literatura pertinente. É, pois, de importância teórica saber quais as relações que se poderiam encontrar entre elas.

Como hipótese de pesquisa, sugere-se que há uma relação curvilínea entre qualidade e concordância em forma de J, e uma relação direta entre qualidade e visibilidade (Tabela 1-a). Espera-se, pois, que nas unidades onde se observa a alta qualidade de pesquisa, seja observado também alto nível de visibilidade e de concordância de julgamento sobre a qualidade. As unidades de nível intermediário de qualidade, devem apresentar nível intermediário de visibilidade e baixo nível de concordância. Finalmente, aquelas de baixo nível de qualidade, devem apresentar nível baixo de visibilidade e nível intermediário de concordância.

As justificativas para tais previsões são as seguintes: a existência de padrão explícito de qualidade na comunidade e a valorização positiva desse padrão fazem com que as unidades que o possuem sejam ao mesmo tempo muito visíveis e recebam um julgamento homogêneo. As demais unidades têm, assim, níveis mais baixos em ambas as características. As unidades de nível intermediário não são tão visíveis, porque não servem como ideal de referência, nem são recompensadas com tantas alusões positivas pelo seu desempenho. Deve haver baixa concordância, porque alguns juízes podem ser levados a errar por um extremo e outros, por outro. Nas unidades de baixo nível de qualidade, a visibilidade é ainda mais baixa, porém o nível de concordância é intermediário, porque não funcionam nenhuma das razões que o levariam para os extremos.

A Tabela 1-b apresenta as médias dos Índices de Visibilidade e Concordância de acordo com o nível de qualidade das unidades. Os dados confirmam a predição teórica (é oportuno lembrar que, quanto mais baixo o índice de concordância, maior o nível de concordância).

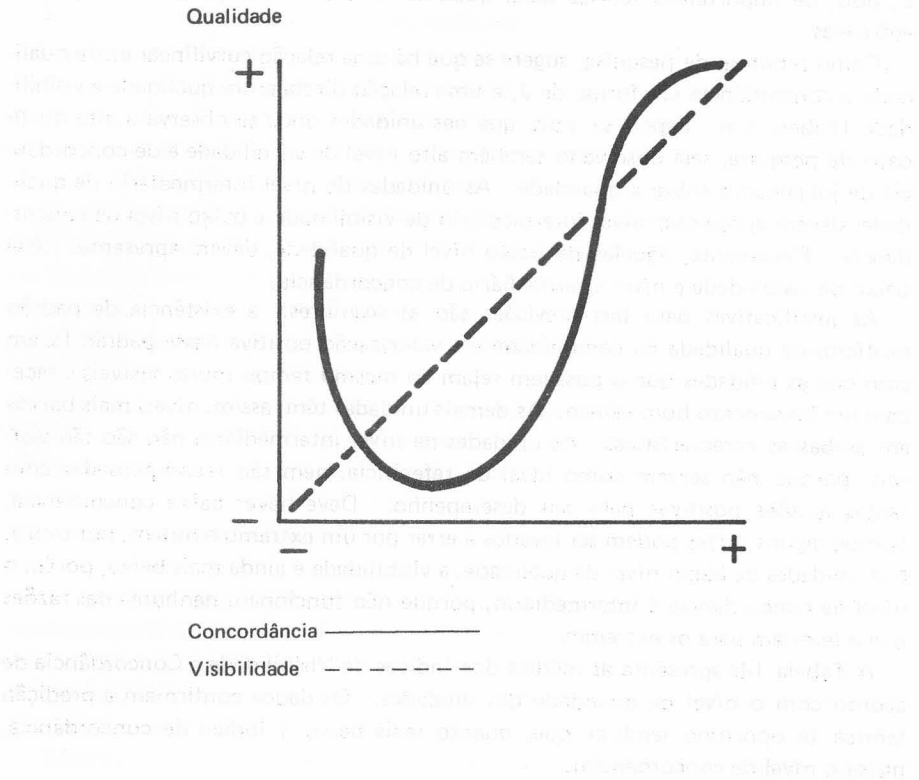
VALIDAÇÃO DA MEDIDA DE QUALIDADE DA PESQUISA

Fidedignidade

“Uma escala é fidedigna quando, aplicada à mesma amostra, produz consistentemente o mesmo resultado”. (Goode & Hatt, 1964). A fidedignidade da escala foi medida através do processo de teste-reteste. Foi pedido a seis dos juízes que classificassem novamente as unidades de pesquisa. Os resultados foram correlacionados com a classificação feita por eles no primeiro teste, e produziram um coeficiente de correlação de Pearson $r = 0,890$. Este resultado deve ser interpretado como o coeficiente de validade ou de fidedignidade do Índice (Kerlinger, 1980). É notável que tenha atingido nível tão alto em replicação com apenas seis juízes, após trans-

TABELA 1. Relação entre qualidade da pesquisa, índice de concordância e visibilidade, EMBRAPA - 1983.

a. Relação teórica



b. Relação empírica

Nível de qualidade	Índice de concordância (\bar{X})	Índice de visibilidade (\bar{X})
Superior (N = 8)	0,543	82,0
Médio (N = 16)	0,625	70,0
Inferior (N = 8)	0,617	64,8

Fonte: Dados da pesquisa

corrido um período médio de seis meses da primeira aplicação. O Índice de Concordância, por sua vez, é uma medida de fidedignidade do Índice de Qualidade, porque mede a consistência da classificação de cada unidade. Este índice já foi discutido acima. Além disso, o número relativamente grande de juízes (44) e pequeno de discriminações solicitadas (três), indicam que o Índice de Qualidade da pesquisa possui substancial fidedignidade.

Validade

A validade de uma medida diz respeito a saber se de fato estamos medindo o que nos propomos medir. A classificação das unidades pelos juízes produziu uma dispersão de 1,4 pontos entre a mínima e a máxima, de um total possível de 2 pontos, demonstrando que as classificações não foram indiscriminantes. A ordem de qualidade resultante da aplicação do Índice de Qualidade foi mostrada a alguns dos juízes, que concordaram ser ela um espelho da qualidade da pesquisa.

A validade lógica da medida foi considerada em item anterior, quando foram expostos os fundamentos teóricos e metodológicos para sua construção. Além disso, deve ser levada em consideração a validade do construto (Kerlinger, 1980), a qual pergunta que propriedades do objeto o instrumento está medindo.

Existem muitas alternativas de medir a validade do construto, através da determinação de que características estão ligadas ao desempenho da medida.

Tem sido fartamente demonstrado que a educação se correlaciona positivamente com diversas medidas de desempenho ocupacional, principalmente com salários, atividades de publicações e nível de cargo no emprego (Strauss & Tarr, 1982), as quais representam aspectos da qualidade da pesquisa ou estão com esta correlacionadas. Portanto, se o Índice de Qualidade da Pesquisa está medindo o que se propõe medir, deve existir correlação positiva entre o mesmo e alguma medida de educação nas unidades de pesquisa.

O Departamento de Recursos Humanos da EMBRAPA havia criado o Índice de Especialização para medir a densidade educacional (EMBRAPA, 1979). O índice foi calculado como a média ponderada (1 = BS, 2 = MS, e 3 = PhD) do nível educacional dos pesquisadores da unidade. As pesquisas anteriormente feitas levam a esperar que, quanto mais alta a densidade educacional da unidade de pesquisa, mais alta a qualidade da pesquisa. Usando do Índice de Especialização referente aos anos de 1977 e 1982, encontramos a correlação de $r = -0,579$ entre os Índices de Qualidade e de Especialização 1977 e de $r = -0,471$ entre aquele e o Índice de Especialização 1982. Os dois Índices de Especialização, são altamente (mas não perfeitamente) correlacionados entre si ($r = 0,725$).

A existência de correlação mais alta do Índice da Qualidade com o Índice de Especialização mais antigo (1977) do que com o mais recente (1982) é esperada devido ao retardamento entre o treinamento e o início do fluxo de benefícios dele decorrentes (Avila et alii, 1983).

Em síntese, concluímos em favor da validade do Índice de Qualidade, porque este discrimina de modo consistente, previsível, teórica e praticamente coerente, os objetivos que foram submetidos à mensuração.

Validade externa

A validade externa procura determinar até que ponto os resultados de uma pesquisa são transferíveis para situações e condições exteriores à mesma. Pela natureza do Índice de Qualidade da Pesquisa, o problema deve ser dividido em duas partes: a metodologia e os resultados substantivos.

Como metodologia, não há por que não aplicá-la em outros contextos que satisfaçam as condições de correta aplicabilidade. Seria viável, por exemplo, para medir a qualidade do produto das Empresas Estaduais de Pesquisa Agropecuária, ou para comparar o ensino oferecido por educandários em uma cidade. O êxito da transferibilidade depende, sobretudo, da existência de um grupo de especialistas ubicado em posição que lhes permita obter a visão necessária de todo o sistema a que esteja sendo aplicado. A comparabilidade dos resultados é direta, no que se refere às posições relativas dos objetos de cada grupo. Para que haja comparabilidade dos julgamentos absolutos é necessário, contudo, evidenciar que dois ou mais painéis de juízes usaram critérios de qualidade equivalentes.

Os resultados substantivos da pesquisa, por sua vez, referem-se a grande parte do Sistema Cooperativo de Pesquisa Agropecuária do Brasil. Ainda assim, não servem para responder diretamente às questões sobre a qualidade da pesquisa no Sistema como um todo. Aliás, esta não foi a intenção proposta.

A média do Índice de Qualidade poderá ser interpretada futuramente como um indicador da qualidade da pesquisa da EMBRAPA, se comparada através do tempo, ou com outras organizações, caso haja suficiente evidência de que os padrões de julgamento são equivalentes.

Em compensação, os resultados permitem responder o problema crucial para a administração da pesquisa, que é a variância de qualidade interna à organização. Neste contexto, não só o Índice de Qualidade, como também o de Concordância e o de Visibilidade, serão instrumentos relevantes para a continuação das pesquisas de recursos humanos e de organização. Interpretá-los no contexto macro-social é uma alternativa que está sendo explorada (Acosta Hoyos, 1984).

Diferentemente dos índices de status das ocupações que são invariáveis durante longos períodos e através de culturas diferentes, é de esperar que o Índice de Qualidade da Pesquisa apresente variações através do tempo. Assim, é conveniente que a mensuração seja repetida periodicamente (três a cinco anos) de modo a atualizar a percepção da realidade e estudar as modificações porventura observadas.

QUASE UM APÊNDICE: OS JUIZES

Os dados coletados para medir qualidade da pesquisa agropecuária podem ser

analisados em direção inversa à que usamos até aqui.

Ficou evidente que existe certa concordância nos resultados dos julgamentos, a qual pode ser atribuída à existência de uma base valorativa comum a todos os juízes. E as discrepâncias, a que podem ser atribuídas? Parte delas provavelmente é determinada pelas idiossincrasias pessoais. Parte, terá sido efeito de erro de mensuração e/ou de má informação, como acontece em qualquer pesquisa. Outra parte pode ser devida sistematicamente a características dos juízes ou a seu relacionamento com a organização. A análise que se segue procura evidência sobre a influência das características dos juízes em seus julgamentos.

Foram examinadas quatro características: a localização do juiz na estrutura organizacional da EMBRAPA, o nível educacional, a experiência internacional e o conhecimento do sistema.

Os juízes que trabalham na sede da Empresa são ligeiramente mais rigorosos em seus julgamentos do que os chefes de unidades. A média dos julgamentos dos primeiros foi de 1,89, isto é, um pouco superior ao conceito de qualidade "média", enquanto a dos chefes de centros foi de 1,69 e a dos chefes de UEPAEs, 1,68.

A relação entre rigor de julgamento e educação não se mostrou linear. Seria de esperar que, quanto mais alto o nível educacional, maior a identificação com os critérios da comunidade científica e, portanto, mais rigorosa a aplicação destes em forma de julgamento. Os juízes cujo nível educacional é o de doutorado são mais rigorosos (1,79) do que aqueles com mestrado (1,60). Os que só possuem o nível de bacharelato, porém, são os mais rigorosos de todos (2,02), quando se poderia esperar que fossem os mais lenientes. Uma possível explicação advém do acesso diferencial à pós-graduação que tiveram essas pessoas na EMBRAPA. Dos pesquisadores mais antigos, hoje ocupantes de posições elevadas na estrutura da organização, apenas aqueles que já haviam conquistado sólido renome e reconhecimento não foram enviados para fazer cursos de pós-graduação. Por causa dessas características como cientistas, é provável que eles adotem critérios de julgamento os mais rigorosos, o que se evidencia na mensuração da qualidade da pesquisa e, portanto, na diferença de médias que foi observada.

Os juízes com experiência internacional, isto é, aqueles que fizeram algum curso formal no exterior, quase não se distinguem em rigorismo de seus colegas que estudaram apenas no país (médias, respectivamente, de 1,79 e 1,73). A maior presença, entre estes, do grupo de juízes sem pós-graduação com certeza influencia fortemente a magnitude e a direção da diferença.

O conhecimento do sistema foi medido pelo número de julgamentos emitidos por cada juiz. Aqueles que se consideraram aptos a atribuir julgamentos a metade ou mais das unidades de pesquisa, foram considerados de conhecimento elevado. Os que conhecem menos o sistema julgaram com mais leniência a qualidade da pesquisa das unidades (1,68), enquanto os que mais o conhecem foram mais rigorosos (1,83). Este achado não é independente do fato de que as pessoas localizadas na sede conhecem o sistema melhor do que as demais.

A evidência sugere que há relação entre as características dos juízes e os julgamentos por eles emitidos, sendo a educação o maior correlato de tais diferenças. Todas elas, porém, são muito menores do que as diferenças identificadas pelos juízes entre as diversas unidades de pesquisa.

CONCLUSÃO

O presente trabalho é uma proposta de mensuração da qualidade da pesquisa agropecuária que se executa na EMBRAPA. Parte da premissa (testada em seguida) de que os critérios para julgamento de qualidade da pesquisa são compartilhados pelos pesquisadores e pelo pessoal de direção da Empresa. Esses critérios se baseiam, tanto no julgamento de utilidade que a pesquisa apresenta para a agricultura, como no dos aspectos acadêmicos da mesma.

O foco de mensuração escolhido é a unidade organizacional em que a Empresa está dividida, ou seja, o centro de pesquisa. O conceito de qualidade da pesquisa aproxima-se, por isso, do conceito de efetividade (desempenho, sucesso organizacional), tal qual se usa nos estudos de organizações complexas. Por outro lado, aproxima-se do conceito de status, como é usado em estudos de estratificação social mas, no presente caso, tendo como referencial a organização. Refere-se à imagem que o público especializado tem das unidades de pesquisa, a qual se baseia nas tecnologias e comunicações científicas produzidas por esta.

Foi usado um painel de juízes formado pelos chefes de 30 unidades de pesquisa agropecuária e por 14 pessoas que ocupam postos de direção e controle da pesquisa na sede da EMBRAPA. Suas opiniões provavelmente refletem corretamente a qualidade da pesquisa na EMBRAPA e seguramente influenciam o julgamento de outras pessoas sobre a mesma.

O Índice de Qualidade representa a média dos julgamentos feitos pelos juízes, oscilando teoricamente entre a qualidade 1 = superior e 3 = inferior. Os resultados da classificação das unidades de pesquisa foram revistos por diversas pessoas, tanto dentre as que serviram de juízes, como por outras conhecedoras da situação da pesquisa agropecuária, que os consideraram descrição adequada da realidade. A média do Índice de Qualidade (1,810) pode ser interpretada como a qualidade do sistema em geral, assim como é vista pelos juízes. Esta se localiza no âmbito da qualidade "média", tendendo, porém, para "superior".

Além do Índice de Qualidade, foi calculado o Índice de Concordância e o de Visibilidade. O primeiro é o desvio padrão dos julgamentos atribuídos pelos juízes a cada unidade de pesquisa e o segundo é a proporção de juízes que julgaram cada unidade. Os resultados mostram que há maior visibilidade das unidades onde se faz pesquisa de qualidade mais elevada, que das demais. A concordância é mais alta a respeito das melhores unidades, é menor a respeito das unidades médias, sendo intermediária a respeito das unidades que receberam julgamentos mais baixos.

A fidedignidade da medida de qualidade foi calculada usando o processo do

teste-reteste. Seis juízes classificaram novamente todas as unidades de pesquisa, com um intervalo médio de 6 meses entre as duas classificações. A correlação com a primeira classificação feita por eles ($r = 0,890$) sugere alto nível de fidedignidade para o Índice.

Foram examinadas diferentes formas de validade, e a evidência sugere concluir que o Índice de Qualidade está realmente medindo a qualidade da pesquisa. Levantadas hipóteses sobre sua correlação com índices que medem variáveis teoricamente relacionadas à qualidade da pesquisa, foram observadas as correlações na direção predita e nos níveis teoricamente esperados. Deve-se concluir que o Índice de Qualidade, além de ser fidedigno, discrimina de modo constante, previsível e coerente os objetos que foram submetidos à mensuração.

Por causa das características teóricas e sociométricas que o índice apresenta, é possível aplicá-lo a outras instâncias semelhantes onde se queira medir qualidade da pesquisa, usar os resultados da mensuração para estudar as diferenças e semelhanças que levam as unidades de pesquisa a apresentarem os níveis de qualidade por ele registrados e, até, transpor seus princípios lógicos e procedimentos teóricos para a mensuração de qualidade de unidades organizacionais que apresentem características similares às aquelas apresentadas pelas organizações de pesquisa.

Finalmente, foram estudadas as classificações feitas pelos juízes, de modo a evidenciar até que ponto as condições sócio-organizacionais em que eles estão inseridos influenciaram os seus julgamentos. Ficou evidente que há alguma relação entre tais fatores, sendo o nível educacional, provavelmente, o fator mais importante dessa relação que, contudo, não é linear.

Quanto às características técnicas, os índices podem ser melhorados de duas maneiras. Em primeiro lugar, pode-se inverter a escala, de modo que à melhor qualidade correspondam os escores mais elevados e vice-versa. Isto facilitará o uso dos mesmos, embora nada de essencial acrescente. Em segundo, é aconselhável usar escalas mais sensíveis, isto é, com maior possibilidade de discernimento, de modo a diminuir o erro da mensuração. Uma escala de 0 a 100 pontos, ou uma escala tipo Galileu (Woeffel & Fink, 1982), usando, neste caso, como termos de comparação, a qualidade da pesquisa da Unidade mais visível, seriam soluções apropriadas. Além disso, é aconselhável que os índices sejam correlacionados com indicadores mais tradicionais de qualidade, como seja, o número de publicações e de citações.

Em suma, a explicação dos julgamentos sobre a qualidade da pesquisa agropecuária produzida pelas Unidades da EMBRAPA certamente não será um mero exercício acadêmico de mensuração. Além de servir como elemento para fundamentar estudos futuros sobre recursos humanos, organização e qualidade de vida no trabalho, deverá suscitar interrogações que, esperamos, se irão aprofundar, sobre o sentido do que foi medido e principalmente sobre como melhorar sempre mais a qualidade do que está sendo medido e também das técnicas de mensuração.

O Índice de Qualidade procura captar o fenômeno, ao nível do impacto na pró-

pria organização, do que foi realizado pelas unidades. Este é o momento que mais interessa à organização, porque representa a parte do trabalho de pesquisa visível aos especialistas e é globalizante como percepção. Uma linha de prosseguimento da análise, será determinar como se gera esse impacto, e que relação ele pode ter com o que se passa dentro da unidade. Além disso, identificar de que modo se complementam as influências de fatores substantivos da pesquisa, com fatores circunstanciais, tais como a comunicação, a oportunidade da tecnologia, o produto pesquisado, são explicações que levarão ao alargamento do conhecimento e possuem grande potencial de aplicação para a administração da ciência e tecnologia. Estudos sobre o ciclo de produção da tecnologia e seu impacto, retardado ou imediato, no status das organizações de pesquisa são outra linha de potencial fecundidade. Um estudo criterioso sobre o transcurso de tempo entre a introdução de mudanças na organização e o impacto sobre o status desta, não só seria de utilidade para a administração, como certamente significaria um avanço considerável para a teoria das organizações e da estratificação social.

AGRADECIMENTOS

Uma versão preliminar deste trabalho foi submetida à apreciação dos integrantes do "painel de juízes", assim como a outros pesquisadores. Tantos são eles, que se tornou impossível a costumeira citação. Grande parte dos comentários generosamente enviados determinou modificações no texto, as quais foram incorporados a esta versão, ou se transformaram em sugestões para estudos complementares. O produto final é, porém, responsabilidade única dos autores, que agradecem a colaboração de todos os que tornaram possível este trabalho. Agradecemos ainda, especialmente a Suzana Maria Valle Lima pelo trabalho de computação.

REFERÊNCIAS

- ACOSTA HOYOS, L.E. *Análise da qualidade da pesquisa agropecuária através dos impactos na qualidade de vida*. Brasília, 1984. Projeto de Dissertação de Doutorado, Fundação Escola da Sociologia e Política de São Paulo.
- ALVES, E.R.A.; PASTORE, J.; PASTORE, A.C. *Coletânea de trabalhos sobre a EMBRAPA*. Brasília, EMBRAPA-DID, 1980.
- ANDERSON, J.H. Survey of recent agricultural research evaluation literature. In: WORKSHOP sobre metodologias de avaliação socioeconômica da pesquisa agropecuária. Brasília, EMBRAPA, 1983. v.3, p. 3-1 a 3-35.
- ÁVILA, A.F.D. et alii. *Formação do capital humano e retorno dos investimentos em treinamento na EMBRAPA*. Brasília, EMBRAPA-DEP/DRH, 1983.
- BLUMENSCHNEIN, A. *Princípios da pesquisa no sistema EMBRAPA*. Brasília, EMBRAPA, 1978.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA. Departamento de Recursos Humanos. **Relatório de atividades, 1978**. Brasília, 1979.

EVENSON, R.E. & WRIGHT, B.D. An evaluation of methods examining the quality of agricultural research. In: WORKSHOP sobre metodologias de avaliação socioeconômica da pesquisa agropecuária. Brasília, EMBRAPA, 1983. v.3, p. 12-1 a 12-73.

GASTAL, E. **Enfoque de sistemas de programação da pesquisa agropecuária**. Rio de Janeiro, IICA, 1980.

GOODE, W.J. & HATT, P.K. **Métodos em pesquisa social**. São Paulo, Nacional, 1964.

KERLINGER, F.N. **Metodologia da pesquisa em Ciências Sociais**. São Paulo, E.P.V., 1980.

NORTON, G.W. & DAVIS, J.S. Evaluation returns to agricultural research: a review. *Am. J. agric. Econ.*, 63(4):685-99, Nov. 1981.

PELZ, D.C. & ANDREWS, F.M. **Scientists in organizations**. Ann Arbor, Mi., Institute for Social Research, University of Michigan, 1976.

PRICE, J.L. **Handbook of organizational measurement**. Lexington, Mass, Heath, 1972.

QUIRINO, T.R. **A socialização ocupacional do pesquisador agropecuário**. Brasília, EMBRAPA-DID, 1981.

STRAUSS, R.P. & TARR, W.J. Salary patterns of agricultural economists in the early 1980. *Am. J. agric. Econ.*, 64(4):1053-61, 1982.

WOELFEL, J. & FINK, E.L. **The measurement of communication processes: Galileo theory and method**. New York, Academic Press, 1980.

WRIGHT, B.D. Evaluation of research quality. In: WORKSHOP sobre metodologias de avaliação socioeconômica da pesquisa agropecuária. Brasília, EMBRAPA, 1983. v.3, p. 13-1 a 13-20.

ANEXO - Qualidade da Pesquisa Agropecuária na EMBRAPA - 1983, determinada por um painel de 44 juízes.

Unidade	QUALIDADE			CONCORDANCIA			VISIBILIDADE		
	Quartil	Ordem	Índice	Quartil	Ordem	Índice	Quartil	Ordem	Índice
I	1º	1	1,129	1º	1	0,341	2º	15	72,1
II	1º	2	1,162	1º	2	0,442	1º	3	86,0
III	1º	3	1,264	1º	3	0,448	2º	9	79,1
IV	1º	4	1,389	1º	7	0,549	1º	5	83,7
V	1º	5	1,417	3º	23	0,649	1º	5	83,7
VI	1º	6	1,472	4º	29	0,696	1º	5	83,7
VII	1º	7	1,485	1º	8	0,566	2º	11,5	76,7
VIII	1º	8	1,513	4º	26	0,683	1º	1	90,7
IX	2º	9	1,607	2º	9	0,567	3º	22	65,1
X	2º	10	1,636	3º	17	0,603	2º	11,5	76,7
XI	2º	11	1,643	3º	19	0,621	3º	22	65,1
XII	2º	12	1,656	2º	15,5	0,602	2º	13,5	74,4
XIII	2º	13	1,676	3º	20	0,638	2º	9	79,1
XIV	2º	14	1,714	2º	14	0,600	3º	22	65,1
XV	2º	15	1,733	2º	13	0,583	3º	16,5	69,8
XVI	2º	16	1,739	4º	27	0,689	4º	32	53,5
XVII	3º	17	1,785	2º	10,5	0,568	3º	22	65,1
XVIII	3º	18	1,821	1º	6	0,548	3º	22	65,1
XIX	3º	19	1,867	4º	30	0,730	3º	16,5	69,8
XX	3º	20	1,884	3º	24	0,653	4º	29,5	60,5
XXI	3º	21	1,894	4º	32	0,764	1º	2	88,4
XXII	3º	22	1,971	2º	10,5	0,568	1º	7	81,4
XXIII	3º	23	2,031	1º	5	0,538	2º	13,5	74,4
XXIV	3º/4º	24,5	2,036	4º	31	0,744	3º	22	65,1
XXV	3º/4º	24,5	2,036	4º	28	0,693	3º	22	65,1
XXVI	4º	26	2,148	2º	15,5	0,602	4º	27	62,8
XXVII	4º	27	2,241	1º	4	0,511	3º	18	67,4
XXVIII	4º	28	2,294	4º	25	0,676	2º	9	79,1
XXIX	4º	29	2,296	3º	18	0,609	4º	27	62,8
XXX	4º	30	2,400	3º	21	0,645	4º	31	58,1
XXXI	4º	31	2,423	3º	22	0,643	4º	29,5	60,5
XXXII	4º	32	2,556	2º	12	0,577	4º	27	62,8
MÉDIA DOS ÍNDICES	-	-	1,810	-	-	0,604	-	-	71,7

* O índice de Qualidade corresponde à média das classificações dadas pelos Juízes que usaram a seguinte escala: 1 - Superior, 2 - Média, 3 - Inferior. O Índice de Concordância corresponde ao desvio padrão do Índice de Qualidade. O Índice de Visibilidade corresponde à porcentagem de Juízes que se dispuseram a julgar cada unidade. A coluna "ordem" foi obtida pelo ordenamento decrescente do atributo a que se refere. A coluna "quartil" mostra a posição relativa da Unidade na distribuição do atributo a que se refere.