siderably more difficult, and some species may never be amenable to such breeding schemes.

This general breeding plan opens up new opportunities for applied geneticists. It can be modified in imaginative and useful ways to contribute to world food resources by means of the genetic preservation and improvement of economically important wild stocks. The improvement of breeding stocks for catch fisheries is an important and as yet unfulfilled role of aquaculture. We believe that the time lag to successful application of the plan with many commercially important aquatic and marine species is primarily a function of investment in problem-oriented research and development.

References and Notes

1. S. H. Wittwer, *Science* **199**, 375 (1978).
2. G. R. Spangler, N. R. Payne, J. E. Thorpe, J. M. Byrne, H. A. Regier, W. J. Christie, *J. Fish. Res. Board Can.* **34**, 1983 (1977).
3. J. J. Gwahaba, *East Afr. Wildl. J.* **11**, 317 (1973). *Tilapia nilotica* is one of the commercially important fish species of the African genus *Tilapia*.
4. R. P. Silliman [*U.S. Fish Wildl. Serv. Fish. Bull.* **73**, 495 (1975)] showed the same results as Gwahaba in a controlled laboratory experiment with another widespread commercial species, *Tilapia mossambica*.
5. J. Turner, *J. Fish. Res. Board Can.* **34**, 232 (1977).
6. "Proceedings of a Conference on the Effects of Exploitation, Eutrophication and Introduction on Salmonid Communities," *J. Fish. Res. Board Can.* **29** (No. 2) (1972); "Proceedings of the 1976 Percid International Symposium," *ibid.* **34** (No. 10) (1977).
7. R. Moav, T. Brody, G. Wohlfarth, G. Hulata, in *Proceedings of the FAO Technical Conference on Aquaculture* (Kyoto, Japan, 1976).
8. R. Vincent, *Trans. Am. Fish. Soc.* **89** (No. 1), 35 (1960); S. Smith, *Can. Fish Cult.* **20**, 7 (1957).
9. W. Flick and D. Webster, *J. Fish. Res. Board Can.* **33**, 1525 (1976).
10. M. J. Whitten and G. G. Foster, *Annu. Rev. Entomol.* **20**, 461 (1975). An identical approach may be used to eradicate or reduce pest fish populations, and it has been initiated in the fight against the common carp in Australia.
11. L. D. Donaldson, in *Marine Aquaculture*, W. J. McNeil, Ed. (Oregon State Univ. Press, Corvalis, 1970); pp. 65–74; J. R. Calaprice, *J. Fish. Res. Board Can.* **33**, 1068 (1976).
12. J. E. Bardach *et al.*, *Aquaculture: The Farming and Husbandry of Freshwater and Marine Organisms* (Wiley, New York, 1972).
13. The carp was chosen as an example because (i) more is known about genetic control of the growth of carp than of any other fish species and (ii) the carp has been the subject of our intensive genetic research for 20 years, and for the last 7 years we have been studying genetic differences between the cultivated European and Chinese big-belly races of the carp in relation to their evolution and use in breeding programs. These studies provide the only documented empirical evidence that we know of concerning the present subject of fish domestication. The cultivated European carp has been selected by European fish breeders for hundreds of years and it may be considered a truly domesticated breed. The Chinese big-belly carp has been grown in ponds even longer than its European relative, but it has not been selected for improved economic value and it continuously absorbed immigrants caught in rivers and stocked in ponds for cultivation. Also, traditional reproduction and harvesting techniques used with the big-belly carp, as a rule, did not differ in major features between farms and rivers. Consequently, we feel justified in considering the big-belly a wild organism (*16*).
14. R. Moav and G. Wohlfarth, *Genetics* **82**, 83 (1976); R. Moav, M. Soller, G. Hulata, G. Wohlfarth, *Theor. Appl. Genet.* **47**, 285 (1976).
15. G. Wohlfarth, R. Moav, G. Hulata, A. Beiles, *Aquaculture* **5**, 375 (1975).
16. G. Wohlfarth, R. Moav, G. Hulata, *Heredity* **34**, 341 (1975).
17. W. Steffens, *Biol. Zentralbl.* **93**, 129 (1974).
18. The curvilinearity of the additive component may result from its nonlinear relation to traits such as growth rate, viability, and seine catchability [see R. Moav, *Anim. Prod.* **8**, 193 (1966)].
19. Here, polyculture implies the mixing together in one pond of several species of fish. For full details of this experiment, see R. Moav, G. Wohlfarth, G. L. Schroeder, G. Hulata, H. Barash [*Aquaculture* **10**, 25 (1977)].
20. When comparing the results presented in Fig. 4A with the equivalent hypothetical example in Fig. 3B, the following points should be borne to mind. (i) The $BC_{w_1} \times BC_{w_1}$ group of Fig. 4A had lost half of the heterosis of its parental generation BC_{w_1}. (ii) Weight gain is only one of the major components of overall economic value. A second major component is viability, which would favor W over D. (iii) The correction for differences in initial weights in Fig. 4A eliminated a major effect of competition that would have favored the W side of the diagram. For a detailed discussion of the last point, see G. Wohlfarth and R. Moav [*Aquaculture* **1**, 7 (1972)].
21. G. Hulata, R. Moav, G. Wohlfarth, *J. Fish Biol.* **9**, 499 (1976).
22. For an example of the application of electrophoretic markers in studying hybridization under uncontrolled circumstances, see B. May, F. M. Utter, and F. W. Allendorf [*J. Hered.* **66**, 227 (1975)]. For a detailed discussion of the relative advantages of electrophoretic and morphological markers, see R. Moav, T. Brody, G. Wohlfarth, and G. Hulata [*Aquaculture* **9**, 217 (1975)].
23. The two genotypes with three W alleles may be found in both BC_{W_1} and F_2 at expected intragroup frequencies of 0.5 and 0.25, respectively (Table 2), so the frequency of BC_{W_1} may also be estimated.

$$(BC_{W_1}) = 2[A_wA_wB_wB_d + A_wA_dB_wB_w - {}^1/_4(F_2)]$$

Similarly

$$(BC_{D_1}) = 2[A_dA_dB_dB_w + A_wA_dB_dB_d - {}^1/_4(F_2)]$$

$$(F_1) = A_wA_dB_wB_d - {}^1/_4(BC_{W_1}) - {}^1/_4(BC_{D_1})$$

$$(W) = A_wA_wB_wB_w - {}^1/_4(BC_{W_1}) - {}^1/_{16}(F_2)$$

$$(D) = A_dA_dB_dB_d - {}^1/_4(BC_{D_1}) - {}^1/_{16}(F_2)$$

Note that the equations above provide estimates of the proportions of these groups among the caught fish. To calculate the equivalent proportions in the water, these estimates have to be multiplied by the appropriate relative catchability coefficients.
24. R. Suzuki, *Gyoruigaku Zasshi (Jpn. J. Ichthyol.)* **20**, 235 (1973).

25. We acknowledge the help of G. Wohlfarth, J. Bardach, and S. Malecha in the preparation of this manuscript. This research was supported by a grant from the United States–Israel Binational Science Foundation, Jerusalem, Israel. Contribution from the Agricultural Research Organization, Volcani Center, Bet Dagan, Israel: 1977 series, No. 204-E.

Research Involving Human Subjects

The performance of institutional review boards is assessed in this empirical study.

Bradford H. Gray, Robert A. Cooke, Arnold S. Tannenbaum

Research involving human subjects raises ethical and legal issues of sufficiently serious and widespread concern that an increasingly comprehensive mechanism has been developed through which the judgments of researchers are reviewed. Institutions seeking funds under the Public Health Service Act for research involving human subjects are required under the National Research Act of 1974 to establish committees (called "institutional review boards" in the Act and commonly referred to as IRB's) to review such research conducted at or sponsored by the institution. However, such committees existed at most institutions prior to this statutory requirement, because of Public Health Service and Department of Health, Education, and Welfare (HEW) requirements dating back to 1966, and many institutions had review committees even earlier (*1*, *2*). Under HEW regulations (*3*), IRB's are supposed to review research proposals to determine whether subjects will be placed at risk and, if so, whether the risks to the subjects are outweighed by the sum of the benefits to subjects and the importance of the knowledge sought, whether the rights and welfare of subjects are protected, and whether "legally effective informed consent" will be obtained by adequate and appropriate means. Institutional review boards also

Dr. Gray is professional associate at the Institute of Medicine, National Academy of Sciences, Washington, D.C. 20418. During the conduct of this study, Dr. Gray was on the staff of the National Commission for the Protection of Human Subjects. Dr. Cooke is associate research scientist and Dr. Tannenbaum is program director in the Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor 48106.

0036/8075/78/0922-1094$01.75/0

bear a largely undefined responsibility for "continuing review" of the conduct of research. An IRB is to be composed of at least five individuals of varying backgrounds and must include individuals who are able "to ascertain the acceptability of proposals in terms of institutional commitments and regulations, applicable law, standards of professional conduct and practice, and community attitudes" (*4*).

Summary. This article reports a study of the activities and performance of institutional review boards to protect human research subjects. Researchers and institutional review board members were generally supportive of the review system, although substantial criticisms were also heard. Institutional review boards had some direct impact on half of the proposals reviewed by requiring either modification of or additional information about proposed research. The data, however, raise questions about the effectiveness of some review board actions, for example, with regard to informed consent. Some policy implications of the study are presented.

Institutional review boards have been regarded as an important social invention because of interest in the value conflicts that they are intended to mediate, because of growing concerns about the regulation of powerful professions, and because they are seen as a decentralized regulatory model that may be adaptable to other situations. Yet, despite the complexity of these committees' tasks and the importance of both their effectiveness and credibility, they existed as a federal requirement for almost 10 years without being subjected to any serious evaluation by the government, although the rules and guidelines under which they operated were changed several times. During that time, hundreds of millions of dollars worth of research affecting tens of thousands of subjects passed through these review committees.

The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research was created by the National Research Act and directed to make recommendations regarding IRB's. Only superficial information about IRB's was available in HEW files, and the existing literature provided only a limited picture of review committees and the research they approve. In the only previous survey of review committees, conducted in 1969 by Barber *et al.* (*1*), data were limited to biomedical institutions and to the responses of a single official at each institution in the sample. The remainder of the existing literature on human subjects review committees was based upon studies of particular review committees (*5*, *6*). These studies, however, provided some reason for concern (*7*). Because of the key role of IRB's in the existing system of protecting human subjects and their prominent place in the commission's mandate and because existing studies were either dated, of limited depth, or based on the experience of a single committee, the commission contracted with the University of Michigan for the conduct of the study summarized in this article.

The Data on IRB's

The study described herein focused on review procedures and research projects at a probability sample of 61 institutions drawn from the more than 420 institutions with review committees approved by HEW (*8*). It covered research reviewed by these committees between 1 July 1974 and 30 June 1975. Interviews were conducted between December 1975 and July 1976 with more than 2000 research investigators whose proposals had been reviewed, over 800 IRB members or persons especially knowledgeable about the boards in the sample, and almost 1000 subjects or third parties (for example, parents) who consented on subjects' behalf.

Institutional review boards operate in a wide variety of institutional environments and differ widely from one another in their work loads and in the types of research they review. Medical schools (including those that share IRB's with universities) accounted for 59 percent of the research reviewed by IRB's in the sample. Universities (with IRB's separate from those for medical schools) and hospitals accounted for 18 percent and 15 percent, respectively. Most of the remaining research was conducted in "institutions for the mentally infirm" (*9*), although some was conducted in research institutions or in dental or nursing schools.

Approximately 60 percent of the studies reviewed by IRB's were primarily biomedical, most frequently involving the administration of drugs, chemical agents, or blood products (28 percent of all research by IRB's), or the study of samples of bodily fluids or tissues (25 percent). Investigators in many of these studies reported that the major intervention (for example, the drug administration) would have occurred even if the patient had not been involved in the study. Only 1 percent of the projects reviewed by IRB's involved surgery, perhaps reflecting the difficulty of defining "research" in the surgical context. Behavioral research—most frequently based on interviews, questionnaires, testing, or observation—accounted for about one-third of the research reviewed by IRB's; about a fifth of the behavioral research entailed the study of an intervention such as social or psychological therapy, behavior modification techniques, or educational innovations. The remaining small fraction of the research reviewed by IRB's (about 6 percent) involved secondary analysis of data or the study of bodily fluids or tissues that had been obtained for other purposes.

An average IRB reviewed 43 proposals per year. However, a review board at a small institution may not receive even a single proposal in a given year, while IRB's in major medical schools or universities review hundreds of proposals. The number of members on IRB's in the sample ranged from 5 to 55, with an average of 14. Some boards in the sample met as few as two times and some met as many as 51 times per year, with an average of almost 10 meetings per year. The average IRB expended 760 member-hours per year on board work, and this figure ranged as high as 5000 member-hours for one board. Institutional review boards met for almost 1 hour per proposal (*10*), and the total number of member-hours per proposal (both in meetings and outside of meetings) averaged 38 hours but ranged up to 270 hours at one review board.

Composition of IRB's

The majority of members of IRB's in the sample were biomedical scientists (50 percent) or behavioral scientists (21 percent); the remainder included administrators, lawyers, nurses, members of the clergy, and others. Approximately 90 percent of the IRB's included biomedical researchers, behavioral researchers, full-time administrators, and "community representatives." About three-fourths of the IRB's included a lawyer; this was most common in medical school IRB's, while lawyers were present on fewer than one-third of the IRB's at institutions for the mentally infirm. At least one member of all IRB's was not otherwise affiliated with the institution. Half of the

Table 1. Review board action on research proposals.

Type of institution	N	Percentage of projects					
		Modified by board	Modified informally*	For which board sought more information†	For which modification or information not required	For which no data available	Total
Universities	514	29	7	16	46	2	100
Medical schools	1425	32	8	9	46	5	100
Hospitals	254	42	4	10	37	7	100
Institutions for mentally infirm	101	33	6	11	43	7	100
Other	95	28	4	3	52	13	100
All	2389	33	7	10	44	6	100

*Includes projects not modified by formal board action but which were modified as a result of informal discussions with IRB members. †Includes only projects not modified by either formal actions or informal discussion but for which the IRB requested information beyond that originally provided by the investigator.

IRB's included racial or ethnic minorities, and 88 percent included women. Fewer than 5 percent of IRB members said that they had any special training for their role (for example, by attending seminars or workshops), although most said they had received a briefing or some written instructions (for example, the HEW regulations).

Members of IRB's who were not behavioral or biomedical researchers generally reported themselves to be less active and less influential than other IRB members. Nevertheless, almost all members indicated that viewpoints of all members were sought and considered in IRB decisions, and almost 90 percent of members expressed satisfaction with their accomplishments on the IRB.

Policies and Procedures of IRB's

While there are a few common denominators among IRB's—for example, almost all boards discuss proposals in convened meetings, and most institutions do not confine the review requirement to funded research—the diversity of their policies and procedures is much more striking than the similarities. About two-thirds of the IRB's had a procedure to screen out proposals that did not need review. About half of the IRB's assigned proposals to individual members for intensive review, and about one-fourth reported delegating some responsibility to subcommittees for intensive review. About half of the IRB's took formal votes on all proposals, and almost all took formal votes on at least some occasions. Two-thirds of the IRB's accepted majority approval as satisfactory; one-fourth required unanimity. More than one-fourth said that investigators always attended the meetings at which their proposals were discussed, and more than 80 percent reported that this happened at least occasionally. Half of the IRB's had provisions for investigators to appeal board decisions.

Modifications of Research Proposals

Data from the survey of 300 biomedical institutions by Barber *et al.* (*2*, p. 160) indicated that in 1969 most IRB's were not very active in the modification of proposed research. For example, 34 percent of the IRB's had never modified or rejected a project, and only 16 percent had required revision of more than 10 percent of the projects reviewed. These low rates could have been due to (i) high-quality proposals that needed no changes or (ii) the poor performance of IRB's. On the basis of an intensive study of a single, very active review board, one of us (B.H.G.) argued that the latter explanation was more plausible (*7*).

Although data from the present study are not precisely comparable with the Barber data because of differences in sampling procedure, item wording, and categories used in the analyses, there is little doubt that IRB's have become more active, at least as such activity is reflected in their requiring investigators to modify proposed research. Information provided by research investigators indicates that 40 percent of the proposals reviewed by IRB's were modified (Table 1). Most of these modifications occurred in the formal review process, although some projects were modified as a result of informal contacts between investigators and IRB members. Modifications regarding informed consent were required in one-fourth of the proposals, while modifications regarding scientific design, subject selection, risks and discomforts, and confidentiality were each made in a small number of proposals (Table 2). More information was sought by IRB's on about almost one-third of the proposals submitted for review; in about two-thirds of these proposals, the IRB's also made a substantive request for modification.

In the percentage of proposals that they modified, IRB's varied markedly. Modifications in every proposal reviewed were reported for 14 percent of the boards, while 22 percent of the boards modified no more than one-third of the proposals reviewed. Boards also differed in the variety of modifications they made. Nineteen percent made only one type of substantive modification, while 7 percent made all six types of modification identified in the survey (that is, regarding consent, risks, scientific design, subject selection, confidentiality, and "other" modifications). The median number of types of modifications by IRB's was 2.5.

Risks and Benefits of Approved Research

More than half (55 percent) of the projects for which information was available were expected by the investigators to be of benefit to the research subjects (*11*). There was little difference between biomedical research and research that involved a behavioral intervention in the frequency with which benefit to subjects was expected, although the nature (medical or psychological) of the expected benefits differed. Fewer than one-third of the behavioral projects that did not involve the study of an intervention were expected to benefit subjects.

Estimates of the probability and magnitude of the possible harms to subjects were also provided by investigators. One-fourth of the investigators judged their projects to be without risk, and another fourth judged their research to have no more than a "very-low" probability of "minor" medical or psychological complications. The remaining half of the research involved at least a "low" probability of minor complications or a "very-low" probability of "serious" medical or psychological complications.

These estimates of risk should not be treated as objective assessments of the degree of risk present in research. The assessment of independent raters would undoubtedly differ in some cases from the assessment of investigators themselves. That is, after all, one of the rationales for the review process. Nevertheless, the validity in the aggregate of the investigators' estimates of the riskiness of their research receives some confirmation from the fact that injuries to subjects were more likely to be reported

Table 2. Actions formally required of the investigator by the review board. The data are expressed as percentages; however, percentages need not add to 100 since respondents might indicate fewer or more than one action required. The N's vary slightly within columns because of missing data. The percentages exclude missing data.

Type of request	Type of institution					
	Universities (N = 514)	Medical schools (N = 1425)	Hospitals (N = 254)	Institutions for mentally infirm- (N = 101)	Other (N = 95)	All (N = 2389)
More information	33	30	39	28	21	32
Modification in consent forms and procedures	19	25	31	14	13	24
Modification in scientific design	*	2	6	8	1	3
Modification in subject selection	*	3	5	7	1	3
Modification regarding risks and discomfort	3	4	4	7	9	4
Modification regarding confidentiality	6	2	3	6	6	3
Other modifications	5	3	7	7	9	5
Informal suggestion for modifications†	13	15	13	19	15	15

*Less than 0.5 percent but greater than zero. †Row includes all projects modified as a result of informal discussions with IRB members, whether or not the review board formally requested modifications.

in studies in which risks were assessed as relatively high (see Table 3). Table 3 shows that as the risk of projects increased, so did the likelihood that the project would benefit the subjects. Only about one-third of the "no risk" projects were expected to benefit subjects, while at the other end of the risk scale 80 percent of the projects were expected by the investigators to benefit subjects (*12*).

Overall, investigators reported the occurrence of actual harm to subjects in 3 percent of the projects (*13*). These harms were generally reported as trivial or only temporarily disabling. Three investigators, however, reported fatal effects; in each of two projects one subject died and in one project three subjects died. Each of these projects involved cancer research, and in two of the projects some subjects were in near terminal condition at the time of their participation in the research.

There were indications that IRB's that review relatively risky research are more careful in their reviews. For example, among the medical school and hospital IRB's, those that reviewed more relatively high-risk research reported discussing a more comprehensive set of issues during the review of proposals, and the rate of modification in proposals was greater than in boards that reviewed more lower-risk research. This correspondence between risk and IRB activity was not found in universities, perhaps because there is less variation in risk in the research reviewed therein.

Selection of Subjects in Approved Projects

By and large, IRB's accepted investigators' plans for selection of subjects. However, changes were required in 3 percent of the projects, usually by limiting or restricting the sample in some way. "Patients" served as subjects in 76 percent of the projects approved by IRB's in medical schools and in 86 percent of projects in hospitals. In almost half of these projects, the subjects were the investigator's own patients. Patients were subjects in only 17 percent of the projects in universities, where projects most frequently involved college students (37 percent) as subjects. Subjects in most research were selected because of some specific condition or characteristic. For patients this usually meant that their disease was a selection criterion; in research in universities, the most common selection criteria were demographic characteristics such as age or educational situation. Persons identified as "patients" served as subjects in three-fourths of the projects expected to benefit subjects and in half of the other studies.

Projects in which investigators reported relatively high proportions of (i) males, (ii) persons between 41 and 64 years of age, and (iii) high or middle income persons were more likely than other projects to be above average in risk. Overall, although more investigators described their subjects as "low income" persons than as "high income" persons, there was no evidence that low income persons were particularly likely to be selected either for relatively risky research or for research that was not expected to benefit subjects. Projects involving substantial proportions of children or older people were more likely to be expected to benefit the subjects than were projects that drew more heavily on 19 to 40 year olds.

Informed Consent

Informed consent is the focus of considerable activity by IRB's; yet it remains a problem. In one-fourth of the projects, investigators reported that the IRB had required that they make changes in their plans for obtaining informed consent. Virtually all of these changes pertained to the content of consent forms—most commonly through the addition of materials—rather than to the way in which consent was obtained; in fewer than 1 percent of the studies did IRB's require changes regarding the timing of the consent process, who obtained consent, the setting in which consent

Table 3. Risk, benefit, and availability of treatment for harmful effects. Data are expressed as percentage of projects.

Relative risk level*	N	Expected to benefit subjects	Harmful effects reported	Treatment reported available
No risk	710	34	0	14†
Very low risk	446	52	1	31
Low risk	459	63	3	52
Moderate risk	483	80	12	81

*Medical and psychological risks, as assessed by investigators. The risk index was constructed by arraying the studies in which risk was reported ("no risk" studies thus represent a distinct category), using weighting procedures that reflect both the seriousness and probability of harm. The resulting arrayal was then, to the degree possible, trichotomized. Studies are thus categorized according to their risk in relation to other studies, not according to an independent standard of "very low risk," "low risk," and "moderate risk." †In some cases, investigators probably reported institutional policies regarding treatment of injuries, rather than provisions made for a specific study; hence, provisions for treatment for harm were reported to some studies that the investigators indicated were free of risk.

Table 4. Readability of consent forms approved by IRB's. The table is based on those projects for which a consent form was available. Percentages exclude missing data. Readability is measured by the Flesch Readability Yardstick. The "reading-ease score" for a selected passage is based on word length (the average number of syllables per 100 words) and sentence length (the average number of the words per sentence) [see (*16*)]. The classification used in table (for example, scholarly/academic, pulp fiction) were delineated by Flesch in 1948 (*16*).

Type of institution	*N*	Very difficult (scientific/ professional)	Difficult (scholarly/ academic)	Fairly difficult (*Atlantic Monthly*)	Standard (*Time*)	Fairly easy (slick fiction)	Easy (pulp fiction)	Very easy (comics)	Total
Universities	219	18	49	20	10	3	*	0	100
Medical schools	1011	21	56	19	4	*	*	0	100
Hospitals	159	22	61	8	8	*	0	0	100
Other	137	24	55	16	2	3	0	0	100
All	1526	21	56	17	5	1	*	0	100

*Less than 0.5 percent but greater than zero.

would be obtained, or the presence of a witness (*14*).

The obtaining of consent. Investigators reported that informed consent was obtained in almost 90 percent of the projects. Usually such consent was obtained in writing. The major reasons cited for not obtaining consent were that the return of questionnaires implied consent, that only routine procedures or treatments were being used, or that the study was based exclusively upon existing records, data, or materials gathered for other purposes (*15*).

In 15 percent of the studies, investigators reported that some information was withheld from subjects. This occurred most frequently in studies conducted in universities and least frequently in projects conducted in medical schools, and as often in projects expected to benefit subjects as in other studies. The reason given for withholding information was usually to eliminate sources of bias in the study or because it was believed that the subject would not understand the information. The information not disclosed usually pertained to the purpose of specific procedures in the study or to the identity of the medication or treatment being used with particular subjects (as in double-blind research designs). In a few projects (2 percent) investigators reported that subjects were given information that was untrue. Most of these projects were conducted in universities. The false information usually concerned the purpose of the procedures used in the study, and the reasons again pertained to the avoidance of bias in the data.

Consent forms. Despite the general use of consent forms and evidence of the concern of IRB's regarding such forms, consent forms tended to be incomplete and difficult to understand, according to our analysis of the content and readability of the actual forms used in the research. On an index composed of six consent elements specified in the HEW regulations—the purpose of the research, the procedures involved, the risks, the benefits, a statement that subjects are free to withdraw from the research, and an invitation to ask questions—only 18 percent of the forms were complete or nearly complete. Twenty-one percent of the forms from hospitals and medical schools were complete or nearly so, while this was true of less than 10 percent of the forms from universities and other institutions. Descriptions by investigators of the topics covered in oral explanations added only negligibly to the report of information that was transmitted to subjects.

Some elements received more coverage than others in consent forms. The procedures of the research were not mentioned in 10 percent of the forms; the purpose was not mentioned in 23 percent; and the benefits of the research (or absence of benefits to subjects) were not mentioned in 45 percent. Risk was not mentioned in 30 percent of the forms, and 70 percent of these cases were in studies which were described by investigators as entailing at least a very low probability of minor harm to subjects. Even in consent forms in which these various elements were mentioned, fewer than half of the forms provided a detailed description. In some cases, these topics were mentioned only in statements saying "I certify that I have been informed of the purpose, procedures, and risks and benefits of this study." A statement regarding withdrawal from the study was not present in 22 percent of the consent forms; however, many of these may have been from studies in which the active participation of subjects ended quickly. An offer to answer questions appeared in more than half of the consent forms. A description of alternative treatments might have been expected in studies which were expected to be of benefit to subjects; however, this occurred in fewer than 20 percent of the forms in such studies. Similarly, consent forms from projects described by investigators as including an "experimental" element might have been expected to mention this. About 60 percent of the forms from such projects, however, did not call attention to the experimental nature of the project through the use of words such as "experiment," "research," or "investigation."

A "reading ease score" was computed for each consent form according to a standard measure, the Flesch Readability Yardstick (*16*). Consent forms tended to be written in academic or scientific language that may be difficult for the lay person to understand (see Table 4). Descriptions of the procedures used in the research tended to be somewhat more readable than descriptions of the purpose or risks of the research; but overall, fewer than 7 percent of the consent forms were in language as simple as is found, for example, in *Time* magazine. In more than three-fourths of the consent forms, fewer than 10 percent of the technical or medical terms were explained in lay language. It is questionable whether many subjects would find most consent forms useful to them in making decisions regarding participation in research. No information is available on the degree to which the difficult language of the consent forms is mitigated by oral explanations in simpler terms.

There was no tendency for the more complete consent forms to be either more or less difficult to read than were the less complete consent forms. The correlation between the two measures was low and negative ($r = -.09$).

Comparisons were made of the pre- and postreview versions of consent forms from the same projects in an attempt to elucidate why IRB's required many modifications in consent forms, yet approved forms that were frequently incomplete and difficult to read. No significant difference was found on the average readability or completeness scores between consent forms as submitted to the IRB and the consent forms as approved by the IRB. Confining attention to consent forms which were reported to have been changed at the request of the

IRB's, the average difference in completeness between the two versions of consent forms was 0.076 (on a scale that could range from −3 to +3), and the average difference in readability was 0.007 (on a scale that could range from −6 to +6). Thus, as readability and completeness were measured, no significant improvement could be found in consent forms that the IRB had required to be modified.

Furthermore, an examination of consent forms submitted for review showed there to be no significant differences (in the expected direction) in readability or completeness between forms which produced an IRB request for modification and those for which the IRB made no such request. Confining attention to consent forms submitted for review (that is, *not* to consent forms approved by IRB's), the mean completeness score of forms which the IRB requested to be changed was 1.84 and the mean completeness score for forms for which no changes were requested, 1.60 (completeness scores here ranged from 0 for incomplete to 3 for complete). On these same forms submitted for review the mean readability score on forms that the IRB requested be changed was 2.09, while the readability score for those for which no changes were requested was 2.11 (readability scores here ranged from 1 for very difficult to read to 7 for very easy to read). Thus, the less readable and less complete consent forms were no more likely than the relatively readable and complete forms to have elicited a request from the IRB for modification.

Involvement of IRB's After Initial Review

Perhaps the most common criticism of IRB's has pertained to their lack of involvement with research after the initial approval. The study showed that most IRB's approve at least some projects with the stipulation that they be reviewed again after intervals ranging from 1 month to 3 years, but usually after 1 year. Only half of the boards reported having either a formal or informal policy regarding the reporting of injuries to subjects. In most instances, investigators were supposed to notify the IRB in the event of injuries to subjects; a few IRB's reported that a study would be halted or reviewed again if injuries occurred. Somewhat surprisingly, reports from more than one-third of the IRB's indicated that they had at some time designated someone to observe the manner in which research was conducted; half of these boards said that this was done routinely, and the others reported that projects were observed only under certain circumstances, such as when there was particular risk, when children were involved, or when there had been problems in the past.

The Performance of IRB's

In examining the performance of IRB's, we looked at differences among them in the extent to which each (i) is comprehensive in its discussions of proposals (a score constructed by combining the responses of members of each IRB to a list of topics discussed by their IRB), (ii) has procedures to monitor the progress of research, (iii) makes modifications in proposals, (iv) approves readable and complete consent forms, (v) is judged by IRB members to do a good job, and (vi) is viewed positively by investigators.

Although a high score on any particular measure may not indicate an effective board, an IRB that scores high on all of these aspects could presumably be judged to be an effective board, and one that scores low on all of these aspects is presumably an ineffective board. However, no such patterns among the criteria emerged in the analysis of the data. Instead, in most cases an IRB's score on one of the measures was unrelated to its score on other measures. Thus, for example, there was no relation between evaluations of a board by its members and evaluations by the investigators whose research it reviews. Of the few statistically significant relationships found between indicators of performance, almost as many were negative as were positive.

The boards that made the most common types of modifications in proposals tended to receive lower evaluations from investigators. Thus, IRB's that made frequent requests for more information from investigators were evaluated in less positive terms by investigators. Similarly, at institutions where IRB's made relatively frequent modifications concerning consent, investigators more frequently disagreed with the statement that the IRB protects the rights and welfare of human subjects. These findings suggest there may be a trade-off between IRB activity and investigator acceptance, particularly when investigators do not see a link between the IRB's actions and the protection of subjects. Such trade-offs among the criteria of performance were found only infrequently, however.

The data were analyzed to see the extent to which variations in procedures, policies, and composition of IRB's were associated with differences on the various measures of performance. Few significant relationships were found, and among these few no consistent pattern emerged.

The operation of the review process was viewed more favorably than unfavorably by most research investigators and IRB members (see Table 5). However, a substantial minority, particularly of the investigators, felt that the review procedure is an unwarranted intrusion on the investigators' autonomy, that the IRB gets into inappropriate areas, that it makes judgments it is not qualified to make, and that it impedes research. The attitudes of the board members were slightly more positive than those of investigators. Among the investigators surveyed, the behavioral researchers had the least favorable attitudes. The problem (from a list of ten problems) most frequently cited by board members was getting members together for meetings. More than one-fourth of the IRB members indicated as problems the need for rapid action to meet deadlines imposed by funding agencies, the lack of precise HEW guidelines, and the time spent unnecessarily reviewing research with little risk.

Attitudes of Research Subjects

Investigators who found it appropriate to cooperate in this aspect of our research sent letters to their subjects indicating that the Survey Research Center wished to interview them about their experience in research. Only those subjects who returned a postcard indicating their willingness to be interviewed were contacted. This procedure was employed to protect the privacy of the subjects of the research under study, and it complicated the inherent difficulties of contacting such a sample. We were unable to obtain a true probability sample of research subjects, and the sample cannot be considered representative. Furthermore, periods of up to a year had elapsed since some subjects' participation. These data, therefore, must be interpreted with caution.

Most subjects or third parties recalled giving consent for participation, but one out of ten indicated that they did not understand that they were to be involved in "research." The majority, however, felt that they had been given clear, sufficient, and accurate information about the project in which they participated. The single most prevalent reason for subjects' participation was the ex-

Table 5. Attitudes of different types of investigators and review committee members toward the review process.

Statement	Percentage agreeing with each statement					
	Review board members			Research investigators		
	Biomedical sciences (N = 370)*	Behavioral and social sciences (N = 135)*	Other (N = 220)*	Biomedical sciences (N = 940)*	Behavioral and social sciences (N = 395)	Other (N = 180)
The human subjects review procedure has protected the rights and welfare of human subjects—at least to some extent	99	99	99	99	96	98
The review procedure has improved the quality of scientific research done at this institution—at least to some extent	78	62	70	69	55	83
The review procedure runs with reasonable efficiency—at least to some extent	99	96	99	96	94	94
The review procedure is an unwarranted intrusion on an investigator's autonomy—at least to some extent	13	11	6	25	38	23
The review committee gets into areas which are not appropriate to its function—at least to some extent	39	24	27	50	49	39
The review committee makes judgments that it is not qualified to make—at least to some extent	28	21	20	43	49	25
The review procedure has impeded the progress of research done at this institution—at least to some extent	26	30	22	43	54	36

*The N's are approximate since nonresponse varied from item to item.

pectation of medical, psychological, or educational benefits. Almost all of the respondents (98 percent) felt that participation was voluntary; most felt positively about the experience; and two-thirds felt that they (or the subject) benefited directly. Thirteen percent, however, said that they had experienced unexpected difficulties. About 70 percent said they would be very willing to participate in a similar study again. Many of those who were less than willing gave as their reasons the time and travel involved, the unpleasant procedures, or the side effects. Other respondents said that their decision to participate in another study would depend on the nature of the research or its benefits to subjects.

Subjects and third parties who consented on their behalf offered a number of suggestions and comments, including the desirability for additional information about the research (expressed by 19 percent) and the need for more care or courtesy on the part of investigators in their treatment of subjects (expressed by 11 percent). Here are some sample suggestions from these subjects.

[P]eople could explain research more. . . .

. . . [G]ive participants as much knowledge as they want.

I am a physician. They could have given me more information. My education is equal to theirs. They treated me like an idiot.

. . . [I]t would have been nice to have gotten a letter explaining the results.

. . . [R]esearchers should make very clear the length of the study and what is involved.

They should warn people more emphatically about the possibilities of unforeseen side effects. They could talk to patients and get their perspective. I felt like a number in the hospital. . . .

There were also miscellaneous suggestions, such as to increase the benefits and reduce the risks of research and to do *more* research and make it available to more people, but these comments were less frequent than those regarding the conduct of the research. The majority of subjects did not offer any suggestions for improvement in the conduct of research.

Conclusions

A substantial effort goes into the review process for protecting human subjects. Although we cannot be precise about the total magnitude of this effort, a projection from our sample suggests that IRB members spent more than a third of a million person hours on IRB activities during the year of this study. Additional effort came from administrative assistants and from the researchers whose studies were reviewed. Although a judgment about whether this effort is worth the cost depends on the importance that one attaches to the issues addressed by IRB's as well as the alternatives that one sees as reasonable, the results of this study are of relevance.

Both the frequency with which IRB's require modification of proposals and the evaluations of IRB members and investigators suggest that the boards play a useful and valuable role. Institutional review boards have some direct impact on more than half of the proposals that they review, by requesting either modification of or additional information about proposed research. Board members and researchers are virtually unanimous in agreeing that the review procedure helps to protect the rights and welfare of human subjects. Furthermore, most of these persons agree that the institutional review process is a reasonably efficient approach to the task. Of course, a substantial minority of investigators believe that their review board gets into inappropriate areas, makes judgments that it is not qualified to make, and impedes research. On balance, however, fewer than 10 percent of the investigators felt that the difficulties of the review procedure outweighed its benefits in protecting human subjects. Most researchers, as well as board members, apparently recognize a need for the review of research, accept the legitimacy of IRB's, and are prepared to play their part in supporting the work of these boards.

This general level of acceptance and support bodes well for the review process. Obviously, the process will not work well unless the participants are willing to make it work, and this willingness seems to prevail at most institutions. However, we believe that an effective system requires more than the goodwill and common sense on which the present system relies almost exclusively.

Our data on informed consent illustrate the need for more effective actions by review boards. Most boards apparently confine their attention to the pieces of paper to be used to document informed consent rather than to the overall

process by which informed consent is to be sought. The consent forms themselves tend to be incomplete and difficult for ordinary persons to understand; some forms could probably not be understood by most subjects. We were unable to document that review boards help to improve forms in this regard. Forms that are difficult to understand when first submitted to boards for review are no more understandable after they pass the review and, therefore, they are as difficult when finally given to subjects as they were when first submitted to the board. The problem of readibility cannot be solely attributed to the use of medical and technical terminology; some of the difficulty, according to our analysis, is related to the complexity of sentence structure and the nature of many of the nontechnical terms that researchers use. The communicativeness of these forms, therefore, could be improved without sacrificing content, though consent forms should not ordinarily be relied upon as the primary device for providing information to subjects (*6*).

A second source of data about this general problem comes from the subjects of research themselves. Our "sample" of subjects is very select, and we must be cautious in drawing inferences on the basis of information provided only by those whom we were able to interview. Nonetheless, the information obtained from this group of respondents is of some value. While most of these respondents felt that they had been well informed and well treated in the research, one-third of them offered suggestions about how researchers might improve the way they do studies involving human beings. Most of these suggestions implied the need for researchers to communicate more effectively and to treat subjects with greater concern and sensitivity.

Although the ethical conduct of research involving human subjects requires that researchers have a certain degree of skill in communicating with and relating supportively to others, training in the techniques of effective communication is not an ordinary part of the education of researchers (*17*). Thus, it is probably not surprising that the data show a need for IRB's to attend to investigators' plans for obtaining informed consent from subjects. However, few IRB members have had any special training for that role, and IRB's are not performing this function effectively. Communication skills could be defined as a legitimate part of the methodological training of scientists in fields that employ human subjects, and efforts could be made to expand and make more systematic the training programs for IRB members and researchers that have been undertaken by various groups in recent years.

Improvements in the effectiveness of IRB's and in the ethical conduct of research might also be expected if IRB's received more feedback about their performance. Information from subjects about their experiences in research could be most useful to IRB's in improving the quality of informed consent. Much can be learned from subjects who were badly informed or who felt coerced to participate in research, as well as from those who have had good experiences. It also seems likely that IRB's would find useful some information about how they compare to other IRB's or to some standards of performance. The former comparison could come through the development of a newsletter or through workshops or similar activities in which members of different IRB's could communicate with one another. The latter comparison could result from a monitoring process that is educational in orientation. At present, IRB's have only their own past experience and the HEW regulations as a guide to their behavior. The value dilemmas at stake and the level of effort that is being expended by IRB's are of sufficient magnitude to justify serious efforts to improve their effectiveness (*18*).

References and Notes

1. B. Barber, J. J. Lally, J. Makarushka, D. Sullivan, *Research on Human Subjects: Problems of Social Control in Medical Experimentation* (Russell Sage Foundation, New York, 1973). Barber *et al.* (p. 148) report that 70 percent of the biomedical institutions they surveyed had a review procedure prior to the 1966 Public Health Service (PHS) requirement. Some of the committees were probably established in immediate anticipation of the PHS requirements, although more than one-third of medical schools had review procedures as early as 1960, according to L. Welt [*Conn. Med.* **25**, 75 (1961)].
2. Review requirements for clinical research at the National Institutes of Health's Clinical Center date from the early 1950's [M. S. Frankel, *Ethics Sci. Med.* **2**, 43 (1975); see also W. Curran, *Daedalus* **98**, 566 (1969)].
3. *Code of Federal Regulations*, title 45, part 46.
4. For an examination of some problems with this formulation, see B. H. Gray, *Am. J. Psychiatry* **134**, 907 (1977).
5. D. H. Cowan, *Case West. Reserve Univ. Law Rev.* **25**, 533 (1975); S. E. Marcy, thesis, Yale University School of Public Health (1974); K. Melmon, M. Grossman, R. C. Morris, *N. Engl. J. Med.* **282**, 427 (1970); E. J. Millstein, *The DHEW Requirements for the Protection of Human Subjects: Analysis and Impact at the University of California* (Research Management Improvement Project, University of California, Berkeley, 1974).
6. B. H. Gray, *Human Subjects in Medical Experimentation: A Sociological Study of the Conduct and Regulation of Clinical Research* (Wiley-Interscience, New York, 1975), chap. 3.
7. ———, *Med. Care* **13**, 318 (1975).
8. The study was confined to institutions from which HEW had accepted a "general assurance" of compliance with HEW regulations for protection of human subjects. A detailed description of the procedure used to obtain the sample, which was stratified by type of institution and weighted by research volume at institutions, can be found in R.A. Cooke, A. S. Tannenbaum, B. H. Gray, "A survey of institutional review boards and research involving human subjects," which is published in the *Appendix* to the *Report and Recommendations on Institutional Review Boards* of the Commission for the Protection of Human Subjects (Government Printing Office, Washington, D.C., September 1978), pp. 293–302.
9. This term comes from the National Research Act (P.L. 93-348) and is not of our choosing.
10. Proposals per meeting is used as a convenient rate for expressing IRB activity, but the figure should not be taken to mean that the average proposal is discussed for almost an hour since IRB meetings are not confined to the review of new proposed research.
11. This includes projects which were reported by investigators as "primarily intended to benefit subjects directly" or which the investigators estimated had a medium or high (as against no, very low, or low) probability of providing medical or psychological benefit to subjects.
12. The fact that 20 percent of the studies in the highest of our risk categories were not expected to benefit subjects does not necessarily indicate that unethical research is being conducted. As these matters are generally assessed, that question would depend on the importance of the knowledge to be gained in the research and the quality of the subject's consent. No separate analysis of these matters has been done to date on this specific subset of projects.
13. This rate, which is based on projects as the unit of analysis, is not precisely comparable with data from a recent HEW survey of injuries to research subjects, which used subjects as the unit of analysis. However, the rate of injury to subjects suggested by the two studies is similar. In the HEW survey of investigators who had received support from the National Institutes of Health and the Alcoholism, Drug Abuse, and Mental Health Administration, injuries (most of which were "trivial" or "temporarily disabling") were reported to have occurred to approximately 3.7 percent of the subjects. See the report of the *HEW Secretary's Task Force on the Compensation of Injured Research Subjects* (Department of Health, Education, and Welfare, Washington, D.C., January 1977), pp. IV-2 to IV-11. The task force concluded that "the risks of participation in nontherapeutic research may be no greater than those of treatment in other settings." The task force recommended that subjects injured in research conducted or supported by the Public Health Service should be compensated for certain injuries suffered in research. Only 40 percent of the institutions in the present study reported having a policy regarding the treatment of or compensation for injuries to subjects.
14. For a detailed empirical study showing the inadequacy of seeking consent by relying on consent forms, while ignoring the setting and circumstances in which consent is obtained, see Gray [(*6*), particularly chap. 8].
15. In some cases investigators who reported that they did not obtain informed consent may have meant only that they did not use a consent form. The confusion of the substance of consent with its documentation is not an uncommon error and may have negative implications for informed consent [→ B. H. Gray, *Ann. Am. Acad. Polit. Soc. Sci.* **437**, 37 (1978)].
16. R. Flesch, *J. Appl. Psychol.* **32**, 221 (1948). The "reading-ease score" is based on word length, that is, the average number of syllables per 100 words; and sentence length, that is, the average number of words per sentence.
17. For a recent article on the ethical socialization of physician-investigators, → J. J. Lally, *Ann. Am. Acad. Polit. Soc. Sci.* **437**, 86 (1978).
18. Many recommendations for the improvement of IRB's are included in the *Report and Recommendations on Institutional Review Boards* (Government Printing Office, Washington, D.C., September 1978) that has now been issued by the National Commission for the Protection of Human Subjects. The separate *Appendix* to that report contains a much more detailed presentation of the findings of the surveys on which this article is based.

19. This research was supported by the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research under NIH contract NOI-HU-6-2110. The findings and interpretations reported here are the sole responsibility of the authors. We thank M. Balter, B. Barber, S. Greenhouse, J. Lally, J. Makarushka, and D. Sullivan for advice in the planning of this study, and T. Berckmans, F. Brigilia, J. Donnelly, C. Goble, S. Lawson, D. McCulloch, J. Robertson, J. Scheffler, and J. Stanton for their contributions to this article.