

**Twenty-fourth Annual Meeting of
The Joint Committee on Standards for Educational Evaluation**

October 1–3, 1998

Yarrow Center

Augusta, MI

Attendees:

Joint Committee Members:

Dr. Edith Beatty (CESS)
Dr. Rolf Blank, (CCSSO)
Dr. Ruth Ekstrom (ETS)
Dr. Arlen Gullickson (Member-at-large)
Dr. Kevin Hollenbeck (NSBA)
Dr. Charles Moore (NASSP)
Mr. Wayne Rietberg (NAESP)
Dr. James R. Sanders, Chair
Dr. Gary L. Wegenke (AASA)
Dr. Robert Wilson (CSSE)
Dr. Donald Yarbrough (NCME)

Student Evaluation Standards Validation Panel:

Dr. Todd Rogers, Chair
Dr. Daniel Stufflebeam

Joint Committee Staff:

Ms. Mary Ramlow, Secretary
Dr. Jennifer Fager, Student Evaluation Standards Project Manager
Dr. Pamela Zeller, *Personnel Evaluation Standards* Coordinator

**JOINT COMMITTEE ON STANDARDS FOR EDUCATIONAL EVALUATION
MEETING MINUTES
OCTOBER 1—3, 1998
KALAMAZOO, MI**

Meeting Objectives

1. To conduct the business of The Joint Committee on Standards for Educational Evaluation
2. To review the project plan for developing standards for evaluation of students, to report on the progress of the project, and to review upcoming project work.
3. To review the project plan for revising *The Personnel Evaluation Standards* and to review the nominees for the validation panel.
4. To select, revise, or develop standards for student evaluations based on the work of the Panel of Writers.
5. To plan for work following the Fall, 1998 meeting
 - a. Selection of the Validation Panel for *The Personnel Evaluation Standards* revision
 - b. Identification of National and International Review Panel nominees for the student evaluation standards project.
 - c. Completion of specifications for the first draft of the student evaluation standards and the form for reviewing it.
 - d. Review of plans for reporting on progress to boards and members of sponsoring organizations.
6. To receive a report from the Validation Panel for student evaluation standards project.
7. To elect a new chairman for the Joint Committee to replace Dr. Sanders.
8. To develop a general report of the meeting for distribution to the sponsoring and cooperating organizations.
9. To address other matters as needed.

Thursday, October 1, 1998

Dr. Sanders welcomed everyone to the 24th annual meeting of the Joint Committee. Dr. Sanders described the history and purposes of The Joint Committee. He indicated that the committee is now working on two major projects, one to develop standards for evaluations of students and the second to revise the 1988 *Personnel Evaluation Standards*.

Members of The Joint Committee introduced themselves and the organization they represented. The two members representing the validation panel also introduced themselves. Dr. Sanders indicated that some representatives from the W. K. Kellogg Foundation may also join the meeting later, as it is funding two of the Joint Committee's projects.

Dr. Sanders described the different roles and tasks of the various meeting participants.

The agenda was discussed and adopted.

Chair's Report / Business Meeting

Dr. Sanders described the funding received for the two main projects. He indicated that the W. K. Kellogg Foundation provided funding that will support four years for the student evaluation standards project. The Kellogg Foundation also provided funding for one year for revision of *The Personnel Evaluation Standards*, with the understanding that if progress is demonstrated on the revision, other funding will be forthcoming for the second and third years of that effort.

Dr. Sanders reported that *The Program Evaluation Standards* are in the process of being translated into German and Spanish.

Dr. Sanders stated requests continue to be received to reprint the *Standards*, especially *The Program Evaluation Standards*, for textbooks, research articles, and dissertations.

A decision needs to be made concerning membership in the American National Standards Institute (ANSI). When the Joint Committee joined ANSI in 1989, the understanding was that it would not be charged membership dues. ANSI has changed their policy, and wants the Joint Committee to pay membership dues of nearly \$2,000 per year. Also, ANSI has begun an audit of standards setting organizations, and the organization is required to pay for the audit, which is estimated to be about \$4,000 for the Joint Committee. Dr. Sanders indicated that he has appealed and demonstrated with a balance sheet that the Joint Committee may have a profit each year from royalties of \$1,000 to \$2,000, but the membership and the audit process would bankrupt the Joint Committee. ANSI representatives said they understood, but they stood firm, since they have to require all members to pay dues and the cost of audits. An alternative to terminating membership would be to seek an ANSI member who would be willing to submit our standards as their standards, i.e., we would develop the standards and give them to another ANSI member that would submit them. When the standards are approved, they will be American National Standards, but the other organization will be listed as the accredited standard setting body.

A second alternative is for the Joint Committee to terminate its relationship with ANSI. The benefit of belonging to ANSI is credibility and some limited visibility, both nationally and internationally. They do provide workshops and bulletins and give us a chance to critique other standards, but most

of the standards we get are corporate standards from the business world, and not in areas where we have expertise.

Dr. Sanders indicated that a decision has to be made about the future of the Joint Committee with ANSI and opened discussion on the matter.

Dr. Gullickson asked about timing and frequency of the audit process. Dr. Sanders stated that the audit would occur in December 1998, and every 5 to 10 years thereafter. Dr. Gullickson suggested that all member organizations could be asked to share the cost of membership and audit.

Dr. Sanders stated that he believes the benefits of ANSI membership are very limited. He indicated that APA , AERA, and NCME test standards committee decided not to apply for ANSI membership because of the cost.

Dr. Gullickson stated that he believed the credibility that comes with ANSI membership and approval of the standards is important and worth retaining. Dr. Stufflebeam also stated that he believes ANSI membership and accreditation is important to the work of the Joint Committee. Dr. Stufflebeam explained that being recognized by ANSI means that this Joint Committee is the only recognized body in the U.S. for setting standards for evaluation in education. Dr. Stufflebeam expressed concern that many other organizations are developing standards, and he would like to see the Joint Committee continue to grow, become stronger and be recognized as the group to set standards. He indicated that giving up ANSI membership may lead to parallel efforts that would not serve the public schools very well.

A question was raised about cost involved in obtaining another sponsor to submit the Joint Committee standards for approval by ANSI. Another question raised was if the Joint Committee were to go with an umbrella organization, would they get the royalties from publications? Dr. Sanders stated he did not know about royalties, but the other organization would have their name on the standards, and there may be some loss of control over the standards.

Discussion was held about what ANSI member organizations possibly could be used Dr. Sanders replied that ASTD, the training organization, might have an interest. We could consider another training or human resource development organization that is a member, that is more in line with what education is about than going to a large corporation.

Dr. Gullickson discussed the possibility of assessing dues to each organization to pay the cost of ANSI membership and audit. He also expressed concern that some organizations have not been actively involved, and a dues requirement may have negative or positive consequences.

Another member indicated he had experience with another joint committee that did assess dues to its member organizations and that worked very well.

A question was raised about whether funding for the standards projects is dependent upon ANSI membership, and Dr. Sanders indicated that is not the case at the present time. Dr. Sanders also indicated that another option for funding is to ask Kellogg or another foundation for support to pay for ANSI membership.

Dr. Sanders asked for two votes, one for the continuation with ANSI, and one for funding options.

Dr. Sanders requested suggestions for funding options. Suggestions and discussion included:

- service fee type billing of \$250 for each sponsoring organization
- NCME and AEA strongly argued against service fee in the past
- Kellogg Foundation may fund ANSI membership for the next three years (as long as there is a project in the works). However, once a grant is terminated, they probably would not give an annual grant for continued membership.
- there was discussion about eliminating organizations who will not support the work of the Joint Committee through membership fees and regular participation of their representative.
- there was discussion about having levels of membership, active and associate membership, with different fees
- it was pointed out that previously there was opposition to a service fee, but there was not a pressing need for additional funding for ANSI.

Dr. Sanders asked for volunteers for a task force to work on funding sources to cover ANSI membership costs. Arlen Gullickson, Edith Beatty, and Wayne Rietberg agreed to gather options and make a recommendation to the full group on Saturday.

Dr. Sanders indicated the committee needs to make a decision about one of the sponsoring organizations. CORPA, the Council on Recognition of Postsecondary Accreditation, has gone out of business. A motion was made by Kevin Hollenbeck to remove CORPA from the Joint Committee list of sponsoring organizations. The motion was seconded by Ruth Ekstrom to remove CORPA from the list of sponsoring organizations. A vote was held and the motion approved.

Dr. Sanders stated that the Council of Great City Schools has asked to become a sponsoring organization of the Joint Committee. They represent the 50 largest school districts in the country and schools located in each state's capital city. Dr. Sanders indicated that this organization would bring the urban city public school perspective to the Joint Committee, which would be desirable. A motion was made by Ruth Ekstrom and seconded by Gary Wegenke to extend an invitation for membership to the Council of Great City Schools. A voice vote was held and the motion approved.

Financial Report

Mary Ramlow presented the financial report. Individual sales reports for *The Personnel Evaluation Standards* and *The Program Evaluation Standards* books were presented. Mary pointed out that *The Personnel Evaluation Standards* sales have slowed, and it is hoped that the revision will generate new interest and increase sales. Mary also stated that the funding of projects by the Kellogg Foundation will offset expenses considerably this year.

There was a question about expenses in years when there is no project funding, and Dr. Sanders indicated that the annual meeting costs approximately \$3,000 and general operating costs may be another \$3,000 to \$4,000, and in the past the royalties have just about balanced the expenses. With funding for projects to offset expenses for the next three years, the Joint Committee will be able to increase its equity.

**JOINT COMMITTEE
BANK ACCOUNT STATEMENT AS OF
AUGUST 31, 1998**

Balance as of August 31, 1997		\$15,626.25
<u>Income</u>		
Royalty Payment from Sage Publishers	\$4,708.59	
Interest 7/1 – 9/30/97	\$150.45	
Interest 10/1 – 12/31/97	\$146.58	
Interest 1/1 – 3/31/98	\$137.88	
Interest 4/1 – 6/30/98	\$158.01	
Revenue from sale of Step Guides and Training Manuals	<u>\$440.78</u>	
Income Total		\$5,742.29
<u>Expenses</u>		
Payment for expenses to The Evaluation Center:		
Duplicating-\$842.65;		
Telephone-\$99.02;		
Postage-\$31.87;		
Office Supplies-\$61.80;		
Personnel-\$850.22;	\$1,885.56	
1997 Annual Meeting Expenses		
Dr. Gullickson-\$40.05		
Dr. Sanders-\$194.68		
Mary Ramlow-\$15.50	\$250.23	
1998 Delaware Annual Registered Agent Fee	\$99.00	
1997 Annual Franchise Tax Report	<u>\$19.99</u>	
Expense Total		<u>\$2,254.78</u>
 NEW BALANCE AS OF AUGUST 31, 1998		 <u>\$19,113.76</u>

SALES REPORT
THE PROGRAM EVALUATION STANDARDS

Published – 1981

	<u>SALES</u>	<u>RUNNING TOTAL</u>
Published by McGraw-Hill, Inc.		
9/1/83 – 12/31/83	538	7,547
1/1/84 – 12/31/84	1,390	8,937
1/1/85 – 12/31/85	1,236	10,173
1/1/86 – 12/31/86	1,033	11,206
1/1/87 – 12/31/87	682	11,888
1/1/88 – 12/31/88	1,042	12,930
1/1/89 – 12/31/89	987	13,917

Change in overseer of account to Ira Seibel as of 1/1/90

1/1/90 – 12/31/90	1,279	15,196
1/1/91 – 12/31/91	1,772	16,968
1/1/92 – 12/31/92	579	17,547
1/1/93 – 6/30/93	219	17,766
7/1/93 – 12/31/93	25	17,791

***The Program Evaluation Standards* revised and printed by Sage Publications, Inc.**

1/1/94 – 12/31/94	2,612	2,612
1/1/95 – 12/31/95	2,398	5,010
1/1/96 – 12/31/96	2,249	7,259
1/1/97 – 12/31/97	1,185	8,444

SALES REPORT
THE PERSONNEL EVALUATION STANDARDS

SAGE PUBLICATIONS

Published – 10/88

	<u>SALES</u>	<u>RUNNING TOTAL</u>
As of 12/31/88	1,175	1,175
1/1/89 – 8/31/89	3,778	4,953
9/1/89 – 12/31/89	874	5,827
1/1/90 – 12/31/90	1,001	6,828
1/1/91 – 12/31/91	963	7,791
1/1/92 – 12/31/92	341	8,132
1/1/93 – 12/31/93	192	8,324
1/1/94 – 12/31/94	736	9,060
1/1/95 – 12/31/95	321	9,381
1/1/96 – 12/31/96	453	9,834
1/1/97 – 12/31/97	302	10,136

Report on Projects

Dr. Jennifer Fager reported on the Student Evaluation Standards project, and referred to the notebook that was distributed. Dr. Fager indicated that members of the Joint Committee provided approximately 50 names of individuals who would be willing to write draft standards. Listserv requests were sent out to all of the AERA divisions, AEA, and others. She stated that the panel of writers came from diverse areas providing broad perspectives. Dr. Fager explained the literature review and bibliography that was completed. It has been submitted to the *Educational Measurement Issues and Practices* journal.

Dr. Fager also indicated that Jim Sanders, Dan Stufflebeam, Arlen Gullickson, Dianna Newman and she will be presenting this project at the AEA conference in November. It will also be presented at the NCME conference this year. Dr. Fager stated that she submitted this to AACTE, which is another teacher education organization.

Dr. Sanders indicated that in the coming year, international and national reviewers will be asked to review the first draft of the standards, and in August, results of those reviews will be mailed to members. Also, members will be asked to identify field test sites. Next October at the annual meeting members will be revising the first draft based on the reviews to prepare for the field tests.

Dr. Pam Zeller reported that beginning in August a request was sent to each member of the Joint Committee asking for input and suggestions for revisions on *The Personnel Evaluation Standards*. Also, a notification was put on Evaltalk, the listserv, which referred to The Evaluation Center website and the standards. A few comments were received, but not a great number.

Dr. Sanders asked that each member put a notice in their organization's newsletter soliciting reviews of the standards, and indicated this is a 3-year project and recommendations can be accepted anytime.

A suggestion was made to have member organizations link to The Evaluation Center website, so the standards can be posted and the solicitation for input. Dr. Gullickson stated that member organizations would have to create the link to The Evaluation Center website.

A question was raised about what will be expected of people who volunteer to work on revisions of *The Personnel Evaluation Standards*. Dr. Sanders said we want people who use, or would like to use, the evaluation standards to recommend revisions either for outdatedness of references, to new methodologies that have appeared since 1988 that should be included, or a bias they see in the standards toward one way of thinking about personnel matters, philosophical differences or methodological advances, new references, clarification, ideas for guidelines or common errors, all would be welcome for suggestions on how we can improve the second edition.

Dr. Gullickson said that other places to look for input include all 50 state departments of education, and in the regional laboratories, particularly SERVE, Southeast Regional Vocational Education Laboratory, and the Northwest labs as well.

Dr. Zeller stated that revisions were made in the 1996 Joint Committee annual meeting, and those have now been organized to compare revisions with old material. She indicated many of the standards need to be updated, particularly regarding technology, adding to reference lists, and illustrative cases. It is important to have reviewers keep computer technology in mind, and its impact on personnel standards. Dr. Zeller suggested that it may be helpful to have someone outside the education field look at the standards from the perspective of protection of personnel when disseminating information via e-mail or websites.

Dr. Sanders stated that in 1996 the Joint Committee spent two full days going cover to cover through *The Personnel Evaluation Standards* to determine what changes we thought needed to be made, and that is a major input for revision.

Dr. Sanders said that in the next year the first draft of the second edition will be completed and sent out to international and national review panels. Members need to recommend reviewers from their associations for *The Personnel Evaluation Standards* reviews. Those reviews will be used to prepare a second draft. At the annual meeting next year, the results of the review and a field test plan will be available.

1999 Annual Meeting

Dates and locations for the 1999 annual meeting were discussed. Members agreed on September 29, 30, October 1 and 2, 1999. Discussion was held about changing the meeting place to improve attendance. Other suggestions included Reston, VA, and Hilton Head, SC. Dr. Sanders stated they would check into alternative sites, and notify members later.

Other Business

Dr. Sanders asked for other business. There was a question about the Executive Committee, and Dr. Sanders said that an election for a new member and Chair will be held Saturday. He indicated a Vice Chair is usually proposed by the Chair, usually a member-at-large not representing any organization.

Work Session

Dr. Sanders asked the Joint Committee members to divide into four work groups to work on the Student Evaluation Standards.

1. Utility Standards
Pam Zeller, Facilitator
Rolf Blank
Kevin Hollenbeck

2. Propriety Standards
Gary Wegenke, Facilitator
Wayne Rietberg
Edith Beatty

3. Feasibility Standards (3) and Accuracy Standards (1st 4)
Jennifer Fager, Facilitator
Charles Moore
Bob Wilson

4. Accuracy Standards (last 8)
Arlen Gullickson, Facilitator
Ruth Ekstrom
Don Yarbrough

Dr. Sanders explained the task is to develop a good first draft of each of the 7 or 8 standards. Dr. Sanders said there appears to be some confusion about the focus of the standards, and he clarified by stating that the standards are intended for evaluation of *students*. It is not to evaluate a program, classroom, teacher, school district, statewide assessment program, nor a testing system. Student needs, performance, etc., is the focal point.

We are looking for good support for whatever statements we make about the standards. If you have specific references to the topic, we need to add in those references. Each small group will have original writing to do to fill in where it is just not good enough yet, so that is to be expected. Pieces of submitted standards may be used, and some may need to be created.

A third thing is that some of the standards were off target, but rather than throw them away, it may be useful in another standard. So, you may want to carry a standard over to one of the other groups if it covers another domain.

Dr. Sanders confirmed that members need to consider all types of students, including special education assessment, classroom and large scale assessment, in training programs in corporations, not just K-12, but also higher education and corporate settings, anywhere there is an instructor and student. The term “education” is used very broadly. Members also need to be clear that we are addressing evaluation issues, not developing standards for testing programs.

Dr. Sanders clarified the difference between testing and evaluation stating that testing is a data collection process, and evaluation is the interpretive process, i.e., valuing, determining merit or worth of the student’s performance, and their needs. He also indicated that the recent literature

search was the basis for the student evaluation standards categories. In addition, *The Program Evaluation Standards* and *The Personnel Evaluation Standards* were used to keep wording consistent across the publications. Dr. Sanders indicated the topics are open to revision.

Dr. Sanders explained mechanics of revision process, use of notebook, forms, etc. Dr. Sanders asked the members to break up into small work groups and return at 5:15 p.m.

Break

Members reconvened in large group

Dr. Sanders asked for a brief report from the groups.

Utility Group Report

Pam Zeller stated that issues that came up included (1) evaluation vs. assessment and, (2) the issue of values, i.e., how it fits into the assessment and the whole idea of the evaluation and should it be a standard or should it fit in somewhere else?

Another issue was which things had higher priority, and how they fit together, but there was a necessary overlap of things like values. The way the writers saw it and the way the group saw it was different.

Dr. Sanders addressed the issue of evaluation vs. assessment. He stated that evaluation has two components. One is data collection, which is necessary in order to be informed, i.e., testing, assessment, data collection in many forms. The other component is judgment, i.e., where we arrive at some justifiable conclusions about the student, if they are doing well, excelling, acceptable, other judgmental conclusions we might make. Evaluation is not evaluation if it is missing judgment. It is the judgmental or valuing process that makes evaluation what it is. If we have just data collection, we have just testing, or just observation, without judgment; that is not evaluation. That is the distinction to be made in the book between assessment, testing and other forms of data collection, and evaluation which adds the second component.

Dr. Zeller stated that most writers omitted the judgment component, and were writing about assessment only.

A running list of terms that need to be defined is being recorded.

Discussion was held about confusion over definition of words “assessment” and “evaluation” that are sometimes used interchangeably. It was suggested that “evaluation” is a more stable term than “assessment” which should be avoided. There was disagreement about the use of the word “assessment” and suggestion that the standards are intended for a diverse audience that may define terms differently, and this may be an editorial question which can be addressed later.

It was suggested that illustrative cases should be included for many settings, e.g., K-12, higher education, and private sector cases.

Propriety Standards Group

Dr. Sanders asked Dr. Wegenke if there were any issues that need to be brought up to the large group. Dr. Wegenke stated that many of the written statements were restrictive, and they tried to broaden them somewhat for more flexibility. He suggested that some of the material was so extensive as to be intimidating, and they were attempting to make it more user-friendly by prioritizing.

Dr. Sanders suggested that the draft should be more inclusive at this point, and leave reduction process to a later editing. He also stated that the drafts are just suggestions, and do not have to be used if it is not appropriate. Dr. Sanders said that the final product from the small groups needs to be a clearly written standard on the blank form provided.

The Feasibility/Accuracy Group

Dr. Fager stated that there were no issues for the whole group. This group recommended that the bibliography be sent to all members of the Joint Committee and have them go through that and find references that fit the standards that they worked on. Also, one case (A-1) was excellent, and could be used as a model.

Accuracy Group

Dr. Gullickson reported that the work they reviewed was germane to testing standards, but not to evaluation standards, so there is considerable rewriting to be done.

Dr. Sanders said that some of the groups split up and worked on standards individually to get the first draft done. He also said that because these standards are going to be most useful to teachers, perhaps there should be two teachers on the validation panel. Dr. Fager suggested contacting AACTE and ATE, because they include teachers and teacher educators. Dr. Sanders said he could work with Todd Rogers to get teachers on the validation panel. It was also suggested that a school administrator should also be included on the validation panel.

Dr. Sanders clarified the role of the validation panel, which is to identify and examine the assumptions underlying the work, so they are being analytical. Second, they are to critique and report on the Joint Committee's procedures, that the operating procedures have been followed. They are to assess the applicability of the standards in various contexts, i.e., national and international contexts.

Dr. Gullickson raised the issue that it may not be possible to include a representative from every stakeholder group on the validation panel, and the necessary representation may already be in

place on the Joint Committee and its panels. Dr. Sanders suggested that the group come back on Saturday to the issue of expanding the validation panel.

There was discussion about taking standards home to redraft over the next week, and it was decided that this would be done only if absolutely necessary.

Members noted that references need attention, group members are not familiar with supporting references.

There was also a discussion about excluding jargon, and emphasis on cases that teach/stretch the reading.

The Feasibility/Accuracy Group

Suggestions include: define terms listed. Make sure case studies are not just K-12 classrooms, but include K-16 classrooms, human resource, private sector situations, where appropriate. Also, have the Joint Committee members look at the bibliography to determine which things might support their standards. Also need the Joint Committee members to add to the bibliography.

In terms of designing the guidelines, the focus should be on what an effective evaluator does. On the errors section, the focus should be on what poor evaluators do.

The standards are finished, and there are two cases for each.

Bob Wilson suggested that the supporting references be the best and most current, and one or all members should undertake a professional literature search, and make sure the sources have been read. It was also suggested that a wide scope of sources are used to include all potential users' fields.

Accuracy Standards Group

Arlen Gullickson reported that on all standards they completed overview, guidelines, and errors, and in some there are case study information. But a great deal of the material had to be reconstructed. Two of the standards, qualitative and quantitative analyses, are being combined into one.

Don Yarbrough said that three standards evolved over the course of previous projects. One was systematic data collection, which had moved in the direction of quality control. Then the attention to qualitative and quantitative issues appeared in the 1994 edition. It doesn't make sense now, and not thinking about four to five years from now, to differentiate qualitative from quantitative, with regard to procedural matters that have to do with quality. We thought we could maintain the emphasis on quality in two standards, one to do with systematic data collection and processing, and one having to do with systematic analysis and interpretation. Integrate qualitative and quantitative concerns in both standards with the general motif of quality control.

Try to capture everything in two standards that made organizational sense in a way we thought would be more accessible and understandable to the audience. But again, there is great concern about the cases—both need to be thoroughly edited one more time.

Don Yarbrough also asked if there would be an opportunity for all members to read through these again, and give one another feedback before the release to the next round of reviewers. It was suggested that could be accomplished in a month. This process could include members of the Joint Committee who were not able to attend the annual meeting as well.

Internal consistency could be addressed in a second reading as well.

It was suggested that if this is to be published three to four years in the future, references will need to be updated just prior to publication.

Dr. Gullickson said he has been creating an instrument exchange, intended not for the Joint Committee, but as a vehicle to facilitate exchange of instruments, checklists, tests, surveys, and other things among evaluators. From your school, you can go onto the web, upload the document to our server, and it then becomes available for others to read and make comments on it. And we could use that vehicle as a means for all of us to sit at home to read the various standards, and we could earmark them as standards such-and-such. That would enable you to get the comments so that all of us on the committee could read it and make comments on it after we have read it. And that set of comments could be plugged into the standard, so that everyone on the committee can read what others have said in reaction to the particular standard.

There was general agreement on using an instrument exchange via the web.

Don Yarbrough also raised the issue of how to deal with emerging technology, e.g., computerized testing. That would be a good example for a case, if someone has experience with it.

It was suggested that technology and diversity issues should be embedded in all the standards.

Another issue is grading—although it does not need to be its own standard, it should be addressed. Dr. Sanders indicated that grading is one of the decisions that is made about students based on student evaluations, like tracking, matriculation, promotion, admissions, and placement. All of those are decisions, and perhaps we need a matrix with the different kind of decisions and the standards, and which standards relate to which decisions. It may be all standards are linked to all decisions.

Don Yarbrough stated that grades are unique, in that we make a decision about what grade to give a student. But further, that grade becomes the basis for a number of additional decisions made about students, so it is compound. In a way, a grade is a reduction of information on which further decisions are made.

Dr. Sanders suggested writing grade information into the analysis and interpretation standard. Don Yarbrough said he sees it in a “Justifiable Conclusions” or if there were a recording standard, but it may fit into analysis.

Dr. Gullickson said that if you take the number of grading decisions made in schools and put them in one category, and put everything else on the other side of the scale, they overwhelm everything else. And they drive everything else. And so what Don is saying is really important. At some point we have to come down pretty squarely on where we stand in terms of grading. What are appropriate strategies to employ, how do you go about the process. If there is anything schools and teachers need information on, it is grading practices.

It was suggested that grading practices be written into the introduction to the book, where it can be described in detail, so that teachers using this book understand that the grade is the reduction of many observations, and many tests, and techniques. And we hope the book will help the teacher understand the full range of assessment opportunities available to make sound grading decisions.

Dr. Sanders said that one thing that was done in the other standards was to provide an application of the standards to show how to use the standards in a stepwise process. We could make that application a grading application of the entire set of standards. That would be very large, but it is a good suggestion.

Dr. Gullickson stated that is like the functional table of contents in the other two sets of standards. Grading is one of those places where you would turn to a functional table as well in the design of the evaluation process. Has any attempt been made to decide what are the functional issues that we ought to address once this set of standards is complete, so that when we create these standards, we are aware of that and attend to that matter. Dr. Sanders indicated that no functional table of contents has been developed yet.

Dr. Sanders asked Dr. Ekstrom if any of her colleagues at ETS would be willing to write 1 or 1½ page analysis of a case related to the use of technology and student evaluations. Dr. Ekstrom requested more information about who and in what situation. Dr. Sanders suggested a case where computer based testing was used and affected validity of results. He did not have a specific request, but rather to tap into those who have good ideas in this area.

Steve Wise was recommended as an expert in computer testing. Jennifer Fager agreed to attempt to contact him.

Dr. Tom Wilson, who is principal of Egan High School in the Rosemont School District is a techno-whiz, and has a whole school organized around technology. He did a presentation last year at NASSP, and published a couple of articles in a bulletin regarding technology. Charles Moore agreed to contact him to ask for his involvement in writing a case analysis.

Another case was suggested where other technologies are used for assessment of students with disabilities.

Dr. Gullickson stated that there are many good issues outside computerized testing. He gave the example of the common use of spread sheets for collecting, assembling, managing student grades. The new technologies make it easy to make massive errors by leaving out one small piece of information—not just individual errors but whole classes of errors. It was agreed this would be a good case for the systematic data collection standard.

Another case scenario was suggested, wherein computer testing is used that adapts to the level of difficulty, giving the student increasingly easy or difficult questions depending on whether previous answers are correct or incorrect. How does this affect the student's performance?

Dr. Gullickson raised another issue, as in Kentucky and in Dallas, Texas, they ran into problems in standardized testing, when the teachers changed the student's responses to test items in order to improve the student's scores, because the consequences for the teachers are bad if the students fail. Is that a student evaluation problem? The student is being graded inaccurately by being told they are doing better than they are. There are many "games" being played with standardized assessment, where scores are manipulated in one way or another, and policymakers do not understand what is going on in the assessment process. Dr. Sanders agreed that using these "games" as cases is a good way to expose them.

A suggestion was made to have the cases written, and decide where they fit later. It will be to look at the organization of a school student evaluation program based on interactive technology. That prepares a way for teachers to input in the evaluation process, a way for students to access it, check it, and a way for parents also to access and check grades. Another might be to look at the relationship between this evaluation process and the final grades, i.e., how does all this become synthesized down to a letter grade—all of the inputs by the students, teachers, and parents. And it might make sense to describe a case that has to do with how parent and child can access homework assignments in that process. To what extent does this access affect the evaluation process, evaluation by the parent as well as the teacher?

Dr. Sanders asked the group to reconvene after dinner to discuss Validation Panel and Review Panel business. He asked for members to give suggestions for the kinds of people to get for reviewers, and to work with member organizations in the next month or two to get those people. The process is to send a card to ask them to participate, and once they agree, we send them a packet of information, so the initial contact should be in November or December.

The Personnel Evaluation Standards were distributed to all members and Dr. Sanders explained the changes made to date and tasks for small groups on Friday. Dr. Sanders also stated that a Validation Panel needs to be appointed for *The Personnel Evaluation Standards* revision. He recommended that Diana Pullin be appointed as Chair of the Validation Panel. A motion was made to ask Diana Pullin to chair the Validation Panel and the motion seconded and carried. He also asked for names for the Validation Panel Members from area of personnel psychology,

teaching research, education philosophy, international education/cultural issues, education law, and education administration. Dr. Sanders also requested nominees for Review Panels for the revised *Personnel Standards* and for Student Evaluation Standards first draft. Dr. Sanders described the tasks for the Review Panels.

The ANSI task force then gave their recommendations.

1. Keep ANSI membership and pay the approximate \$6,000 for costs from the bank account as a temporary measure;
2. Indicate to each member association there would be \$1,000 fee that would be waived if it has a member in attendance at the 1999 and each year's meeting. This was in addition to number one, which is a temporary fix. This would also clarify which organizations are willing to give support.
3. Assess each member association approximately \$200 fee to pay ANSI costs, and from a straw vote see if that has real possibilities for members that are going to drop out and withdraw their representative.

Don Yarbrough asked for discussion on the up and down sides of each of these recommendations.

Dr. Gullickson stated that the \$1,000 fee came in as a way to move inactive organizations toward participating. There was concern that if a fee is levied, the Joint Committee might lose a large number of organizations, and it was suggested that the effect of many organizations dropping out should be considered.

Dr. Sanders suggested that rather than just assessing a fee, someone from the Joint Committee should meet with organizations and explain the need for additional support for ANSI costs, and ask if they would be willing to pay a portion of this costs. If they agree, we are set to assess a fee next year. If they say no, we ask for reasons why. Those reasons may be a basis to rethink whether it is worth participating with ANSI.

Dr. Yarbrough stated that committee members should make the case to their organizations of the value of ANSI participation, and that value needs to be clarified. Perhaps a position paper should be written on that. Dr. Sanders stated that he has a letter from ANSI outlining the benefits of membership. There was discussion on whether ANSI membership needs to be "sold" to organizations.

There was agreement that the issue of membership in ANSI and the issue of nonactive Joint Committee sponsoring organizations should be separate. The following questions were suggested: What can we do collectively to find out why the teacher organizations are not participating? What can we do to get them here? What other reasons are there?

Dr. Sanders indicated that the most important variable is who the person is that is appointed as representative, and how responsible they are. There are some members who do not put the Joint

Committee as a high priority. Dr. Sanders indicated he has contacted inactive members by letter, with a poor result.

There was a suggestion that the Executive Committee could negotiate and ask for support in different amounts from different size organizations for continuing ANSI participation. There was discussion about the value and/or importance of retaining all current sponsoring organizations.

Discussion was also held on having different levels of membership. Dr. Sanders explained that there are currently two levels of membership, i.e., sponsoring organizations and cooperating organizations, and he explained the difference.

Dr. Yarborough made a motion to maintain ANSI membership and audit, and allocate sufficient funds from the bank account to cover the cost. The motion was seconded and carried.

Dr. Sanders stated that the Chair and Executive Committee would work with inactive organizations to promote better participation.

Dr. Sanders asked for any other business, and there was none.

Dr. Zeller distributed suggested revisions for *The Personnel Evaluation Standards* and a new standard that has been recommended by Dr. Stufflebeam and Diana Pullin. Dr. Sanders asked the members to review these in preparation for revision work the next day.

The meeting was adjourned for the day.

Dr. Sanders opened the a.m. session by thanking the members for their participation and professionalism.

Dr. Sanders distributed copies of a letter outlining the benefits of ANSI membership and gave a list of benefits:

- membership
- have a voice in the development of national standards
- “Standards Actions” journal
- provide information, services, networking, workshops and conferences on standard setting
- added credibility of ANSI membership

Dr. Sanders distributed an outline for the first draft of the Student Evaluation Standards and asked for feedback. He indicated he was trying to stay consistent with *The Program Evaluation Standards*.

There was discussion about the need for clarification of terms in the beginning of the Student Evaluation Standards whether it requires a special section; part of the preface, background, or introduction.

Dr. Sanders indicated he expected to use the section on “definitions” to clarify terms. He said he would work on the “Introduction” Section, and most of the other sections will remain placeholders for the first draft. Also, the glossary and bibliography at the end will not have completed text. So the first draft would consist of the “Introduction” “Applying the Standards” and the individual standards.

Dr. Sanders asked for reactions to criteria for reviewers. There was discussion about benefits of inclusion of a functional table of contents and definitions of terms in the first draft. Also discussed was inclusion of a section on supporting documentation. Dr. Sanders said there would be 50 to 60 national people and 50 to 60 international people to be reviewers of the first draft.

Dr. Gullickson suggested taking more time to review and construct the draft student evaluation standards, as it has to be in good enough shape for the review to be productive.

Dr. Gullickson argued for spending more time to have all standards reviewed by all members to make certain all members are in agreement on the first draft before it goes out, particularly when substantial changes were made in the standards submitted by the panel of writers. There was concern about alienating some organizations if these substantial changes are misinterpreted. It was also suggested that clearly labeling it as a working draft, not nearly a finished product, would help prevent misinterpretation.

It was suggested that there be a two-step review process, where the first draft is seen by 8 or 10 trusted reviewers, before going to the large scale review.

Dr. Stufflebeam gave some historical information on the review process, and emphasized clarifying the difference between the evaluation standards and test standards.

Review of *The Personnel Evaluation Standards*

Dr. Sanders asked the group to move on to *The Personnel Evaluation Standards* revisions. He indicated the members would break up into the same small groups as they did previously. He gave instructions for review and revision of *The Personnel Evaluation Standards* and asked for progress reports on return to the big group.

Dr. Zeller asked the members to review references and cases, which need to be updated.

Dr. Sanders reconvened the large group, and asked for feedback on the suggested changes and to raise any issues or thoughts.

Don Yarborough reporting for the Accuracy Group (Arlen, Don and Ruth) indicated they continued to talk about some unresolved fundamental issues on the student evaluation standards, and they asked for time to discuss those issues after the other groups had reported on *The Personnel Evaluation Standards* revisions.

Charles Moore reported for the Feasibility Group, (Jennifer, Charles, and Bob) that they reviewed *The Personnel Evaluation Standards*, and there was some confusion about the context of why the changes were proposed or where they came from. They discussed personnel experiences, did some writing, but have many serious issues to continue to work on.

Wayne Reitberg reported for the Propriety Group, (Edith and Wayne), and said that there were not many changes in the area of Propriety, with the exception of the case study. He also suggested that the case was naively written in light of the state of the law in Michigan. The addition of “due process” either as a new standard or in the functional table of contents is an unresolved question.

Pam Zeller clarified that previous suggestions were that “due process” should be added to G, K, P1 and P2 or integrated throughout the standards.

The Utility Group (Pam, Rolf, Kevin) reported that they reviewed the materials and asked for clarification on IEP, and whether it meant individual education plan, as used in special education or whether it was synonymous with IDP, individual development plan, as used with teachers and instructors, which would make more sense in this context. There was also a general comment about creating standards that will remain relevant over time, using broad concepts and principles rather than including trendy jargon or labels.

There was a comment that working on two projects and switching gears was difficult, and options were discussed to prevent overload on the Joint Committee. A suggestion was that the time line be extended on *The Personnel Evaluation Standards* revision project and ask Kellogg for an extension of funding to allow more concentration on the student evaluation project. Dr. Sanders expressed concern that if *The Personnel Evaluation Standards* project is extended too far, the past work will become obsolete.

Implications for revision of standards—each standard should be looked at in terms of how it is related to policy, i.e., that changes in operational policy often have to be made before changes in a system, like evaluation, can occur. Dr. Sanders asked the members to consider that issue and bring in suggestions as to how that could be integrated into the standards or functional table of contents.

Dr. Sanders asked Don Yarborough to report on the unresolved issues from the work on Student Evaluation Standards. Yarborough said the fundamental concerns are:

1. Who is in the audience to be addressed? Addressing the combined needs of K-12 teachers, higher education, school administrators, and business and industry trainers may not be possible. What focus of student learning is being emphasized; is it assumed to be the classroom?
2. Lack of familiarity with other standards is problematic. Do other standards overlap, i.e., do testing standards overlap the accuracy questions? How do we make them

complementary? What *other groups* have standards that address the same issues? Should those be addressed in the group?

3. They had difficulty writing a full *model* student evaluation standard. It would be helpful if the group produced one satisfactory model for members to take with them and work from.
4. Timing of validation panel input, i.e., concern that the members have their input early, to catch issues the members may not see or comprehend.

Todd Rogers agreed to give the Validation Panel Report at 1:00 to address the concern. Several groups offered standards to be used as a model to copy and distribute.

Dr. Sanders explained that by using a higher level of abstraction when dealing with principles and standards, the needs of multiple audiences can be addressed.

There was discussion about including testing standards for teachers as an appendix or a bibliography reference.

Dr. Sanders stated that the primary audience for the student evaluation standards is made up of teachers, whether K-12, including principals, higher education, or business/corporate trainers, they are still teachers, who will be using the standards to evaluate students.

Dr. Gullickson had concern that different levels of teachers require different writing styles, and identification of teacher focus is important in determining the writing style.

Lunch break

Dr. Sanders reconvened the meeting after lunch and summarized previous discussion on clarifying the audience for the student evaluation standards. Dr. Sanders asked for input from individual members as to the audience. The following comments were made.

- ▶ not just teachers, broad audience, like other standards
- ▶ broad, but question of weight for cases, number target at teachers, & other
- ▶ broad audience
- ▶ primarily educators in the teaching process
- ▶ primarily teachers in the classroom, less in business & industry
- ▶ classroom teacher as evaluator of students, effective grading, supervisors, i.e., principal as metaevaluator; wants to see classroom teacher audience as primary focus, to achieve big impact in that arena.
- ▶ sees developing standards as general statements that people who are responsible for doing assessments of students could identify with; believes there is real benefit in generalizing and not localizing the audience; advised to express standards generically, especially to escape from the qualitative/quantitative paradigm wars; important that certain standards are more applicable to certain roles, and the text can indicate that if you are in a certain role, you focus more on certain standards; guidelines have to be reasonably generic, but

less so, because they are step-by-step; some people would use the standards as a place to go, then say, ok, now how do I do it? Those guidelines and common errors together provide that. As soon as you begin to do that, then you almost always have to focus on a particular case or issue, which takes it from the global to the specific, and that is the problematic area; would like to see the cases mostly school based. Would like more emphasis on administrators, as believes they have school policies and procedures that represent outdated mode of assessment, and they need guidance.

- ▶ Using cases from the military or medical field may be helpful to teachers to understand that evaluation is important outside of the school classroom.
- ▶ helpful to distinguish between general guidelines and specific cases; find the right level of generalization/abstraction; hope document leads to some level of enlightenment in the national discussion of student evaluation, and sees that as a major contribution that the Joint Committee could aspire to; often standards are directed toward an audience that is sophisticated beyond the level of the national dialogue.

Dr. Sanders summarized the discussion stating that there is some disagreement about the primary audience, but agree that audience is primarily teachers of people at different ages and settings, and that other Joint Committee standards will be used as general model. Then use the guidelines and common errors as places where more specific things are addressed.

Validation Panel

Dr. Sanders asked Dr. Rogers to present the Validation Panel report. Dr. Rogers complimented the group on their hard work. He said this is the first time the Joint Committee has taken on two major projects at the same time, and is concerned about not knowing what the impact will be on the quality of the two documents, unless there are some structural changes made in the way the work is done. Suggestions included:

- If all 16 members were present, there could be two groups working on different projects.
- Increasing the length of the meeting
- Special ad hoc task forces to take on certain tasks that the Joint Committee might otherwise do, e.g., if there are no cases from the military or business, acquire those through personal contacts with people in the business field. Cases could be solicited rather than trying to craft those.

Some changes need to be made to assure good quality products.

Need to be clear on the foci of the standards. How will these standards relate to other standards, e.g., in the illustrative cases there is something that deals with assessing a young person with special needs. That is covered in other standards.

The second issue is what the audience is for these standards and need to be clear about characteristics and education of audience in relation to measurement and evaluation. Need care in language used, concrete vs. abstract, and the more concrete—the more understandable. Perhaps illustrations provides sufficient concreteness.

In the work that was reviewed, need to reduce wordiness and overuse of passive voice. If one of your audiences is teachers, watch don't write over their heads.

Boundaries for project and focus, as discussed—is this including all levels of education or just certain levels?

May be wise for the Joint Committee to reconsider its decision to cover the whole spectrum and to consider setting boundaries to include student evaluation within K-12. If other areas are addressed, let the guidelines and cases illustrate that the standards apply here. As in Canada, modify the standards by adding statement, “With appropriate modification, you can use these standards at higher levels.”

Suggest determining balance or percentage of cases from K-12 and higher education.

Definition of terms has been discussed, and suggest inclusion in introduction, section on definition of terms.

Language agreement, i.e., children vs. students; teacher vs. instructor, or evaluator.

Lack of stakeholder identification in the standards. For example, in F2 the term stakeholder is used, but there has been no stakeholder identification. At the front of the Utility Standards may be the appropriate place to put it.

Could some standards could be combined, e.g., balanced evaluation (complete & fair previously) given that this all about fairness, why drop the word “fair”? Also, needs orientation – teachers serve students, and retention of term “descriptive service orientation” might be better.

Some summary statement of standards need to be rewritten, e.g., Priority 1, 3, & 6; Utility 1, 3, 6, & 7; and Feasibility 2 & 3.

On Common Errors and Guidelines, often the common errors stated are the reverse of the guideline. Yet given the limited amount of space, need the two of them complimentary to make the point. The group needs to come to some decision about how many positive and how many negative items you want, and suggestion is that there are more positive to model best practice, as that has more instructional value.

Areas overlooked, if using K-12--have something addressing observation, anecdotal reporting, the informal assessment that go on, particularly at the lower grade level. As we go up, it tends to be more formal. There is a shift in the assessment from the qualitative to the quantitative. There is a shifting of the mentality of the teachers when you go up the grade levels.

We need an case example dealing with inclusion of kids who are different, and how should they be assessed.

Do we need to address the Freedom of Information Act? Some teachers don't know what to do, because they don't know what parents can have access to.

On scoring, combining, ranking, rating, and grading, how do you get a term mark? What should be added together?

Cut off scores, frames of reference for interpretation—we didn't find anything there. We found allusions to a standard of performance, like 80 percent, but only in one illustration. The teacher takes a term mark and translates it into a grade, and there are different frames of reference to use for that. There should be some context for what they are. Perhaps the guideline, "Know what frame of reference you are using and the limitations, and take them into account in your interpretations."

"Payment by results" is in effect in the U.S., where someone gets paid if they get good results. What are the implications for student performance?

The assessment of special students identified by psychological testers, e.g., in Canada we have aboriginal students, who are asked to go to regular schools, and these kids really suffer because they think differently.

Dr. Stufflebeam stated that Egon Guba and others in the naturalistic field have talked about the evaluator as instrument. That metaphor fits well for the classroom, if we can think about the teacher as an evaluation instrument, and we might think quite differently about some of our technical concepts.

Continuous improvement by each student as a standard. There is some research out now that if you use a cut score kind of standard, you do great violence to the learning of kids at both ends of the standard. As an alternative to that, there are recommendations that we ought to look at continuous progress for every child for the standard. I think it would be a healthy contribution if we could reflect that some way in these standards.

Dr. Stufflebeam also stated that the example sent to the panel of writers was a fatal flaw, in that it was not a good representation of what was needed for student evaluation standards. Dr. Stufflebeam suggested that this step be revised and repeated to get to the desired final result.

Also, the issue of audience, locus, and focus that has been discussed needs to have a consensus. Without that, the likelihood is small of this being a successful effort. In addition, the absent members have not been informed or had input into the discussion. He urged the group to back up and make these fundamental decisions.

Dr. Rogers added that he is confused about where the committee stands on the issues. He referred to the two previous validation panel reports, that refer to a common failure to solicit views of all of the representatives. Teachers in the classroom, principals, students, and parents, should be included. One suggestion is to use master's students who are working as teachers,

going to school at night. Also, there are principals who are working on a degree. These are accessible people that could be contacted quickly. Dr. Rogers strongly urged the committee to obtain views of all stakeholder groups.

Take some of the illustrative cases, make vignettes and have people react to them in that case, i.e., what is wrong with this, if it is a negative one, or what is the best practice? The Validation Panel encourages the staff, executive committee, and the whole Joint Committee to really look ways of accessing all of the various groups after you have settled the issue of locus.

Comments and questions were invited.

Dr. Sanders thanked the Validation Panel and asked for reactions to their report.

An issue was raised as to whether the Validation Panel thought the standards were a complete, exhaustive, nonoverlapping set of statements. Dr. Stufflebeam replied that he had a list of topics that might be considered: (1) legal liability; (2) standard relating to scoring, combining, rating, ranking, and grading; (3) a standard that clarifies the teacher model involved in the classroom, i.e., teachers use different models—growth model, mastery, curriculum embedded evaluation; test score standards, etc., and there was not a standard that dealt clearly with that issue; (4) scoring and grading of qualitative information, that may be dealt with in accuracy; (5) something on the notion of incentives, e.g., teaching the test to raise standardized test scores. This is going on around the country, and should be addressed here. Those are some examples of what might be missing standards.

Stakeholder identification now is contained in the political viability standard. If that were moved up front to Stakeholder identification and engagement, it would fit better, and you would have a better start. But there are still political issues that need to be addressed, so that may have to be rewritten.

Dr. Stufflebeam referred to a recent article by Michael Scriven wherein he suggested changing the labeling of the feasibility standards and call them “viability” standards, i.e., Procedural Viability; Political Viability, Economic Viability; and Legal Viability. So that would be worth your time to look again at the list.

Dr. Sanders indicated items in the notebooks and items sent to the panel of writers is only a first attempt at content. But like any good work, it is open to restructuring, redefinition, etc. He sees this as first step in the process and the Joint Committee is not locked into anything at this point.

Dr. Stufflebeam indicated that there is a tremendous burden on the Chair of the Joint Committee to pull the group together and use resources as best can be done. Your agenda says you are about to elect a new Chairperson. That new Chairperson is going to pick up that burden. The two grants from Kellogg are linked to Dr. Sanders as principle investigator. The Chair is going to have a tough job going on without authority and responsibility over the two grants. That should be addressed while there is time.

Return to the Student Evaluation Standards Project

Don Yarborough requested a prioritization of discussion items. Discussion was held and the following was agreed upon:

1. Discussion of the Chair and members, and the broader issue of leadership
2. Commitment to both projects
3. Locus of audience
4. Examination of the sufficiency of the existing standards, whether to combine or add some.
5. Gaining active participation from all members

Dr. Sanders opened discussion on election of the new chair. He indicated that in the past, the Chair, Dr. Sanders, has been the principal investigator. When a new Chair is elected, those roles will be separate, and concern has been expressed about how that will affect the Joint Committee. Dr. Sanders indicated that he would suggest submission of a request to Kellogg to transfer the principal investigator position to the newly elected chair, and give the rationale for that.

Dr. Beatty read the list of duties of the Chair from the Operating Procedures and Principles of the Joint Committee. It was noted that historically (past 25 years) the Chair has been from The Evaluation Center at WMU, as it has been willing to house and support The Joint Committee.

Election of Chair and Executive Committee

The floor was opened for nominations for a new Chair of the Joint Committee. Dr. Yarbrough nominated Dr. Arlen Gullickson. The nomination was seconded by Dr. Beatty. No other nominations were made and the nominations were closed. A ballot was held and Dr. Gullickson was elected as Chair.

The floor was then opened for nominations for the open Executive Committee position and Dr. Wilson was nominated by Dr. Yarbrough. The nomination was seconded by Dr. Beatty. No other nominations were made and the nominations were closed. A ballot was held and Dr. Wilson was elected to the Executive Committee.

There was discussion of the sponsoring organizations who have not been participating and reasons presented to the Joint Committee. Dr. Sanders indicated there were adequate explanations for four of the organizations that did not send a representative this year. Three organizations have chosen not to participate, and the new Chair should negotiate with them to get them reinvented.

Thanks and appreciation were given to Dr. Sanders for all his past leadership and work as Chair.

There was discussion about working on two projects at the same time. Dr. Sanders described the remaining work for the Joint Committee on the revisions for *The Personnel Evaluation Standards* as required by the Operating Procedures and ANSI.

Options for restructuring the annual meeting were discussed.

Dr. Sanders stated that members will be asked to review both projects between meetings due to having two funded projects and taking meeting time for other discussions. Dr. Sanders suggested that both projects go forward, with communication by email, phone, or listserv. He also suggested the Committee stay with the Student Evaluation Standards issues for the afternoon.

There was discussion on the focus as specifically a K-12 school settings vs. broad focus, including other settings where students are evaluated.

There were two issues on focus in the material from the panel of writers. One is who the standards are about, and that is where there was a big split, i.e., student vs. institution. It has to be seen as a focus on the student, but the panel of writers did not view it that way. Then there is the focus of who the document is going to be distributed to. Both issues need to be addressed.

Dr. Sanders stated that the guidelines, common errors, and overview are generalizable, and do not determine the setting. The setting is determined only in the illustrative cases. He also stated that the committee may chose cases which are primarily in the K-12 setting, but he would like to see some cases from higher education and professional development where students are adults.

There was a suggestion for inclusion of two illustrative cases for each standard, one in a K-12 setting, and one in another setting, medicine, business, or military. Also, there should be a balance between positive and negative examples.

Gender of the evaluator should be varied, using equal he/she references.

Dr. Gullickson voiced a concern that by including cases from settings other than classroom teachers, the primary audience of teachers will not have the full scope of information.

Another suggestion was made to include 3 cases, 2 from classroom perspective, and 1 from an alternative setting.

It was also suggested that K-12 teachers could benefit from seeing how these standards apply to a broad range of situations, that it may not necessarily be irrelevant.

Use student to define multiple ages and educator vs. teacher or instructor.

Another suggestion was that teachers may not be enthusiastic about using the new standards when they have so many assessment guidelines that are extremely prescriptive already to deal with. Also, in the field of special education, it is a team that evaluates for placement. The team includes parents, who are the child's first educator.

A question was raised about the parameters of the project, i.e., if a broad scope is assumed, where are the limits? What is outside the scope of the student evaluation standards?

Dr. Sanders summarized that most of the committee members were agreeing that the focus should be on students in institutional settings, prekindergarten, K-12, and higher education.

Dr. Sanders explained that when the first program standards were developed, there were writers from noneducation groups such as AEA, APA, and NLPES, saying they found the standards very useful, but asked for cases that would cover things outside of schools. The Joint Committee tried to respond to that by expanding the cases that were nonschool cases. He also suggested that the Student Evaluation Standards could start at the school level as well, and wait to see if other groups find them useful.

It was also pointed out that the more narrow the focus, the more likely that the Joint Committee might replicate standards that are already in existence. Dr. Sanders disagreed, stating that these standards are principles, and only narrowed in the illustrative cases. This is a casebook for anyone who evaluates students. The distinction is that we have the “judgmental” piece that no other standards deal with. The Student Evaluation Standards also are intended to be a comprehensive set which cover prekindergarten, K-12, and higher education. No existing standards are as comprehensive, including guidelines, common errors, and illustrative cases.

Break

Bob Wilson gave the following case example:

. . . each kid takes a company and follows that company, develops a portfolio of work around that company around the course of the semester, and half-way through it, the parents come in for parents night, with 4 of these sessions going on at the same time, where the kid is telling the parent what he or she has learned during the course of the semester. The parents are given a little sheet to help them ask questions that are relevant. Sessions are supposed to last 20 minutes. Sessions frequently go on for an hour; sometimes a lot longer than that, and he makes coffee and gives it to people. I have been in on those sessions. I have seen parents cry. For some it is the longest conversation they have had with their child in a long time. There was a 21-year-old single mother, who invited her parents, even though she wasn't living at home anymore, who came to hear what she had done. They are marvelous stories. The principal came one time and was blown away because he thought David was running all of the conferences, and went back contented because no parents ever complained about the reporting that goes on in that class. There is a 100 percent turnout. At the end of the semester, each one of these students has a time allotted him, and following a rubric that has been created for how these projects must look, to present the case for their company to David in a one-on-one situation, at the end of which a mark is negotiated out of a 100 percentage, that is put on the records, and that is the end of the story. Except the other teachers think he is getting away with stuff, so they insist that he also has to supervise the exams in the gym with all the other teachers. And they think it is somehow disreputable what he does—that he is too lazy to mark during the semester.

Now there are 2 or 3 more teachers who kind of like the idea, and he is worming his way into the culture of the school. We found that you cannot change the policies and procedures of the school

by working with administration. You change it from the other way around. It gets changed, and the teachers change.

I read the effective reporting from the point of view of David, who is one of the most inspiring and effective teachers I have ever seen, and asked how would he use this in his class.

Dr. Sanders asked, "How would he show up if we were evaluating his practices? Would he come out looking well or not so good?"

Dr. Wilson came to the conclusion that he would read this and interpret it quite differently than the way many other teachers would, but still find it acceptable, because he is still concerned about the accuracy of the assessments that he does, and so on. There isn't any mention in the front about portfolio assessment or student conference. And these are, if you don't know this, these are the most widespread innovations in reporting assessment that are going on in our country. I think he would find that okay.

When he got to the guidelines, I think they are heavily influenced by the fact that the teacher who is doing all of it are the educators who are doing all of it. And the student, which a theme we have had from a previous discussion, doesn't seem to be placed . . . But with that change, I think it flies fairly well. Once again, the cases are not that. And I wonder if a case like David, being one of the cases, would also be an educative function that we would build into these standards, and probably by the time you get published, will resonate with people who are also trying to . . .

Dr. Sanders commented that what some of us have been saying, is that we need to expand people's thinking through the use of the standards as well as just provide standards.

Dr. Sanders indicated he would like to use this case and Bob Wilson agreed to send it to him.

It was also noted that guidelines on portfolio reporting should be covered. The students have to have adequate information about what to put in their portfolio. They found that white students tended to put the things they got the highest marks on in it, and minority kids pick the things that have been the most difficult for them to do, but they felt a sense of accomplishment about it. Of course, when you look at, evaluate the portfolio, there was a big difference.

Dr. Wilson noted that in the system he evaluated, they didn't do any selection of what was in the portfolio. They just organized it by time, and the parents came in and the average time was 2 hours, and one of them gave 8 hours, across 4 times, the grandparents came in.

Dr. Gullickson noted one of the things he would like to see in Effective Reporting is the care and attention to matters of accuracy, for example, are commensurate with the significance of the judgment. He thinks the typical teacher, for example, would spend relatively little time writing comments on homework pieces, when the teacher will follow up with comments to the class illuminating that area and use things as a marker. They don't worry about being accurate and precise on a daily basis in terms of homework, and you would worry more about summing things

up for written comment to give to the student and parents at the end of the term. So the size of the decision, the implications that go with that decision, ought to be commensurate to the amount of attention given to accuracy in that area.

He also had a comment on guidelines. Point E says: "Indicate to report users how the evaluation information may be useful to them." To bring it a little closer to home, Dr. Gullickson might say something like, "Explain to students the intended purpose of the evaluation statements, conclusions, and how the results will be used." In order for this to be an effective report, the student needs to know how this information should be used by them. Teachers are often very cryptic or terse in their comments.

Cross-references are necessary to other standards.

Dr. Yarbrough questioned the use of the title: Guidelines and Common Errors, particularly the Common Errors portion. It is useful to include errors that are not common, but if they happen they are tragic or catastrophic. Perhaps it could be a list of dos and don'ts. This could be considered as the standard is revised.

Dr. Gullickson said he would like to include one case example on standardized testing, and no more. There are many other contexts to include.

Dr. Fager said she would like to look through teacher training materials to see what prescriptive kinds of things they say.

Dr. Yarbrough asked for a decision on whether these standards are complete, exhaustive, nonoverlapping enough to be used as the basis for the final product. He asked to keep in mind the task of expanding or compacting the standards as the committee works, so this does not get lost.

Dr. Sanders reviewed the Validation Panel suggestions of items that could be standards that are not included presently:

- Stakeholder Identification
- Legal Liability
- Change of Feasibility to Viability
- Scoring, Combining, Ranking, Rating, and Grading

Reviewers should be directed to consider comprehensiveness and internal consistency of the standards.

General Meeting Report

The general meeting report copies were distributed and reviewed.

Revisions were made and then the report was approved. This report is for use by the Sponsoring Organizations in making reports to Boards and in newsletters.

Dr. Sanders asked each member to report on plans for Joint Committee issues at an annual meeting or other activities that relate to the work of the Joint Committee.

Dr. Yarborough reported that Drs. Fager, Sanders, and he will present at NCME at the invitation of the President of NCME. They will address the history of Joint Committee work, the rationale and the need for growing out of the literature the need for a set of student evaluation standards. *The personnel and program evaluation standards* will be mentioned but there will not be a lot of time spent on those. The goal is to allow the membership opportunity to approach us and voice their concerns primarily about overlap with or competition with the Joint Standards on Testing.

Dr. Wilson reported that he would be attending a retirement celebration where he will talk with others about the work of the Joint Committee. He will also attend the Board of Directors Annual Meeting and report on the work of the Joint Committee.

Dr. Hollenbeck stated that he would call the editor of the American School Board Journal and see if he could get some publicity in that. He also agreed that an article on the other standards would be welcome. He also indicated he would be interested in attending the National Evaluation Institute in Traverse City in July 1999 to help present there.

Dr. Moore suggested the national convention of principals (NASSP) in New Orleans in the winter of 1999 would be a good opportunity to bring a copy of *The Personnel Evaluation Standards* and *The Program Evaluation Standards*. Professors of Secondary School committee that he serves on will hold an open forum, and that may be a place to reacquaint the professors that are there and secondary principals with these two standards books. Further, at the spring meeting in Washington may be a time to share a report on the progress being made on the Student Evaluation Standards and inquire how these standards may be useful to them in their work, and what they perceive to be the value of the standards. This may help the Joint Committee in drafting future publicity releases.

He also reported that the leadership of NASSP is rapidly changing; the executive director is retiring and many directly below him are leaving. Also, the President-Elect is from Michigan, Lynn Babcock from Livonia. He believes that the national organization does not have much awareness of what the Joint Committee is doing, and he is looking for new ways to raise awareness.

Dr. Ekstrom reported that APA will be meeting during the spring of 1999, and she will send them a report. It would be August 2000, in Washington, before anything could be on the program.

Dr. Beatty reported that she will report to the deputy director of ASCD, and she will send names for the review panel. She will work with Diane Berreth to promote the work of the Joint

Committee. Also, she will be working with a group in New England for a local ASCD conference in Boston in December, where she will do networking and discussion.

Dr. Sanders asked all members to get names of reviewers and validation panel member names to him in the next couple of weeks or as soon as possible.

Dr. Sanders proposed next steps on the two projects.

Members will review the proposed revisions on *The Personnel Evaluation Standards* within the next week to ten days, and send reactions, suggestions, to Dr. Zeller by email. She and Dr. Sanders will then move ahead with what they have and try to get a manuscript ready to mail out to everyone. Where there are cases that need to be written, or things to fill in, they will solicit additional writing assistance.

On the Student Evaluation Standards, the suggestions from the meeting will be typed up, Dr. Sanders will work on the introduction. He will do some editing, and send to all members the compilation for comments and editing. He asked members to review and return those within a couple weeks. He also confirmed that the first draft would be sent to a small group of reviewers for comment before sending the first draft out for review by the larger national and international review panels sometime after the first of the year.

Also, he confirmed that teachers will not be added to the validation panel, but will be included as reviewers.

Mary Ramlow requested updated directory information.

A motion to adjourn was made, seconded, and carried.

Dr. Sanders adjourned the meeting.