



aea
EUROPE

**4th Annual Conference
Lyon 2003**

Assessment Challenges
for Democratic Societies

In association with the
Agence Nationale de Lutte contre l'Illettrisme (ANLCI).



Assessment Challenges for Democratic Societies

Contents

The Fourth Annual AEA-Europe Conference	4
Programme of Conference	5
Keynote session - Abstracts and Speaker Profiles	8
Discussion Groups - Abstracts	18
About AEA-Europe	21
Council Members	22

In association with the
Agence Nationale de Lutte contre l'Illettrisme (ANLCI).

The Fourth Annual AEA-Europe Conference

Assessment Challenges for Democratic Societies

6 – 8 November 2003

CONFERENCE VENUE

6 November: ENS-Sciences, Lyon

7 – 8 November: ENS-Lettres et Sciences Humaines, Lyon

INTRODUCTION

The Association for Educational assessment- Europe arranges a conference each year to further its aims of fostering communication in the area of educational assessment. The three previous conferences have been held in Prague (2000), Krakow (2001) and Frascati (2002).

In 2003, the conference is being held in Lyon, France. Lyon is the home of the Agence Nationale de Lutte contre l'Illettrisme (ANLCI), an organisation set up by the French government in order to coordinate actions for combating illiteracy. Because the theme of our conference this year is Assessment Challenges for Democratic Societies, the Council thought it appropriate and interesting to accept the invitation of the ANLCI to hold our conference in harmony with theirs by sharing a common day for both on 6th November.

Our first day will be held in the ENS-Sciences, with a programme jointly decided by AEA-Europe and the ANLCI. That is why there will be representatives of the French government and a larger audience on this particular day. On the other two days our conference will be held at the ENS-Lettres et Sciences Humaines and our proceedings will take the usual form. In addition, for those interested, the ANLCI has organised a particular session on 5th November in order to illustrate some of the measures that France and other countries have adopted to promote literacy.

The AEA-Europe executive committee decided to modify the usual structure of our conference for two reasons. On the one hand, we decided to enrich our offering by opening up to a complex reality such as the one represented by the ANLCI. On the other, we thought it useful to present AEA-Europe and its activities to an audience of French specialists.

We would like to thank the ANLCI and the City of Lyon for their interest in our conference. In particular, the ANLCI has offered the premises and equipment for the conference and discussion groups, and has also made a contribution to the gala dinner. They have also sponsored the accommodation of some delegates. The City of Lyon has offered a splendid hall for the dinner on 6th November and also a free guided tour of the Old Town in the afternoon of 8th November.

AEA-Europe is also grateful to Stephen Austin Ltd and to DRS Data and Research Services PLC for providing conference bags and folders.

Thursday November 6

Common day AEA-Europe & ANLCI

8.00 - 9.30	Registration and coffee	
9.30 - 10.15	Welcome	Emma Nardi, AEA-Europe Marie-Thérèse Geffroy, ANLCI
10.15 - 10.45	Opening address	Luc Ferry, Minister of Education, France
10.45 - 11.15	Introduction to Key Competencies	Dominique Simone Rychen, DeSeCo Secretariat, Swiss Federal Statistical Office, Switzerland
11.15-12.00	The "Information and Everyday Life" survey : results and prospects	France Guérin-Pace, Institut National d'Etudes Démographiques (INED), France Fabrice Murat, National Institute for Statistics and Economic Studies (INSEE), France
12.00 - 13.30	Lunch	
13.30 - 14.15	Exploring Relationships Among Reading Components Across Literacy Skill Distributions	Irwin Kirsch, Educational Testing Service (ETS), USA Kentaro Yamamoto, Educational Testing Service (ETS), USA John Strucker, National Center for the Study of Adult Learning & Literacy (NCSALL), USA
14.15 - 15.00	The magnitude of the "low literacy" problem in OECD countries and its social and economic consequences: Empirical Evidence from PISA and IALS	Scott Murray, Statistics Canada, Canada
15.00 - 15.30	Coffee break	
15.30 - 16.15	Measuring numeracy skills: do we know how to do it? Do we even know what 'numeracy' means?	Alison Wolf, National Research and Development Centre for Adult Literacy and Numeracy, England
16.15 - 17.00	Presentation of LAMP (Literacy assessment and Monitoring Programme)	Bénédicte Terryn, the UNESCO Institute for Statistics, Canada
17.00 - 17.15	Low levels of literacy: a major concern in the debate on the future of schools in France	Claude Thélot, France
19.30 -	Dinner at the City Hall	

Friday November 7

9.00 - 10.30 Benchmarking the Quality of Education Sally Thomas, Graduate School of Education,
University of Bristol, UK

Panel: Alejandro Tiana (Spain) and
Gerard Bonnet (Ministry of Education, France)

10.30 - 11.00 Coffee break

11.00 - 12.45 Using Portfolios for Educational
Discussion Assessment: Two Case Studies
groups 1-3

Val Klenowski, Institute of Education, England

Grace Grima and Debbie Chetcuti, University
of Malta

Target Setting within Global
Education Programmes

Harvey Goldstein, Institute of Education,
England

Media Coverage of Assessment Issues

Roger Murphy and Paul Warmington,
University of Nottingham, England

13.00 - 14.00 Lunch

14.00 - 15.00 Examinations reform in Russia -
Unified State Examinations

Victor N. Simkin, The Russian Ministry of
Education, Institute for General Education, Russia

15.00 - 15.30 Coffee break

15.30 - 16.30 Reading achievement of primary
school students in thirty five countries:
Reflections on the results of the
progress in the International Reading
Literacy Study - PIRLS

Gabriella Pavan De Gregorio, Istituto Nazionale
per la Valutazione del Sistema dell'Istruzione
(INVALSI), ITALY

16.30 - 17.30 AEA-Europe Business meeting

Saturday November 8

9.30 - 11.00 Discussion groups 4-6	Diagnostic Assessment in Early Years	Marian Sainsbury, the National Foundation for Educational Research (NFER), England
	The Use of Sophisticated Item Formats in Computer-Based Assessment	Andrew Boyle and Dougal Hutchison, the National Foundation for Educational Research (NFER), England
	Monitoring Learning and Achievement: linking research, policy and practice. The Scottish Experience	Carolyn Hutchinson et al, Scottish Executive Education Department (SEED), Scotland

11.00 - 11.30 Coffee break

11.30 - 12.30 Trends in the predictive validity of high school grades and aptitude tests. Jan-Eric Gustafsson, Department of Education, Göteborg University, Sweden

12.30 - 13.00 Concluding remarks and farewell

13.30 - 14.00 Lunch

Afternoon Tour (optional)

Keynote Sessions – Abstracts

THURSDAY, NOVEMBER 6

**Thursday
10.45-11.15**

Introduction to Key Competencies: An overarching conceptual framework for competencies that matter - Some Results from the OECD Project DeSeCo

Dominique Simone Rychen, (DeSeCo Secretariat, Swiss Federal Statistical Office, Switzerland)

Adults of the 21st century are expected to fulfil many different roles - as students, learners, workers, citizens, family members, consumers. They must act and navigate in and across various contexts and meet multiple complex challenges in a responsible manner.

What competencies - apart from reading, writing, and computing - are necessary for individuals to lead an overall successful life and for society to face the challenges of the present and the future? With the myriad of conceivable individual and social differences in and across countries, is there any ground for defining a universally applicable set of key competencies? And if so based on what normative, definitional and conceptual criteria can a limited set of key competencies be identified? Such questions were at the heart of the OECD Project DeSeCo *Definition and Selection of Competencies: Theoretical and Conceptual Foundations* launched at the end of 1997 and carried out under the leadership of Switzerland.

Based upon a review of existing work in the area of competence and on inputs of experts from a variety of backgrounds, the concept of key competence as defined in DeSeCo relies on three general criteria. Key competencies

- contribute to highly valued outcomes at the individual and societal level;
- are instrumental for meeting important, complex demands and challenges in a wide spectrum of contexts; and
- are important for all individuals.

The three categories of key competencies - *interacting in socially heterogeneous groups, acting autonomously, and using tools interactively* - grounded in a holistic model of competence, constitute a core element of DeSeCo's overarching conceptual frame of reference for developing and assessing key competencies in an international setting.

The DeSeCo framework sets an overarching conceptual context for international assessments, such as the Programme for International Student Assessment (PISA), the International Adult Literacy Survey (IALS), and the Adult Literacy and Life Skills Survey (ALL) and could provide a reference point for the development of education and training program for all stages of lifelong learning.

Dominique Simone Rychen

Dominique Simone Rychen is a Senior Program Officer at the Swiss Federal Statistical Office and the Program Manager of the OECD Project, Definition and Selection of Competencies: Theoretical and Conceptual Foundations (DeSeCo; www.deseco.admin.ch). She has been responsible for coordinating the various research activities within DeSeCo and for reporting the findings to the OECD.

She is co-editor of *Defining and Selecting Key Competencies* (Hogrefe & Huber, 2001) and DeSeCo's final report *Key Competencies for a Successful Life and a Well-Functioning Society* (Hogrefe & Huber, 2003), and the author of various articles related to the topic of key competencies.

Previously, she worked on indicator development related to continuing education, labour market, workplace, and informal learning in particular, and co-authored two publications related to continuing education. She holds a master's degree in political sciences at the University of Lausanne, Switzerland.

**Thursday
11.15-12.00**

The “Information and Everyday Life” survey (INSEE): initial results and prospects

Fabrice Murat (INSEE) & France Guérin-Pace (INED)

As the results of the International Adult Literacy Survey (IALS) show, the evaluation of the writing competence of an adult population raises quite a few problems. In conducting the “Information and Daily Life” (IVQ) study, the accent was on maintaining the survey respondents' motivation as regards a survey protocol that took account of the specific nature of the population and which was based on various evaluation supports. In particular, a preliminary phase of orientation allowed providing the survey respondents with a quick set of exercises adapted to their level. The evaluation sheet measured the weakest competencies and was created by the PsyEF laboratory (University of Lyon 2) in cooperation with INSEE, INED and ANLCL.

In November 2002, about 2000 people took part in the IVQ survey. The analyses conducted on the survey data try to overcome the simple hierarchy of levels of competence, showing the heterogeneous nature of the results according to the nature of the tasks carried out. It is also possible to work out differentiated profiles of writing competence. These analyses also highlight the links with difficulties in reading and the individual processes involved, as well as their impact on the conduct of daily life.

Fabrice Murat

Fabrice Murat is Head of the FQP (*Formation et Qualification Professionnelle*) and IVQ (*Information et Vie Quotidienne*) surveys at the National Institute for Statistics and Economic Studies (INSEE). Previously, he was Head of the INES model for the simulation of fiscal and social policies, also at INSEE.

He was educated at Ecole Nationale de la Statistique et de l'Administration Economique and subsequently completed a Higher studies diploma in Sociology on “*Les différents usages des évaluations de compétences*”.

His recent publications include

- *Les compétences des élèves à l'entrée en cours préparatoire Études réalisées à partir du panel d'écoliers recruté en 1997*, les dossiers d'Education & Formations n. 132, Ministère de l'Education nationale, de l'enseignement supérieur et de la recherche, Direction de l'évaluation et de la prospective, 2002.
- *La place du projet professionnel dans les inégalités de réussite scolaire à 15 ans, in France, Portrait Social*, INSEE, 2002.
- *La mesure des compétences: les logiques contradictoires des évaluations internationales in Actes des journées de la méthodologie statistique: 4 et 5 décembre 2000*, INSEE, 2002
- *Motivation et performances scolaires: les filles creusent l'écart*, Insee Première n. 886, INSEE, 2003

France Guérin-Pace

France Guérin-Pace is a Research Appointee at the Institut National d'Etudes Démographiques, Paris, having previously been a lecturer in statistics at Universities of Paris 1 and Paris 7. She earlier undertook a higher studies diploma in Statistics, Higher University Institute of Paris, University of Paris 6 and then a Doctorate in Geography on "*La dynamique d'un système de peuplement: évolution de la population des villes françaises de 1831 à 1982*", at the University of Paris 7.

Her recent publications include:

- BESSE J. M., GUÉRIN-PACE F., 2002: Une évaluation des compétences sur l'écrit: l'enquête "Information et vie quotidienne", *Economie et Humanisme*, n. 363, pp. 17-21..
- GUÉRIN-PACE F., BLUM A., 1999: L'illusion comparative - Les logiques d'élaboration et d'utilisation d'une enquête internationale sur l'illettrisme, *Population*, 54, pp.271-302.
- BLUM A. GUÉRIN-PACE F., 2000: Weaknesses and defects of IALS, in Siobhán Carey (ed), *Measuring Adult Literacy - The International Adult Literacy Survey in the European Context*, Office for National Statistics, London, pp. 68-98.

Thursday
13.30-14.15

Exploring Relationships Among Reading Components Across Literacy Skill Distributions

Irwin S. Kirsch, Kentaro Yamamoto (ETS, USA) and John Strucker (NCSALL, USA)

For some, the twenty-first century marks the passing of the age of the industrial worker to the age of knowledge worker. This passage brings with it an increased emphasis on the development of new knowledge and investments in human capital. This is because what one knows and can do (literacy skill) is strongly associated with social, educational and economic outcomes. Although education increases one's literacy skills, these skills play a critical role in influencing educational success. These skills influence what courses students take, whether they graduate from high school, whether they attend college, what they choose to study if they attend college, and the types of degrees they obtain. These skills are also associated with success in the changing labour markets. Changes in the structure of jobs by occupation and industry and in accompanying job duties have increased the demands for better-educated and more literate workers with stronger communication and critical thinking skills. The cumulative impact of these changes in the job market has increased the economic premiums associated with formal schooling, literacy proficiencies and technical skills.

Despite the growing importance of acquiring new knowledge and investments in human capital, data from the 1992 National Adult Literacy Survey (NALS) and the 1994 International Adult Literacy Survey (IALS) showed that some 20% of American adults performed at the lowest of five literacy levels on the NALS and IALS scales. In 2003, the United States is participating in a second international adult literacy assessment, the Adult Literacy and Lifeskills survey (ALL), which will provide data on changes in literacy since 1994 among adults in America as well as other participating countries. At the same time, the U.S. will also carry out a study of adult education participants (Adult Education and Literacy study, AEL) in order to report on the literacy levels of learners enrolled in adult education programs. Collectively these studies will provide policy makers and others with important information for understanding the distribution of literacy skills in adult populations. Yet, little is known about the underlying component skills especially among lower performing adults or the relationships among these skills across the ability distribution. This knowledge is thought to be important for the development of better diagnostic measures as well as for improving instruction. By using assessment

tasks derived from the international Adult Literacy and Lifeskills survey (ALL) as well as a battery of reading component skill tests, the Components of Reading Study will describe how a set of component skills relate to literacy performance.

Based on previous reading research, a variety of measures were selected and administered along with measures of prose and document literacy used in IALS and ALL in face-to-face interviews. Two of the questions being addressed in this paper are: 1) How do these component skills relate to each other and to literacy performance; and 2) Are the relationships between these component skills and literacy constant across the ability distribution. These instruments were administered to over 900 adults between the ages of 16 and 65 who were enrolled in adult education programs in the United States at the time of the survey.

This presentation will focus on describing the overall study and the measures that were used along with the results that were obtained. Analyses will be presented that show the relationships among the component skills and with the measures of prose and document literacy. Implications of the findings for future test development and instructional practice will also be discussed.

Irwin S. Kirsch

Irwin S. Kirsch is a Senior Research Director and Director of the Center for Global Assessment at Educational Testing Service. He earned his Ph.D. in Educational Measurement, Reading / Literacy from the University of Delaware in 1982. Since joining ETS in 1984, he has directed a number of large-scale assessments including the first International Adult Literacy Survey. Other large-scale assessments that he has directed include the National Adult Literacy Survey, the DOL Workplace Literacy Assessment, and the NAEP Young Adult Literacy Survey. Among his activities, Mr. Kirsch chairs the Reading Subject Matter Expert Group (SMEG) for an international survey of 15-year olds assessing their preparedness for adult life, known as the OECD Programme for International Student Assessment (PISA). In addition to project responsibilities, Mr. Kirsch serves as a member of the national policy group on setting standards for adult education; a multi-year project supported by the National Institute for Literacy. Among his publications is a monthly column on reading/literacy he co-authored for the *Journal of Reading* from 1989 to 1993. His research interests include the psychology of literacy, issues of comparability and interpretability in large-scale assessments, and using technology to link learning and assessment.

Thursday
14.15-15.00

The magnitude of the "low" literacy problem in OECD countries and its social and economic consequences: Empirical evidence from PISA and IALS

T. Scott Murray (Statistics Canada, Canada)

The governments of Canada and the US have, over the past 15 years, invested considerable resources in the development of methods to assess the literacy and numeracy skills of heterogeneous populations of adults. Embodied in the International Adult Literacy Survey (IALS) and the Adult Literacy and Life Skills Survey (ALL) these methods have yielded a wealth of comparative data on the distribution of literacy, the social and economic outcomes that are associated with objectively determined proficiency levels and the factors that influence, either positively or negatively, the level and distribution of skill in some 35 economically developed nations. While these surveys provide information on the number of individuals with low skills and their demographic characteristics, they reveal little about what these individuals can and can't do with the printed word. The ALL Level 1 study was developed to profile the component reading skills of low skilled individuals with a view to understanding their relationship to the acquisition of reading fluency and to identify population sub-groups whose component skills profile warrant particular pedagogical approaches.

The objective of this paper is threefold. First, it presents an overview of the design of the IALS and ALL studies with a focus on how the design responds to the need to generate valid, reliable, comparable and interpretable results. Second, the paper provides an overview of the evidence related to validity, reliability and interpretability, including what is known about the relative importance of various sources of measurement error.

The paper concludes with an overview of how the Level 1 study will compliment the base ALL assessment, including what is currently known about the measurement properties of the component skill measures.

Scott Murray

Scott Murray currently holds the post of Director General, Social and Institutional Statistics. Prior to this appointment to this position in 1999, he spent roughly 23 years in the Special Surveys Division at Statistics Canada including a period of 5 years as Director. Mr. Murray has specialized in the design and conduct of large-scale ad hoc surveys to meet emerging public policy issues. His own work has included studies of volunteer activities, child care usage, longitudinal labour market activity and international comparative work in the area of the assessment of adult skill and participation in adult education and training. Currently International Study Director for the International Adult Literacy Survey (IALS) and the Adult Literacy and Life Skills Survey (ALL). Mr. Murray holds an Honors BA in Business Administration from the University of Western Ontario and is overly fond of claret.

**Thursday
15.30-16.15**

Measuring Numeracy Skills: do we know how to do it? Do we even know what "numeracy" means?

Alison Wolf (National Research and Development Centre for Adult Literacy and Numeracy, England)

'Numeracy' has become a popular term in recent years, but its meaning is far more disputed than is the case with literacy, and measurement problems are consequently even greater. Many critics are concerned that a policy focus on numeracy ignores important messages from the labour market about what mathematical skills are really needed in contemporary society; and defining differences between 'adult' numeracy and the lower levels of school-based mathematics curricula has proven very hard; . Close correlations between, for example, 'quantitative' and 'prose' literacies in international surveys may say as much about the conventions of item writing as about the distribution of skills in the adult population. This address will consider competing notions of 'numeracy' and relate them to emerging research on the labour market, to developments and controversies in school mathematics curricula, and to the challenges of developing tests of adult skills within, as well as across, societies.

Alison Wolf

Alison Wolf is now Professor of Management and Professional Development at King's College, London. She is also a member of the National Research and Development Centre for Adult Literacy and Numeracy, England (NRDC) and Visiting Professorial Fellow at the Institute of Education, University of London.

Alison's research focuses on the relationship between education and training, and the skill requirements of the labour market, with particular focus on basic skills and non-specialist mathematics. She is currently Programme Director of *Extending 'Skills for Life': identifying effective basic skills provision in the workplace*, a five-year longitudinal study tracking learners in approximately 50 workplaces and funded by the Economic & Social Research Council (UK) and the NRDC. She is also a programme director within the Centre for the Economics of Education (responsible for research relating to demand for education and skills). She is the Executive Editor of *Assessment in Education*, and a

Member of the Editorial Board of *Politiques d'Éducation et de Formation* (Paris). Her most recent book is *Does Education Matter?* Penguin 2002 and she is a regular columnist for the Times Higher Education Supplement.

**Thursday
16.15-17.00**

Presentation of LAMP (Literacy Assessment and Monitoring Programme)

Bénédicte Terryn (the UNESCO Institute for Statistics, Canada)

UNESCO Institute for Statistics (UIS) compiles estimates of literacy for the majority of countries in the world based on countries' data. However, currently, most national literacy statistics are based principally on a mix of self-declarations and of educational attainment proxies, so that estimates do not reflect the true level of literacy in the world's population. Given the importance of the UN Literacy Decade, and the importance of literacy as an indicator for adult learning and as the outcome of basic education in the Millennium Development Goals and the Education For All goals, the UIS decided to prioritize development in this area. The Literacy Assessment and Monitoring Programme (LAMP) seeks to improve both what and how literacy is measured, and to provide countries with the tools for collecting adequate literacy statistics, in order to support evidence-based policy-making at national and international level, and to inform the design of improved literacy programmes in UNESCO's Member States.

Bénédicte Terryn

Bénédicte Terryn studied mathematics in Louvain-la-Neuve (Belgium) and Warwick University (UK). She then moved into statistics and worked for four years for the UK Government Statistical Service: successively on the Labour Force Survey, the methodology review of the Retail Price Index, and managing land use statistics at the then Department of the Environment Transport and the Regions. Since September 2002 she has been working at the UNESCO Institute for Statistics, in Montreal (Canada) on literacy and in particular the development of the Literacy Assessment and Monitoring Programme (LAMP). In parallel she is completing an MSc in Official Statistics at the University of Southampton (UK)

The UNESCO Institute for Statistics is the statistical branch of the United Nations Organization for Education, Science and Culture (UNESCO). It was established in July 1999 to meet the growing needs of UNESCO Member States and the international community for a wider range of policy-relevant, timely, and reliable statistics in the fields of education, science and technology, culture and communication.

**Thursday
17.00-17.15**

Low levels of literacy: a major concern in the debate on the future of schools in France

Claude Thélot

Claude Thélot is an Advisor at the Court of Auditors and president of the commission dealing with the French national debate on the future of the school system, Claude Thélot is a polytechnic engineer who also graduated at the *École Nationale de la Statistique et de l'Administration Économique* (ENSAE). Over the years he has had many responsibilities in the French education system, notably, between 1990 and 1997, as director of evaluation and prospects at the French Ministry of Education, and also as president of the High Council of School Evaluation, from January 2001 to February 2003. His many publications include: *L'évaluation du système éducatif* (1994), *Le travail en France, 1800-2000* in cooperation with O. Marchand (1997), *Réussir l'école* with Philippe Joutard (1999), *Les écrivains racontent l'école* (2001) and *L'origine des génies* (2003). Claude Thélot is also an associate professor of economics at the University of Paris V - René Descartes.

FRIDAY, NOVEMBER 7**Friday**
09.00-10.30**Benchmarking the Quality of Education***Sally Thomas (Graduate School of Education, University of Bristol, UK)**Panel: Gerard Bonnet (Ministry of Education, France) and Alejandro Tiana, (Spain)*

This presentation will explore the main findings from the final (unpublished) report of an EU Socrates funded project: Benchmarking the Quality of Education (December 2002). This project addressed a number of fundamental issues in educational assessment related to international comparative data, including how indicator data is used and how its usefulness for policy makers and practitioners can be enhanced. These issues we think will be of considerable interest to members of AEA-Europe. The aim is that this presentation will stimulate a timely and important discussion of educational indicators, how indicator methodologies can be improved and other related topics.

Sally Thomas

Sally Thomas, BSc, PhD, AFBPsS, FRSA, is Reader in Education at the Graduate School of Education, University of Bristol - currently one of only two education departments in the UK with the maximum rating for excellence in research. Until December 2000 she was an Associate Director and founding member of the International School Effectiveness and Improvement Centre at the London Institute of Education. Her main publications are on the topic of school effectiveness. In particular she has carried out extensive research studies on 'value added measures of school effectiveness' and the application of these measures for different, or overlapping, purposes including: school improvement, school evaluation and self-evaluation, international indicators and academic knowledge-base research. Recent publications include a co-authored book 'Educational Evaluation And Monitoring; A Systemic Approach' to be published by SWETS in the autumn of 2003.

G rard Bonnet

After an academic training at the *Ecole normale sup rieure de Saint-Cloud* as a linguist specialising in English, he started his career as a lecturer in the French Department of the Queen's University of Belfast (1972 -1978), teaching and doing research in French and English phonetics.

He then joined the French Foreign Service on secondment. His first posting was as director of the French Cultural Centre in Cairo (1978-1982) and his second and final posting was to the French Embassy in New Delhi where he was in charge of co-operation in education for another four years (1982-1986).

He returned to the French Ministry of Education in Paris in 1986 where he has held posts in international relations and the implementation of the reform of initial teacher training. He is currently in charge of European and international relations in the Department of Assessment and Forecasting (*direction de l' valuation et e la prospective - DEP*) of the French Ministry of Education. His role is to promote and organise international co-operation in the field of the quality of education systems. In particular he is closely involved with quality evaluation in the European Union.

He is also a Visiting Fellow at the Institute of Education, University of London, with the Bedford Group for Lifecourse and Statistical Studies.

Alejandro Tiana

Alejandro Tiana is Professor of History of Education and Comparative Education at the Spanish Distance Teaching University (UNED, Spain). He is currently Director General for Innovation and Development of the Ibero-American States Organization for Education, Science and Culture (OEI) and Chair of the International Association for the Evaluation of Educational Achievement (IEA). Among his 17 books and 100 papers on issues related to comparative education and the history of education systems, some may be mentioned: "Systemic Reform: National Policies on Curriculum and Assessment - The Spanish Experience" (in Black, P. and Michel, A., eds., *Learning From Pupil Assessment: International Comparisons*, Los Angeles, CSE - UCLA, 1998, 79-97), "*Usages et mésusages des approches comparatives dans un cadre évaluatif*" (in Sirota, R., dir., *Autour du comparatisme en éducation*, Paris, P.U.F., 2001, 77-93), "*Le monde comme laboratoire éducatif*" (in *Politiques d'éducation et de formation. Analyses et comparaisons internationales*, 2001/3, 2001, 47-57), "*Are Our Young People Prepared?*" (in *Prospects*, XXXII/1, 2002, 39-50).

Friday
10.45-11.15

Examinations Reform in Russia - Unified State Examinations

Victor N. Simkin (The Russian Ministry of Education, Institute for General Education, Russia)

The experiment with the introduction of the Unified State Examination (USE) has been underway for more than three years. The word 'Unified' in the title means that it comprises both school leaving and university entrance examinations and is set in nine different subjects. Before USE was introduced, the school leaving exams had been administered at schools and internally assessed. Each university had set its own entrance examinations. USE is meant to provide an external assessment of the school leavers' achievements.

The main purposes of the introduction of USE were the following:

- to increase accessibility to higher education;
- to form an objective system for the assessment of school leavers' achievements and provide equal opportunities for them in entering higher educational institutions;
- to ensure continuity between secondary and tertiary education;
- to equalize the state certificates of secondary education;
- to establish state control of the quality of general education.

This year almost half of the school leavers (690,000 from about 1,400,000) in Russia sat USE. In my presentation I'll try to give the general picture of USE implementation and prospects for it.

Victor N. Simkin

Dr Victor N. Simkin is the Head of the Foreign Languages Teaching Department at the Russian Ministry of Education Institute for General Education.

A former teacher of English and a deputy headmaster of a Moscow secondary school, he started his academic career as a researcher and lecturer in FLT methodology about fifteen years ago. His main academic interests lie in the field of language testing and assessment. For several years he directed the Centre for Linguistic Education Quality Assessment at Moscow State Linguistic University. Victor is now also the coordinator of the Development of the Unified State Examinations in Foreign Languages Project.

**Friday
15.30-16.30**

Reading Achievement of Primary School Students in Thirty Five Countries: Reflections on the Results of the Progress in International Reading Literacy Study - IEA PIRLS

Gabriella Pavan De Gregorio (INVALSI, Istituto Nazionale per la Valutazione de Sistema, Italy)

The presentation will describe and discuss IEA PIRLS 2001, an international Study of the fourth grade reading achievement promoted by IEA, the International Association for the Evaluation of Educational Achievement.

The Study will be presented from various perspectives: what motivations lay behind the decision to assess levels of reading competence in primary education; what concepts of reading were adopted as a basis for the structure of the Study; what ways, procedures and instruments were used to implement the conceptual design of the Study; what basic indications were obtained from its results.

The presentation of PIRLS 2001 will also include the rationale, description, construction process and application of some of its main instruments: the tasks and questionnaires administered in the sampled schools, and the guides used to score the constructed-response items.

Some significant points will be chosen to reflect on the correlation between students' reading achievement and home, school and national factors. Also in the light of significant comparisons within and between countries, the results will be reported of case studies carried out at the national level on a number of high achieving and low achieving schools, in the aim to explore more thoroughly the causes for children's success and non-success in reading.

Gabriella Pavan De Gregorio

Gabriella Pavan De Gregorio is a former researcher at CEDE, the National Institute for Educational Research in Italy and is now a consultant at INVALSI, the National Institute for the Evaluation of the Education System.

After a degree in foreign languages and literatures, with a specialization in English, she undertook post graduate training in linguistics. She taught English language and literature in high schools, and then became a teacher trainer in language education.

She has co-ordinated various national and international research projects. In particular, she was internal co-ordinator for the IEA Written Composition Study and National research Coordinator for the IEA LES - Language Education Study- and for the IEA PIRLS - Progress in Reading Literacy Study. She was also responsible for the construction of the Italian assessment materials within DIALANG - Diagnosis of Foreign Language Skills - a project financed by the European Commission.

Her research interests include: assessment and evaluation in the humanities subjects; relationships between quantitative and qualitative methods of research in education; training design; and teacher development in primary education.

SATURDAY, NOVEMBER 8**Saturday
11.30-12.30****Trends in the predictive validity of high school grades and aptitude tests***Jan-Eric Gustafsson (Department of Education, Gothenburg University, Sweden)*

Admissions to colleges and universities often involve selection of students because the number of applicants is larger than the number of study places. Admissions systems must, among other things, be fair and effective and it is no simple task to design and evaluate such systems. The purpose of the presentation is to review findings from research on selection to higher education, with a special focus on a new Swedish admissions system, which was implemented in the early 1990s. This system is based on two major selection instruments: high school grades and an aptitude test called the Swedish Scholastic Aptitude Test (SweSAT). In the first part of the presentation the methodological issues involved in investigating the predictive validity of selection instruments are discussed, and a review of empirical results from studies conducted in the USA is presented. It is concluded that high school grades provide a somewhat better prediction than aptitude tests like the SAT, but that the aptitude tests add to the prediction when used in combination with grades. It is also observed, however, that predictive validity varies over time, and factors, which may account for this phenomenon, are discussed. In the second part of the presentation results are presented from a large-scale evaluation study of the Swedish admissions system. The results show high school grades to be a better predictor of achievement than the SweSAT, but the predictive validity is also found to vary over time and over area of study. A model which aims to account for the trends in the predictive validity of high school grades and aptitude tests is proposed.

Jan-Eric Gustafsson

Jan-Eric Gustafsson has been a professor of Education at Gothenburg University, Sweden, since 1986. He has conducted a large number of research projects, and has published within several areas. Most of the research concerns problems of individual differences in educational contexts, and the development of the methodology appropriate for the study of such problems. One of his main areas of research concerns the structure of intelligence. Using confirmatory factor analytic methods to evaluate competing models of the structure of intelligence Gustafsson suggested in the early 1980s a three-level hierarchical model, which at its apex has a general factor which is identical with the ability dimension labelled fluid intelligence. This hierarchical model, which integrates several previously proposed models, has been corroborated in several later studies and extended both in his own research and by other researchers.

This line of basic research also has generated work on several more applied problems in the measurement of aptitude for instruction. Currently much attention is devoted to analyzing measurement characteristics of the Swedish Scholastic Aptitude Test and other instruments for selection to higher education. Another substantive field of research concerns evaluation of education through large-scale follow-up and international comparative studies. Gustafsson thus has been involved in projects reanalyzing data from previous IEA studies, and is currently engaged in the PIRLS project on reading literacy. Another current project is a study of the admission system at Swedish universities, which involves research on the validity of the Swedish Scholastic Aptitude Test, and other selection instruments. That research relies on data from very large, population-based, longitudinal data bases.

Another area of research concerns development of measurement models and multivariate methods. In the late 1970s Gustafsson published several pieces of work on the Rasch model, and he also has done work on classical test theory, and measurement theory within structural equation modelling. Lately most of the methodological work has been concentrated on structural equation modelling, and on making these techniques applicable to problems within the field of education, where phenomena typically have a multilevel character and where incomplete observations are common. In this field contributions also have been made in the form of computer programs, in addition to published research.

Discussion Groups

FRIDAY 11.00-12.45

1) Using Portfolios for Educational Assessment: Two Case Studies

Led by: Grace Grima, Debbie Chetcuti; (University of Malta) and Val Klenowski (Institute of Education, University of London, England)

This workshop will explore, from two different contexts, the use of portfolios for learning and assessment as an integral part of students' learning experiences in Higher Education. Students' perceptions of the relevance and value of portfolios will be discussed as well as how the portfolio is used for both formative and summative purposes.

Contact: Dr Val Klenowski
 Institute of Education
 University of London
 v.klenowski@ioe.ac.uk
 Tel: +44 (0)207 612 6341

2) Target setting within Global Education Programmes

Led by Harvey Goldstein, (Institute of Education, England)

UNESCO's Education for All (EFA) programme is an example of a new type of ambitious project that seeks to meet specific 'targets' defined in terms of cognitive development as well as more conventional aims such as student enrolments.

EFA sets out numerical targets, or performance indicators, for areas such as basic literacy and is designed to monitor each country's progress towards these. There are marked similarities to the use of targets in parts of the Anglophone world such as the UK, where penalties for failure and success are an important part of the system. At an international level such penalties may be linked to aid or debt systems in order to exert control and possibly also to impose certain kinds of 'solution'. With the opening up of Education services within GATS, the development of internationally 'comparable' numerical performance indicators provides one kind of universal trading currency that could expedite the exploitation of education 'markets', especially in the developing world.

The workshop will explore these and related issues and participants are encouraged to contribute from their own experiences and perspectives.

Contact: Harvey Goldstein
 Institute of Education
 University of London
 h.goldstein@ioe.ac.uk

3) Media Coverage of Assessment Issues

Led by Roger Murphy and Paul Warmington (University of Nottingham, England)

In recent years news coverage of A Level and GCSE examination results has seen media reporting on education in the UK assume unprecedented levels. At times educationalists respond by deriding media debates such as these as ritualised, stale and superficial.

However, in taking a defensive response, education professionals often retreat into lay perceptions of media communications: in particular, paying scant attention to the mediating effects of the presentational and production contexts in which news messages are presented. Have educationalists underestimated the importance of the popular media as a key interface between the education sector and the public - and thereby unwittingly contributed to the poor quality of debate?

This workshop draws upon analysis of print and broadcast media coverage of UK examination results in 2002 and 2003. Its aim is to encourage critical analysis of the impact that media coverage of assessment issues has on perceptions of exam standards and to consider the extent to which opportunities might exist for educationalists to influence media coverage of examination results and standards.

Contact: Professor Roger Murphy
 Director, Centre for Developing and Evaluating Lifelong Learning
 Director, Institute for Research into Teaching and Learning in HE
 School of Education, University of Nottingham
 Dearing Building, Jubilee Campus
 Nottingham NG8 1BB
 Tel: +44 (0)115 846 7201
 email roger.murphy@nottingham.ac.uk

SATURDAY 9.30 - 11.00

4) Diagnostic Assessment in Early Years

Led by: Marian Sainsbury, (NFER, England)

A national assessment of the pre-school curriculum has recently been introduced in England. Known as the Foundation Stage Profile, it consists of a series of assessment scales which are completed from teachers' ongoing observations. It is compulsory for children of around five years of age who receive government-funded education, a national cohort of about 600,000. A description of the features of the Foundation Stage Profile will lead on to the discussion of issues surrounding the assessment of very young children. It is hoped that participants will bring experience of working in this area and examples of other assessment approaches to inform a wide-ranging discussion.

Contact: Dr Marian Sainsbury
 Deputy Head, Department of Assessment and Measurement
 National Foundation for Educational Research
 The Mere, Upton Park, Slough SL1 2DQ, England
 Phone +44 (0)1753 574 123
 m.sainsbury@nfer.ac.uk

5) The Use of Sophisticated Item Formats in Computer-Based Assessment.

Led by: Andrew Boyle and Dougal Hutchison (NFER, England)

Many commentators have agreed that Computer-Based Assessment (CBA) is and will be a very significant influence on assessment more generally. Early implementations of CBA often used simple item styles (such as multiple choice). However, more recent assessment developers from a range of educational sectors have attempted to use technology to provide assessments that measure knowledge, skills and abilities in more appropriate ways. Such sophisticated assessment systems hold out the possibility of more authentic assessment, and may even lead to a closer integration between assessment and curriculum activities.

But the move to complex item formats also poses challenges for assessment professionals. Firstly, validity may need to be recast in the light of current knowledge of cognition and learning in particular subjects. Further, the integration of assessment with learning might also require the reconceptualisation of validation techniques. Finally, the use of assessments based on complex items may require sophisticated modelling of data to provide indicators of item and test quality, as well as necessary outputs such as test scores. The workshop will discuss these and other issues and participants' own experiences of such issues will be welcome in the session.

Contact: Andrew Boyle
 Senior Research Officer
 Computer Based Assessment
 National Foundation for Educational Research
 The Mere, Upton Park, Slough SL1 2DQ United Kingdom
 Phone +44 (0)1753 574 123
 Email: a.boyle@nfer.ac.uk

6) Monitoring Learning and Achievement: linking research, policy and practice. The Scottish Experience.

Led by Carolyn Hutchinson (Scottish Executive Education department, Scotland) et al.

The relationship between research, policy and practice in assessment is a complicated business. In the early 1990s the Scottish guidelines Assessment 5-14 stressed the prime importance of formative assessment. However, subsequent evidence suggested that schools did not tackle the practical implications of putting formative assessment into practice.

In the consultation on the future of Assessment 3-14 in Scotland (Hayward, Kane & Cogan, 2000), respondents identified three major issues that would have to be addressed if the principles of Assessment 3-14 were to be put into practice. Firstly, work was needed to understand the practical implications in classrooms. Secondly, the relationship between assessment for learning and assessment for measurement should be revisited. Thirdly, the bureaucracy of assessment had to be minimized to allow a real focus on learning.

The Scottish assessment development programme was initiated in response to the issues raised in the consultation. It aims to bring research, policy and practice in assessment into closer alignment. The first part of this session will summarise the major research influences on the programme. The second part will explore the interrelationship between research and practice in the 170 schools involved in the programme, and consider the implications for developing policy.

Contact: Carolyn Hutchinson
 Head of Assessment Branch,
 Qualifications, Assessment and Curriculum
 Scottish Executive Education Department
 Area 2-A Victoria Quay
 Edinburgh EH6 6QQ
 Tel: +44 (0)131 244 0417
 e-mail: carolyn.hutchinson@scotland.gsi.gov.uk

About the Association for Educational Assessment – Europe

The founding meeting and first general assembly of the Association was held in Arnhem in January 2000. Twelve countries were represented and delegates agreed the foundation of the Association. The Assembly elected the first seven members of the Council. The General Assembly re-convened in Prague in November 2000. Ninety delegates from 29 countries attended the first, extremely successful, conference. The third General Assembly was held in Krakow in 2001, and the fourth in Frascati, Rome, Italy in November 2002. In 2003, the Association has 47 corporate and 55 individual members from 30 countries. The Association received initial financial support from CITO and NFER but is now self financing from its members. The management of the Association is entrusted to the Council. Each year, AEA-Europe organizes a conference on assessment themes relevant to the countries of Europe.

GOALS OF AEA-EUROPE

The general goal of the association is to act as a platform for discussion of developments in educational assessment in Europe, fostering co-operation and facilitating liaison between organisations and persons active in educational assessment across the whole of Europe. AEA-Europe defines educational assessment in its broadest sense including academic, professional and vocational contexts and is equally concerned with both assessment processes and products.

AIMS AND OBJECTIVES

The aims of AEA-Europe are to:

- improve communication among European organizations and institutions interested in educational and occupational assessment through sharing of professional expertise, exchange of knowledge and collaboration between members through conferences and publications;
- provide a framework within which co-operative research, development, implementation and evaluation of projects involving educational assessment can be undertaken;
- foster and enhance collaborative networks and projects between organisations and individuals across the whole of Europe;
- co-operate with other agencies having complementary interests;
- engage in a range of activities that will lead to the improvement of assessment processes and products and their appropriate use by organisations, institutions, agencies and other associations throughout Europe;
- and enhance awareness of assessment processes and products in relation to their impact on learning and understanding.

COUNCIL OF AEA-EUROPE

President

EMMA NARDI

Università Roma Tre, Italy
e.nardi@educ.uniroma3.it

Vice President

ROXANA MIHAIL

National Assessment and Examinations Service, Romania
mrox@fx.ro

General Secretary

STEVEN BAKKER

ETS Europe, The Netherlands
sbakker@ets-europe.org

Treasurer

CHRIS WHETTON

NFER (National Foundation for Educational Research) United Kingdom
c.whetton@nfer.ac.uk

Council Member

JANNETTE ELWOOD

Queens University, Belfast Northern Ireland
j.elwood@qub.ac.uk

Council Member

CHRISTINA WIKSTRÖM

Department of Educational Measurement, Umeå University, Sweden
christina.wikstrom@edmeas.umu.se

Council Member

JOHANNA CRIGHTON

Cambridge Assessment and Development in Education (CADE)
johanna934@aol.com

www.aea-europe.net

FOR MORE INFORMATION, CONTACT:

Dr. Steven Bakker, ETS Europe, Janskerkhof 19, 3512 BM
UTRECHT, The Netherlands. **Tel:** +31 30 230 37 30 **Fax:** +31 30 230 37 31
e-mail: aea-europe@ets-europe.org

