proceedings of

EuroCogSci07

the european cognitive science conference

Delphi, Greece 2007

Editors: Stella Vosniadou
Daniel Kayser
Athanassios Protopapas

Psychology Press

Hellenic Cognitive Science Society

Proceedings of

EuroCogSci07

The European Cognitive Science Conference 2007

European Cultural Center of Delphi
Delphi/Greece
May 23–27, 2007

Hellenic Cognitive Science Society

University of Athens

Jointly organized by
the **Cognitive Science Society**,
the **Hellenic Cognitive Science Society,** and
the **Graduate Program in Basic and Applied Cognitive Science**
National and Kapodistrian University of Athens.

Edited by:
Stella Vosniadou, Daniel Kayser & Athanassios Protopapas

Psychology Press
Taylor & Francis Group
LONDON AND NEW YORK

First published 2007 by Lawrence Erlbaum Associates

Published 2014 by Psychology Press
27 Church Road, Hove, East Sussex, BN3 2FA

and by Psychology Press
711 Third Avenue, New York, NY 10017

Psychology Press is an imprint of the Taylor & Francis Group, an informa business

ISBN: 9-781-84169-696-6 (pbk)

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
A catalogue record for this book has been requested

Table of contents

Papers

Reasoning and cooperation

Emotion

Knowledge construction and change

Morphology

Philosophical issues

Learning

Reasoning and uncertainty

Perception

Analogy

Semantics and pragmatics

Computer and formal models

Memory

Models and representations

Neurocognition

Papers presented as posters

Posters

Preface

This volume contains the invited lectures, invited symposia, symposia, papers and posters presented at the 2nd European Cognitive Science Conference, in Delphi, Greece, May 23–27. The 2007 European Cognitive Science Conference – EuroCogSci07 – is the second European Conference in Cognitive Science to be held under the auspices of the Cognitive Science Society. The first such Conference – EuroCogSci03 – took place in Osnabrück, Germany in 2003 and was organized by Franz Schmalhofer and Richard Young.

EuroCogSci07 follows a long tradition of European Meetings in Cognitive Science including a series of biennial meetings with the title "European Conference on Cognitive Science" (E.C.C.S.). The first E.C.C.S. was held in 1995 in St. Malo (France), the second in Manchester (U.K.), and the third in Sienna (Italy). In 2001, the Cognitive Science Society held its first non-North American Conference in Edinburgh (U.K.), organized by Keith Stenning and Johanna Moore. The success of the meeting prompted European researchers to propose that the Cognitive Science Society holds its meeting in Europe once every three years and also that it sponsors a series of regional European Conferences to be held every four years. The University of Osnabrück was the selected site for the first of these meetings.

The papers presented in this volume range from empirical psychological studies and computational models to philosophical arguments, meta-analyses and even to neuroscientific experimentation. The quality of the work shows that the Cognitive Science Society community in Europe is an exciting and vibrant one. There are 210 contributions by cognitive scientists from 27 different countries, including USA (33), France (29), UK (23), Germany (21), Greece (19), Italy (12), Belgium (11), Japan (6), Spain (5), Bulgaria (5), the Netherlands (5), Australia (4), Canada (4), Cyprus (4), Finland (4), Ireland (4), Switzerland (4), Russia (3), Sweden (3), Norway (2), Poland (2), Singapore (2), Argentina, Austria, Brazil, Israel, and South Korea.

An international program committee with members from 12 European countries, Australia, and the United States, and a panel of 321 reviewers from around the world, helped us in the selection of the best papers. A total of 211 six-page papers and 96 one-page "poster abstracts" were submitted for review. From these, 85 papers were accepted for oral presentation and for publication in this volume. An additional 68 six-page papers and 68 one-page submissions were accepted for poster presentation and publication.

We would like to acknowledge help from the following sources who contributed to the success of the conference: The Cognitive Science Society Board, for inviting us to host the EuroCogSci07 and for providing the framework, expertise, and support; the Program Committee, who assigned submissions to referees, read their resulting reviews and made final recommendations to the chairs; the reviewers, who reviewed the submissions and gave feedback to the committee and to the authors; the Local Organizing Committee, and the Students of the Local Organizing Committee, who helped with the myriad local arrangements for the meeting; and the many volunteers, who contributed to the success of the Conference. Our special thanks go especially to Svetlana-Lito Gerakakis for providing extremely helpful administrative support. We are also thankful to the University of Athens Cognitive Science Lab secretary, S. Efthymiou for secretarial support.

We would also like to thank the Invited Speakers: Margaret Boden, Cristiano Castelfranchi, Jerry Fodor, Catherine Fuchs, Randy Gallistel, Rochel Gelman, Gerd Gigerenzer, and Nancy Nersessian; the Organizers of the Invited Symposia: Tatiana Chernigovskaya, Erik De Corte, Stefan Frank, Peter Gärdenfors, Dedre Gentner, Kenneth Hugdahl, Boicho Kokinov, Konstantinos Moutoussis, and Hedderik van Rijn; all those who submitted proposals, for their considerable effort and for their interest in the conference; the authors and symposium participants, for the preparation and presentation of their work; and all those who attended the conference and made it what it was.

Finally, we wish to acknowledge the following for their financial contributions to the conference, and to thank them for their support: European Office of Aerospace Research and Development – Air Force Office of Scientific Research – United States Air Force Research Laboratory, Association pour la Recherche Cognitive, British Council of Greece, Cognitive Science Society, Education Research Center of Greece, Endolysi-Medical Technologies, French Embassy of Greece, Info-Quest, Institut Français d'Athènes, Istituto Italiano di Cultura di Atene, Laboratoire d'Informatique de Paris-Nord, Greek Ministry of National Education and Religious Affairs, Office of Naval Research, Olympic Airlines, Plus Orthopedics Hellas SA, University of Athens, University of Paris-Nord.

Stella Vosniadou, Daniel Kayser, and Athanassios Protopapas

May 2007

EuroCogSci07

May 23-27, 2007

European Cultural Center of Delphi

Program Chairs

Stella Vosniadou (National and Kapodistrian University of Athens, Greece)
Daniel Kayser (Université Paris Nord, France)

International Program Committee

Maria Teresa Bajo Molina (Spain)
Bruno Bara (Italy)
Cristiano Castelfranchi (Italy)
Fintan Costello (Ireland)
Michel Denis (France)
Robert French (CNRS, France)
Peter Gärdenfors (Sweden)
Dedre Gentner (USA)
Gerd Gigerenzer (Germany)
Christopher Habel (Germany)
Boicho Nikolov Kokinov (Bulgaria)
Jesus M. Larrazabal (Spain)
Daniel Navarro (Australia)
Nancy Nersessian (USA)
Jean – Luc Nespoulous (France)
Mike Oaksford (UK)
Fabio Paglieri (Italy)
Massimo Poesio (UK)
Emmanuel M. Pothos (UK)
Athanassios Raftopoulos (Cyprus)
Franz Schmalhofer (Germany)
Gale Sinatra (USA)
Matti Sintonen (Finland)
Peter Slezak (Australia)
Gerhard Strube (Germany)
John Taplin (Australia)
Maarten Van Someren (Netherlands)
Bernard Victorri (France)
Richard M Young (UK)

Conference Secretariat

Svetlana Lito Gerakakis

Electronic Submission and Review Site Administration

Athanassios Protopapas

Local Organizing Committee

Georgios Giftodimos
Maria Grigoriadou
Georgios Kouroupetroglou
Alexandra Oikonomou
Stathis Psillos
Petros Roussos
Stavroula Samartzi
Fotini Stilianopoulou
Petros Tzelepidis

Graduate students of the Local Organising Committee

Constantinos Christou
Rania Gikopoulou
Kalliopi Ikospentaki
Philippos Katsoulis
Maria Koulianou
Katerina Ligovanli
Georgia Rovatsou
Irini Skopeliti
Stergios Tegos
Xenia Vamvakoussi

Student Volunteers

Panajotis Blitsas
Giorgos Briskolas
Aspasia Kanelou
Thanos Kremizis
Eleftheria-Maria Pattakou
Despina Sakoulogeorga
Maria Stournara
Amalia Syrgounioti

Reviewers

Brandon Abbs
Marianne Abramson
Afra Alishahi
Jedediah Allen
Janet Andrews
Christa S.C. Asterhan
Dilip Athreya
Roger Azevedo
Joscha Bach
Teresa Bajo
Benjamin Balas
Gianluca Baldassarre
Jerry Ball
Bruno G. Bara
Luke Barrington
Lawrence Barsalou
William Bechtel
Giovanni Bennardo
Bettina Berendt
Nadia Berthouze
Lisa Best
Maryse Bianco
Zafer Bilda
Nik Nailah Binti Abdullah
Henrike Blumenfeld
Tibor Bosse
Lewis Bott
Ken Brown
Duncan Brumby
Raff Calitri
Thomas Capo
Laura Carlson
Cristiano Castelfranchi
Stephanie Chambaron
Christine Chan
Franklin Chang
Tatiana Chernigovskaya
Seth Chin-Parker
Key-Sun Choi
Georgios Christou
Evangelia Chrysikou
Adam Chuderski
Bill Clancey
Catherine Clement
Charles Clifton
James Close
Cheryl Cohen
Louise Connell
Fofi Constantinidou
Rick Cooper
Adam Corner
James Corter
Rui Costa
Fintan Costello
Rick Dale
Moritz M. Daum
Eddy Davelaar
Jim Davies
Michel Denis
Natacha Deroost
Cathy Desoto
Barry Devereux
Denise Dillon
Mike Dowman
Kevin Dunbar
Nicholas Duran
Lisa Durrance
Karine Duvignau
Simon Dymond
Jeronimo Dzaack
Inge-Marie Eigsti
Michelle Ellefson
Shira Elqayam
Stephanie Elzer
Tanja Engelmann (née Keller)
Susan Epstein
Martha Evens
Michele Feist
Anne Fernald
Roland Fernandez
Philip Fernbach
Mark Finlayson
Brian Fisher
Robin Flanagan
Kristine Flood
Stephani Foraker
Tim Fosker
Stefan Frank
Bob French
Daniel Freudenthal
Caren Frosch
Danilo Fum
Francesco Gagliardi
David Galbraith
Joao Gama
Peter Gardenfors
Margarida Garrido
Dedre Gentner
Claudia Gianelli
Marco Giunti
Fernand Gobet
Winston Goh
Ariel Goldberg
Laura Gonnerman
Arthur Graesser
Floriana Grasso
Ann Grbavec
Collin Green
Michelle Greene
Thomas Griffiths
Maurice Grinberg
Martin Groen
Joshua Gross
Amal Guha
Michelle Gumbrecht
Glenn Gunzelmann
Christopher Habel
James Hampton
Oliver Hardt
K. David Harrison
Sebastien Helie
Tatjana S. Hilbert
Paul Hodgson
John Hoeks
Elena Hoicka
Jessica Horst
Penka Hristova
Evgenia Hristova
Janet Hui-Wen Hsiao
Sylvie Huet
Markus Huff
Almut Hupbach
Fumihito Ikeda
Armina Janyan
Tiffany Jastrzembski
Melander Johns
Jeffrey Johnson
Linda Kaastra
Daniel Kayser
John Kearns
Frank Keller
William Kennedy
Tim Kenyon
Laura Kertz
Julian Kiverstein
Sachiko Kiyokawa
Alexander Klippel
Harumi Kobayashi
Boicho Kokinov
Lars Konieczny
Anna Koop
Kepa Korta
Alexander Kranjec
Josef Krems
Ben Krose
Sven Kuehne
Kai-Uwe Kühnberger
David Lagnado
David Landy
Jesus Maria Larrazabal
Jorge Larreamendy-Joerns
Sangmok Lee
Hye-Won Lee

Louis Lee
Benoit Lemaire
Margarita Limon
Craig Lindley
Alexandre Linhares
Daniel Little
Manuel Liz
Marcio Lobo Netto
Jeffrey Loewenstein
Deryle Lonsdale
Emiliano Lorini
Max Louwerse
Christopher Lucas
Dermot Lynott
Franηois Lιvy
Matt Macmahon
Carol J. Madden
Joe Magliano
Lorenzo Magnani
Klara Marton
Fulvio Mastrogiovanni
Michael Matessa
Rachel McCloy
Alastair Mckenzie-Kerr
Sean McLennan
David Mendonca
Arron Metcalfe
Kelly Mix
Naomi Miyake
Sergio Moreira
Francesca Morganti
Bradley Morris
Arthi Murugesan
Shahin Nabavian
Mitchell Nathan
Daniel Navarro
Nancy Nersessian
Jean-Luc Nespoulous
Yuk Han Priscilla Ngan
Eric Nichols
Shusaku Nomura
Laura Novick
Padraig O'Seaghdha
Mike Oaksford
Philipp Otto
Bernard Pachoud
Fabio Paglieri
Thomas Palmeri
Sangho Park
Neal Pearlmutter
Bo Pedersen
David Peebles
Amotz Perlman
Georgi Petkov
Giovanni Pezzulo
Raedy Ping
Massimo Poesio
Thierry Poibeau
Fenna Poletiek
Ferran Pons
Emmanuel Pothos
Athanassios Protopapas
Jennifer Queen
Athanassios Raftopoulos
Christopher Ramey
Tânia Ramos
David Reitter
Irene Reppa
Antonio Rizzo
Hannah Rohde
Michael Romano
C. Anton Rytting
Kayo Sakamoto
Ron Salden
Franco Salvetti
Ava Santos
Megan Saylor
Franz Schmalhofer
Ute Schmid
Michael Schneider
Mike Schoelles
Jonathan Schooler
Wolfgang Schoppek
Holger Schultheis
Gerhard Schurz
Roger Schvaneveldt
Angela Schwering
Christopher Schwint
Yasuhiro Shirai
Gale Sinatra
Matti Sintonen
Peter Slezak
Vladimir Sloutsky
Asha Smith
Derek John Smith
Melanie Soderstrom
Hans Spada
David Spurrett
Narayanan Srinivasan
Edward Stabler
Jim Staszewski
Georgia Stephanou
James Stewart
Terry Stewart
Ted Strauss
Gerhard Strube
Tarja Susi
Biswanath Swain
Emmanuel Sylvestre
Niels Taatgen
Hermina J.M. Tabachneck-Schijf
Vanessa Taler
Andrzej Tarlowski
Eric Taylor
Paul Thagard
Jean-Pierre Thibaut
Scott Thomas
Susan Thompson
Robert Thomson
Andree Tiberghien
Maurizio Tirassa
Marc Tomlinson
Jan Treur
Manos Tsakiris
Paolo Turrini
Ryan Tweney
M. Afzal Upal
Frederic Vallee-Tourangeau
Maurits Van Den Noort
Han Van Der Maas
Frank Van Der Velde
Mija Van Der Wege
Tamara Van Gog
Hedderik Van Rijn
Iris Van Rooij
Maarten Van Someren
Argiro Vatakis
Marga Vazquez
Tom Verguts
Bernard Victorri
Stella Vosniadou
Edward Vul
Samuel Waldron
Paul Ward
Daniel Weiss
Markus Werning
Kathryn Wheeler
Jörg Wittwer
Fei Xu
Jie Yan
Ya-Ting Yang
Sule Yildirim
Michael Yip
Richard Young
Johannes Ziegler
Tom Ziemke
Elena Ziori
Iraide Zipitria

Invited Lectures

Using Information Theory to Better Understand Associative Learning

Randy Gallistel (galliste@ruccs.rutgers.edu)
Rutgers Center for Cognitive Science (RuCCS), Psych Bldg Addition, Busch Campus,
152 Frelinghuysen Road, Rutgers University - New Brunswick, Piscataway, NJ 08854-8020 USA

Abstract

Using Shannon's theory of information to quantify the information that a conditioned stimulus (CS) conveys regarding the timing of the next unconditioned stimulus (US) gives a parameter-free, quantitatively rigorous account of background conditioning, blocking, overshadowing and relative validity, while also giving for the first time an empirically valid specification and quantification of the notion of temporal pairing. These results strengthen the idea, dating back to the 1970s, that what drives the learning that occurs in paradigms designed to establish the laws of association formation is not temporal contiguity but rather the learning of the temporal intervals themselves. Learning those intervals is essential to extracting from a protocol the mutual information between two events. The learning that occurs should be conceptualized as the extraction of that mutual information, not the formation of a conductive connection.

Model-based Reasoning in Distributed Cognitive-Cultural Systems: Studies of interdisciplinary research labs

Nancy J. Nersessian (nancyn@cc.gatech.edu)
School of Interactive Computing, Georgia Institute of Technology
Atlanta, GA 30332-0280, USA

Abstract

This paper will examine “model-based reasoning” in the interplay of representation and experiment in the context of two biomedical engineering research laboratories, where problem solving is by means of constructing, manipulating, and revising physical models. Designing, re-designing, and experimenting with *in vitro* simulation models (“devices”) is a signature cognitive practice. These physical models are technological devices that either simulate well-understood mechanisms, such as the forces on arterial vessels from the flow of blood through them, or hypothesized mechanisms, such as how learning takes place among neurons. The devices provide sites of experimentation where *in vitro* models are used to screen and control specific aspects of *in vivo* phenomena that the researchers want to examine. They are constructed and modified in the course of research with respect to problems encountered and changes in understanding. Simulation is an epistemic activity involving exploration, testing, and generation of hypotheses, explanation, prediction, and inference.

In this analysis, I draw on and contribute to research in contemporary cognitive science that construes cognition as a complex system in which cognitive processes are “embodied, “situated” in environments, and “distributed” across people and artifacts. Model-based reasoning in the complex systems of the laboratory is argued to involve simulation processes in which mental and physical models of both the phenomena under investigation and the simulation device are co-constructed, manipulated, and revised. That is, the devices act as ‘hubs’ for interlocking mental models and experimentation. The design and redesign of a device is thus both driven by changes in the mental models and experimental results and lead to changes in mental models and experimental designs. The discovery processes thus run on a hybrid of internal and external structures. Further devices are hubs of interdisciplinary *melding* of cultural, social, material, and cognitive practices. In particular, the mental models are hybrids of various disciplines, and the structure of the device concretely instantiates this hybrid nature. Modelbased reasoning, thus, needs to be understood as being performed within complex cognitive-cultural systems, distributed in space and time.

Gut Feelings: The Intelligence of the Unconscious

Gerd Gigerenzer (sekgigerenzer@mpib-berlin.mpg.de)
Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development
Lentzeallee 94, 14195 Berlin, German

Abstract

We think of intelligence as a deliberate, conscious activity guided by the laws of logic. Yet much of our mental life is unconscious, based on processes alien to logic: gut feelings, or intuitions. We have intuitions about sports, friends, the toothpaste to buy, and other dangerous things. We fall in love, and we sense that the Dow Jones will go up. How do these feelings work? I define an intuition as a judgment that (i) appears quickly in consciousness, (ii) whose underlying process we are not aware of, yet (iii) is strong enough to act upon. I argue that the underlying process can often be described by fast and frugal heuristics, which take advantage of evolved capacities of the brain. Good intuitions behave differently from logical systems: more information or more time does not always lead to better decisions. Moreover, in a moderately unpredictable world, simple heuristics can lead to better judgments than can "optimal" models of decision making. The science of intuition studies the processes underlying snap judgments, and the environmental structures in which they fail and succeed.

References

Gigerenzer, G. (2007). *Gut feelings: The intelligence of the unconscious.* New York: Viking Press.

Early Cognitive Development and Beyond

Rochel Gelman (rgelman@ruccs.rutgers.edu)
Rutgers Center for Cognitive Science (RuCCS), Psych Bldg Addition, Busch Campus, 152 Frelinghuysen Road, Rutgers University - New Brunswick, Piscataway, NJ 08854-8020, USA

Abstract

Any account cognitive development must handle two general facts about knowledge acquisition. Young children, living in reasonably healthy and normal environments of their culture, learn a great deal without formal instruction and often "on the fly". Indeed, they acquire the language of their community, and develop a set of intuitive understandings of natural number arithmetic, the difference between animate and inanimate objects and the role of causality regarding the transformation and movements of different kinds of objects. In the case of early learnings, children benefit from the existence of skeletal, domain-specific structures. They use these to identify examples of data in the environment that are structural maps. In this sense, some early kinds of learning are privileged. Although some of these structures can foster the accumulation of yet more knowledge in a domain, there are clear cases where this is not the case. Indeed, evidence indicates that early learnings can stand in the way of the mastery of new knowledge with understanding. For example, children's knowledge of natural numbers is inconsistent with the task of learning, with understanding about rational numbers and therefore a conceptual change about the nature of natural numbers. The learning problem is tied to the fact that the mathematical structure for rational numbers does not map readily to that for natural numbers. For example, whereas every natural number has an unique next, this is not so for the rational numbers. Additionally when two natural numbers that are >1 are multiplied, the answer is always greater. However, multiplication of two fractions yields a smaller value. The learning task then becomes one of mounting both a new structure with new entities and the rules of combination. We know that the acquisition of new conceptual structures takes work on the part of the learner and a great deal of time. The question then becomes: what fosters the acquisition of new domains of knowledge. I will propose that a variety of learning tools are called upon to help in the creation of new organized domains of understanding.

The History of Cognitive Science: Seven Key Dates

Margaret A. Boden (maggieb@cogs.susx.ac.uk)
Cognitive and Computing Sciences, University of Sussex
Falmer, Brighton, BN1 9QN, UK

Abstract

Cognitive science has seen seven key dates. The first four were 1943, 1956, 1958, and 1960. Important things happened later, too: in 1969, 1986, and 1987. Those seven years (with 1947 and 1979 as runners-up) all saw seminal publications and/or influential interdisciplinary meetings, in which different methodologies and research opportunities were introduced or highlighted--or, in one case, trenchantly attacked. The current profile of the field has been shaped accordingly.

Introduction

Cognitive science has been studied for some seventy years, and covers six different disciplines: AI/A-Life, psychology, neuroscience, linguistics, anthropology, and philosophy. So distilling its history into just seven dates is highly artificial. The thumbnail sketch that follows is based on the book I've recently written on the topic, where everything mentioned here is explored in greater detail (Boden 2006).

The thematic heart of cognitive science is psychology, and its intellectual heart is AI/A-Life. In other words, it's the study of *mind as machine,* its core assumption being that the same scientific concepts apply to minds and mindlike artefacts.

Since the machines in question are of two main types, there are two major theoretical pathways across the field. One is grounded in logical-symbolic computation, or GOFAI (Good Old-Fashioned AI). The other features adaptive, self-organizing, and/or feedback devices--including certain sorts of connectionist system. We may call them the cybernetic/connectionist and the symbolic--but they both arose out of the cybernetics movement of the 1940s, as we'll see. The field's history has been shaped by the contrasts and competition between these two approaches.

Wartime Thoughts

The first key date within cognitive science is 1943. That wartime year saw three influential publications. The most important was Warren McCulloch and Walter Pitts' essay 'A Logical Calculus of the Ideas Immanent in Nervous Activity'.

This combined three hugely exciting, but *prima facie* highly diverse, ideas of the early twentieth century: neurone theory, the Turing machine, and the Russell-Whitehead propositional calculus. The authors argued that these were formally equivalent. That is: every expression of the propositional calculus could be computed by some Turing machine, which in turn could be physically implemented in some definable neural net. Logic, computation, and the brain were all of a piece.

In seeing the mind/brain as a Turing machine, McCulloch and Pitts weren't thinking only of cognition: for *all* psychological processes, they said, "the fundamental relations are those of two valued logic". Even in psychiatry, they added, "Mind no longer goes 'more ghostly than a ghost'". Formal networks should be the psychologist's goal: "specification of the net would contribute all that could be achieved in [psychology, however defined]".

These ideas inspired John von Neumann immediately, leading him to design his computer as a machine grounded in binary (true/false) logic instead of decimal arithmetic. But their influence on theoretical psychology was delayed, for three reasons. First, the paper appeared in an obscure journal which few psychologists saw. Second, it used a rebarbative logical formalism (borrowed from Rudolf Carnap), guaranteed to repel most readers. And last, it had no connection with the various wartime problems dominating psychologists' minds in 1943. Its significance would be widely realized only later.

McCulloch and Pitts here initiated *both* theoretical pathways of cognitive science. On the one hand, their paper led to the psychologically oriented connectionism of the 1940s/1950s, initially implemented in wire-and-solder contraptions, not in general purpose computers. On the other hand, once digital computers arrived a few years later, their paper was seen to imply that language-based meanings and reasoning could be modelled by them. (McCulloch, in fact, had long been a follower of the logical atomists' philosophy of language). That is, it seemed reasonable to hope both that symbolic AI was possible and that it could be seen as theoretical psychology.

(Four years later, these two authors would admit that their precisely structured logical networks, and unvarying neural thresholds, didn't reflect the noisy, error-prone, and damageable nature of the brain: Pitts and McCulloch 1947. So they now outlined a statistical form of connectionism--and even suggested which parts of the brain perform which types of computation. As the pioneering paper in computational neuroscience, and in distributed computing and probabilistic networks too, this might tempt one to add an *eighth* key date to the list. However, their later paper didn't attract many followers. Moreover, they saw it as an "extension" of the earlier one, whose core claim--that neural networks can be *theoretically* mapped onto binary logic--was specifically repeated. Let's mark 1947, then, as an honourable runner-up.)

While McCulloch and Pitts had been writing their ground-breaking paper, three other members of the cybernetic community--including Norbert Wiener himself--had been analysing "purpose and teleology" in terms of negative feedback (Rosenblueth et al. 1943).

This, they said, could be used so as to reduce the differences between the current state and the goal--an idea that was mentioned in the 'Logical Calculus' paper, too. As they put it, "The signals from the goal are used to restrict outputs which would otherwise go beyond the goal" (p. 19). The examples they listed included heat-seeking missiles, and the muscular overshoot seen (in grasping a glass, for instance) in Parkinsonism. In general, adaptive 'goal-seeking' behaviour of humans and animals was assumed to be controlled in this way.

However, these authors thought of the "goal" as a target, rather than a goal. (In their most persuasive example, heat-seeking missiles, it was exactly that.) The key was perception, not intention. Goals (and sub-goals) considered as imaginary, and intended, future states weren't in question. Nor could they be. For there was no mention of internal models, or representations, of the goal--or of current states of the world.

In the very same year, those very matters were being highlighted across the Atlantic in Kenneth Craik's little book *The Nature of Explanation* (1943). This introduced the notion of cerebral models, borrowed from the neurologist Henry Head, into cognitive psychology and the philosophy of mind (Craik described his book as a work of *philosophy).* And it glossed them, for the first time, in terms of the functioning of man-made machines.

The machines Craik had in mind were analogue devices, such as the tidal predictor and the differential analyser. The representational power of cerebral models, he said, lay in the fact that--like the machines just mentioned--each one was "a physical working model which works in the same way as the process it parallels, in the aspects under consideration at any moment" (1943: 51). And he offered some specific hypotheses about the neurophysiology of various analogue "models" for perception.

Although Craik called his approach "a symbolic theory of thought" and referred to "symbolism" in the brain, he seemed to be thinking of representation in general (including language) rather than the logical-computational variety. He died (in an accident) in 1945, so didn't see the rise of GOFAI. Probably, he would have accepted formal-symbolic representations as alternative types of cerebral model. Certainly, many of his followers did. Two early cognitive scientists who acknowledged Craik's inspiration were Richard Gregory (e.g. 1966) and Jerome Bruner (who'd visited Craik's group in England in 1955-56).

Largely as a result of these three publications of 1943, the next quarter-century saw pioneers working on both types of AI and/or computational psychology.

It would be misleading to say that they were working on both sides of the theoretical *divide,* because the unpleasantly antagonistic schism between connectionism and GOFAI, or (more broadly) between bottom-up and top-down approaches, hadn't yet developed. At that time, there was still one intellectual community ("cybernetics"), with shared aims and interests. To be sure, some people were focussing more on adaptation and self-organization, others more on logic and meaning--although a few, such as McCulloch, tackled both. Indeed, the rapid rise of GOFAI was mainly due to its promise, not matched by the adaptationists, to deal with inference and linguistic meaning. In general, however, the two sides communicated freely and agreed to differ on what might be the most promising theoretical approach. Only much later did the community separate into distinct sociological camps, with little love lost between them (see Section V).

The 1950s

The key dates of the following decade were 1956 and 1958. Indeed, 1956 was the *annus mirabilis* of cognitive science. It saw no fewer than six events that raised the spirits of the nascent cognitive scientists, convincing them that something exciting was happening. Four were publications, and two were meetings aimed at consciousness raising in the emerging interdisciplinary community.

The publications included a book reporting an imaginative series of psychological experiments: Bruner's *A Study of Thinking* (Bruner et al. 1956). (The title alone was a provocation, in those behaviourist days.) Bruner posited several information-processing strategies for concept learning, each more or less appropriate depending on the circumstances--and all defined in broadly computational terms. His ideas would be reflected in much early AI and computational psychology.

In addition, there were three papers. The most influential was George Miller's (1956) information-theoretic 'The Magical Number Seven'--which by the mid-1970s had become the most-cited paper in the whole of cognitive psychology. Another described the first computer model of Donald Hebb's "connectionism" (the word was his coinage). This showed that Hebbian theory could be implemented, but only if his 'ft/wt' learning rule was expressed more precisely (Rochester et al. 1956). And--across the ocean--the last was Ullin Place's 'Is Consciousness a Brain Process?' (Place 1956).

Place's paper is the outlier here, for his mind-brain identity theory wasn't a contribution to cognitive science as such: it said nothing about mind as machine. But it was eagerly welcomed by scientifically-minded readers, and its materialist spirit--though not its reductionist letter—was retained when philosophical functionalism replaced it four years later (see Section IV).

As for the two consciousness-raising meetings of 1956, the first was the "Summer Research Project on Artificial Intelligence", at Dartmouth College. Organized by the youngsters Marvin Minsky and John McCarthy, this introduced AI to a wider audience. (It also launched the discipline's name, which has been a philosophical millstone around its neck ever since.)

For instance, Minsky handed out the draft of an insightful review of early AI (Minsky 1956). Published a few years later as 'Steps Toward Artificial Intelligence', this was widely seen as AI's manifesto. Or perhaps one should rather say "as GOFAI's manifesto", for it argued that connectionist AI had fundamental limitations not shared by symbolic AI. It did, however, suggest that a *combination* of neural networks and GOFAI would be needed to emulate human thought—a suggestion that went largely unheeded. (Minsky himself seemed to forget it in the 1960s, as we'll see, but he eventually followed it up in his "society" theory of mind.)

The Summer Project wasn't a meeting in the usual sense, but a two-month period during which about a dozen AI pioneers were located at Dartmouth, and anyone who was interested could drop in. The core group included Arthur Samuel and Oliver Selfridge--and, for the .final week, Allen Newell and Herbert Simon. In the earlier weeks they'd played truant, trying to finish programming their Logic Theorist. This proved theorems (in propositional logic) from the Russell-Whitehead *Principia Mathematica,* and even found a more elegant proof for one of them (Newell et al. 1957).

The Logic Theorist wasn't the first AI program, though it's often described that way. Quite apart from 'toy' programs written by Alan Turing and others, Samuel had implemented a heuristic program for playing checkers (draughts) in 1949, and a learning version was up-and-running early in 1955 (Samuel 1959: 72). It had even featured on American TV in February 1956, six months before the Dartmouth event. Unlike Newell and Simon, however, Samuel attended that meeting without bringing along printout evidence. That's partly why the participants were more enthused by the Logic Theorist. In addition, logic struck most people as more impressive--more 'human'--than draughts.

But the main reason why more interest was aroused by Newell and Simon's program was that it was explicitly intended as a model of human thinking, guided by Gestalt psychology and by their own experiments. In their view, computers and psychology should be seen as equal partners: "artificial intelligence was to borrow ideas from psychology and psychology from artificial intelligence" (Newell and Simon 1972: 883). Buffs on both sides of this disciplinary fence were excited accordingly.

The second 1956 meeting was the IEEE's three-day Symposium on Information Theory, convened at MIT in mid-September--almost back-to-back with the Dartmouth event. This had more direct influence in bringing psychologists into cognitive science. For among the papers given there were Miller's 'Magical Number Seven', Newell and Simon on the Logic Theorist, and Noam Chomsky on formal grammars--which showed that language, considered as structured sentences not just as word strings, can be formally described.

Miller himself instantly put those other two talks together: "I went away from the Symposium with a strong conviction, more intuitive than rational, that human experimental psychology, theoretical linguistics, and computer simulation of cognitive processes were all pieces of a larger whole, and that the future would see progressive elaboration and coordination of their shared concerns" (quoted in Gardner 1985: 29). This epiphany soon led him to play a crucial role in establishing cognitive science as such (see Section IV).

So 1956 was a good year for the field--but it was soon followed by another. In November 1958, a four-day interdisciplinary seminar took place at the National Physical Laboratory (NPL) in London--a resonant venue, given its post-war connection with Turing. Hosted by the psychophysiologist Albert Uttley, this brought other leading neurophysiologists--Horace Barlow, for example--into the discussion.

About two dozen people, almost all now important names in cognitive science, gathered there. Most had experience of interdisciplinary thinking, having done war-work on the design and use of various novel machines. And Craik was a highly respected name--and, for several attendees, an inspiring personal memory. Recognized intellectual leaders such as McCulloch, Frederic Bartlett, and the anatomist J. Z. Young were joined by youngsters who today are at least as famous. And the youngsters served up some very rich fare.

The atmosphere was electric: it was clear that something exciting was happening. The importance of this meeting for both "sides" of cognitive science--and for AI, A-Life, psychology, and neuroscience--can be indicated by listing a few of the talks (see Blake and Uttley 1959). Among NPL's many memorable moments were these: Selfridge on *Pandemonium;* Frank Rosenblatt on perceptrons; Barlow on his 'coding' theory of perception; Gregory on the misuse of brain-ablation studies; Donald MacKay on the need for hybrid (analogue-digital) machines; McCarthy on giving programs "common sense" *via* predicate calculus (and Yehoshua Bar- Hillel's critical reply on what's now called the frame problem); Gordon Pask on his electrochemical model of a developing concept; and, not least, Minsky on heuristic programming--who summarized the AI manifesto circulated at Dartmouth two years earlier.

The NPL meeting was only one of three events which made 1958 special. The others were two highly contrasting papers, both published in the same volume of *Psychological Review* and both--at least for a while--hugely influential.

The first to appear was a theory of human problem solving, based on the Logic Theorist and its successor the General Problem Solver, or GPS (Newell et al. 1958). Even more powerful than the Logic Theorist, GPS whetted the appetite of psychologists who hadn't heard of the Logic Theorist, and enthused those who had still further. They were attracted, too, by the programme of ongoing psychological experimentation initiated by the authors.

The second seminal paper was Rosenblatt's (1958) account of "perceptrons", also featured at NPL but here reaching a much wider audience. This described a class of connectionist computer models based on Hebbian theory, and focused not on problem solving but on pattern recognition. They could learn to distinguish an A from a B, for example.

Although perceptrons excited many people, including youngsters entering AI, they didn't convince everyone whom one might have expected to be sympathetic. Indeed, when cognitive science's manifesto appeared two years later (see Section IV), they were near-invisible: even in those hope-filled pages, parallel processing would be mentioned only in two footnotes. Rosenblatt's hopes were more robust. He saw perceptrons as prefiguring a general theory of human psychology, and was even more daring--some would soon say even more preposterous--than Newell and Simon in his predictions concerning future versions of his machine.

It's noteworthy that these two papers were published in the same Journal. That might have happened ten years later--but not ten years after that. For by then the schism mentioned in Section II had emerged: the field's two pathways had diverged not only theoretically but sociologically too.

Meeting-House, Manifesto, and Mind

Most of the influences mentioned so far were drawn together in two ground-breaking projects of 1960. One was cognitive science's first research institute, the other its manifesto.

Harvard's provocatively named Center for Cognitive Studies was co-founded by Bruner and Miller. Bruner had been running a seminar on these matters for some years, attended (for instance) by the young Chomsky and Jerry Fodor. That had sown important seeds in the local community, for Chomsky later acknowledged Bruner's (neo-Craikian) influence on his positing inner representations of syntactic structure. But in 1960 the new Center put interdisciplinary cognitive science publicly on the academic map.

The name was provocative because it was rejecting behaviourism, then dominant in US psychology. But the word "Cognitive" carried less weight than is often thought, being used simply as an anti-behaviourist shorthand. As Miller later put it: "[We] were setting ourselves off from behaviorism. We wanted something that was *mental* --but "mental psychology" seemed terribly redundant" (Miller 1986: 210). In speaking of "cognition", he said, they weren't intentionally excluding "volition" or "conation", but "just reaching back for common sense". In short, even though in practice most cognitive scientists have focused on cognition, the field has always been concerned in principle with *all* aspects of the mind--as McCulloch and Pitts had urged in 1943.

Besides co-founding the Center, Miller offered another spur to cognitive science in 1960. This was his remarkable book *Plans and the Structure of Behavior,* written with Eugene Galanter and Karl Pribram (MGP for short). The book was (unavoidably) simplistic, and careless to boot. Nevertheless, it was a work of vision. Its declared goal was to discover "whether the cybernetic ideas have any relevance for psychology" (p. 3), and its answer was a confident "Yes".

MGP used the notion of a Plan--simply defined as a TOTE unit (Test-Operate-Test-Exit), or as TOTEs made up of lower-level TOTEs--to sketch mental processes. Their discussion ranged over the whole of psychology. Animals and humans; instinct and learning; language and memory; habit and motor skill; chess and choice; values and facts; self image and social role; knowledge and affect; intention and desire; hope and morality; personality and hypnosis; normal life and psychopathology ... *everything* was included.

Plans was the first book to apply computational ideas so widely. Thanks to the recent work of Newell and Simon and of Chomsky, all of whom were repeatedly cited, the most persuasive parts of the book concerned cognition. But the promises reached beyond the persuasion. Miller and Bruner's intention that the "Cognitive" in "Cognitive Studies" should really be read as "mental"-- *anything* mental--was reflected in this volume.

Even sympathetic readers were almost deafened by the sound of handwaving. However, they were excited too. For some years, the book would function as a manifesto for the new science of the mind. (A good way of judging progress in cognitive science is to compare today's achievements with the hopes and promises expressed therein.)

Meanwhile, a mile or so away from the new Center, another 1960 landmark had been constructed: Hilary Putnam's functionalism (Putnam 1960). For budding cognitive scientists, this new philosophy offered relief, revelation, and promise. It escaped various dilemmas that had plagued the philosophy of mind--including Place's identity theory--through the 1950s. More to the point, it saw Turing computation as the causal process at the core of mental life, and the mind as the 'program' of the brain. By implication, it underwrote the AI-based theoretical psychology that was already emerging.

There were naysayers, of course. Indeed, competing varieties of functionalism would later develop within cognitive science. And there would be plenty of objections from philosophers outside the field. (Putnam himself rejected it, eventually.) Nevertheless, this paper had given sharp philosophical teeth to those who wished to chew the mind in computational terms.

By 1960, then, the field had visibly got off the ground.

A Temporary Glitch

The fifth key date, 1969, marks a publication seen by some people as a step backwards rather than forwards. On that view, the damage caused in 1969 wasn't mended until some twenty years later.

MGP weren't the only ones to be under-impressed by perceptrons: Minsky, with Seymour Papert, had a low opinion of them too. He'd already expressed doubts in his 'Steps' paper. But in the 1960s, when Rosenblatt's ideas were threatening to grab the graduate students, and the funding, he (and others at MIT) felt that sterner measures were called for. The result was an explosive little book called *Perceptrons: An Introduction to Computational Geometry* (Minsky and Papert 1969).

As the sub-title implied, this was a mathematical critique. Minsky and Papert showed that simple parallel processors couldn't do certain things, such as recognizing connectedness, which one might have expected them to do--and which the then-current GOFAI programs could do. And they predicted that more complex versions wouldn't be much better. Admittedly, in 1959-60 Rosenblatt had proved that perceptrons could learn to do whatever they could be programmed to do. His proof was allowed to be both valid and "seductive" (p. 14), but--Minsky and Papert argued--it had little practical relevance in face of the combinatorial explosion. What the widely hailed perceptrons could actually do was highly limited. In short, they were fool's gold.

After this publication, funding for connectionism virtually stopped. In the USA it had started to dry up already, thanks to the circulation of the (even more vitriolic) draft of *Perceptrons* during the early-mid 1960s, and to Minsky's close friendship with the key funder at DARPA (Joseph Licklider).

Carver Mead later spoke of "the twenty-year famine" in connectionism (Anderson and Rosenfeld 1998: 141). But, rightly, he didn't put all of the blame onto Minsky and Papert's shoulders. Rather, he blamed the early-1960s "overhype" about perceptrons--to which they'd been responding.

A Double Renaissance

Both of our last two key dates mark a new visibility, not a new activity. Namely, the public renaissance of connectionism--more precisely, of parallel distributed processing (PDP)--in 1986, and of A-Life a year later. In each case, the new visibility prompted an explosion of further activity that's still expanding.

Connectionism hadn't stopped dead in its tracks in 1969. Throughout the 1970s, important work was done on associative memories and distributed representation. However, it was seen as maverick, and largely ignored. A consciousness-raising meeting was held in La Jolla in 1979 (Hinton and Anderson 1981), but it was highly technical: few newcomers were enticed to join the band.

What mended the damage done to connectionism's reputation by Minsky and Papert's attack was the publication in 1986 of the PDP 'bible' (Rumelhart and McClelland 1986; McClelland and Rumelhart 1986). This was deliberately written, priced, and targetted to attract graduate students away from GOFAI and into the PDP stable. So it did--and it attracted many philosophers too. They valued it because it offered a more plausible account of concepts and conceptual similarity.

Crucially, the bible (alongside some lectures by Stephen Grossberg) also attracted the funding authorities. DARPA organized an urgent five-month review of their past funding policy, which had near-ignored connectionism for two decades. Although Minsky, one of the first invited speakers, refused to withdraw his 1969 criticisms (see Minsky and Papert 1988), the outcome was that DARPA changed their mind. They initiated "a major new program in neural networks beginning in 1989" (DARPA 1988: xxv), and gave Minsky and Papert a coded rebuke: "Neural network research is not new--it is, rather, newly revived from an obscurity and even disrepute which is now understood to have been undeserved" (DARPA 1988: 23). The twenty-year famine was over.

Here, we should note another runner-up for an eighth key date. To do that we have to backtrack seven years, to a masterpiece that paved the way for the connectionist renaissance: Douglas Hofstadter's *Gödel, Escher, Bach* (1979). This was an intoxicating document. It wove music, logic, biology, and Alice in Wonderland into a song of praise for AI/A-Life in general, and parallel distributed processing in particular. It became a cult book, winning the Pulitzer prize and appearing in many languages. (It's still much admired: in 1999 the *New Scientist* invited a dozen people to choose a science book from the last quarter-century to take to a desert island, and three chose this one.)

So why not add 1979 to our list without further ado? Well, for all its brilliance, *GEB* didn't outline a research programme that others could take up. However, it did raise the profile, and indicate the breadth, of cognitive science for the general public. Without its insightful flamboyance to ease the way, acceptance of the much dryer PDP bible would have been less immediate--and much less wide.

The last key date marks a further intellectual renaissance. In 1987 Christopher Langton organized the first conference on "artificial life", at Los Alamos. A-Life, he said, concerned "life as it could be", not just "life as we know it": abstract, preferably formal, descriptions of life were the goal. More generally, the focus was on self-organization and bottom-up processing, in various domains.

He circulated the invitation widely. In the event, a wide spectrum turned up: biologists, biochemists, physicists, mathematicians, AI researchers, neuroscientists, and philosophers (and the journalists turned up too). They discovered--as Langton had hoped--that, despite the superficial differences, they'd been working on closely related issues.

The interdisciplinarity and excitement rivalled the NPL meeting of 1958--and the 1950s Macy meetings of the cybernetics community, too. Indeed, that community was much in people's minds. Ross Ashby, Grey Walter, and Pask were honoured by their A-Life descendants after being near-forgotten for a generation. Now, they're familiar names in cognitive science.

Conclusion

And that, for a while, was that. It's not that nothing went on: cognitive science has continued to advance since 1987. And, increasingly, neuroscientific detail has been brought into formerly abiological zones. But nothing of comparable historical importance has occurred in the last twenty years.

Or rather, nothing that can be recognized *today* as having equal weight. There's plenty of new work out there that's promising, of course--including some which is truly fascinating, not run-of the- mill (see Boden 2006: ch. 17). A few of these examples may turn out to be historical high points. As yet, however, it's too early to tell.

References

Blake, D. V., and Uttley, A. M. (eds.) (1959), *The Mechanization of Thought Processes,* 2 vols. Proceedings of a Symposium held at NPL on 24-27 November 1958. (London: Her Majesty's Stationery Office).

Boden, M. A. (2006), *Mind as Machine: A History of Cognitive Science,* 2 vols. (Oxford: Oxford University Press).

Bruner, J. S., Goodnow, J., and Austin, G. (1956), *A Study of Thinking* (New York: Wiley).

Craik, K. J. W. (1943), *The Nature of Explanation* (Cambridge: Cambridge University Press).

DARPA (1988), *DARPA Neural Network Study: October 1987-February 1988* (Fairfax, Virginia: AFCEA International Press).

Gardner, H. (1985), *The Mind's New Science: A History of the Cognitive Revolution* (New York: Harper Collins).
Gregory, R. L. (1966), *Eye and Brain: The Psychology of Seeing* (London: Weidenfeld and Nicolson).
Hinton, G. E., and Anderson, J. A. (1981), *Parallel Models of Associative Memory* (Hillsdale, N.J.: Lawrence Erlbaum).
Hofstadter, D. R. (1979), *Godel, Escher, Bach: An Eternal Golden Braid* (New York: Basic Books).
McClelland, J. L., Rumelhart, D. E., and the PDP Research Group (1986), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition,* Vol. 2, *Psychological and Biological Models* (Cambridge, Mass.: MIT Press).
McCulloch, W. S., and Pitts, W. H. (1943), 'A Logical Calculus of the Ideas Immanent in Nervous Activity', *Bulletin of Mathematical Biophysics,* 5: 115-133.
Miller, G. A. (1956), 'The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information', *Psychological Review,* 63: 81-97.
Miller, G. A. (1986), 'Interview with George A. Miller', in B. J. Baars (ed.), *The Cognitive Revolution in Psychology* (London: Guilford Press), 200-223.
Miller, G. A., Galanter, E., and Pribram, K. H. (1960), *Plans and the Structure of Behavior* (New York: Holt).
Minsky, M. L. (1956), *Heuristic Aspects of the Artificial Intelligence Problem.* Group Report 34-55 (Lexington, Mass.: MIT Lincoln Laboratories, December). Revised as 'Steps Toward
Arti.cial Intelligence', *Proceedings of the Institute of Radio Engineers,* 49 (1961): 8-30. -9-
Minsky, M. L., and Papert, S. A. (1969), *Perceptrons: An Introduction to Computational Geometry* (Cambridge, Mass.: MIT Press).
Minsky, M. L., and Papert, S. A. (1988), 'Prologue: A View From 1988' and 'Epilogue: The New Connectionism', in *Perceptrons: An Introduction to Computational Geometry,* 2nd edn.(Cambridge, Mass.: MIT Press), viii-xv & 247-280.
Newell, A., Shaw, J. C., and Simon, H. A. (1957), 'Empirical Explorations with the Logic Theory Machine', *Proceedings of the Western Joint Computer Conference,* 15: 218-239.
Newell, A., Shaw, J. C., and Simon, H. A. (1958), 'Elements of a Theory of Human Problem- Solving', *Psychological Review,* 65: 151-166.
Newell, A., and Simon, H. A. (1972), *Human Problem Solving* (Englewood Cliffs, N.J.: Prentice- Hall).
O'Reilly, R. C., and Munakata, Y. (2000), *Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain* (Cambridge, Mass.: MIT Press).
Pitts, W. H., and McCulloch, W. S. (1947), 'How We Know Universals: The Perception of Auditory and Visual Forms', *Bulletin of Mathematical Biophysics,* 9 (1947), 127-147.
Place, U. T. (1956), 'Is Consciousness a Brain Process?', *British Journal of Psychology,* 47: 44-50.
Putnam, H. (1960), 'Minds and Machines', in S. Hook (ed.), *Dimensions of Mind: A Symposium* (New York: New York University Press), 148-179.
Rochester, N., Holland, J. H., Haibt, L. H., and Duda, W. L. (1956), 'Tests on a Cell Assembly Theory of the Action of the Brain, Using a Large Digital Computer', *Institute of Radio Engineers Transactions on Information Theory,* 2: 80-93.
Rosenblatt, F. (1958), 'The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain', *Psychological Review,* 65: 386-408.
Rosenblatt, F. (1962), *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms* (Washington, DC: Spartan).
Rosenblueth, A., Wiener, N., and Bigelow, J. (1943), 'Behavior, Purpose, and Teleology', *Philosophy of Science,* 10: 18-24.
Rumelhart, D. E., McClelland, J. L., and the PDP Research Group (1986), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition,* Vol.1, *Foundations* (Cambridge, Mass.: MIT Press).
Samuel, A. L. (1959), 'Some Studies in Machine Learning Using the Game of Checkers', *IBM Journal of Research and Development,* 3: 211-229. -10-

For a Systematic Theory of Expectations

Cristiano Castelfranchi (cristiano.castalfranchi@istc.cnr.it)
Institute for Cognitive Sciences and Technologies - CNR Rome, Italy

'More geometrico demonstrata'
Spinoza

Abstract

I analyze 'Expectation' as an amalgam of more elementary cognitive components (beliefs and goals). I claim that this produces a unitary 'mental states' with its specific functions. I explain the crucial role of expectations in choices, intentions, attempts, and as the background for several emotions like hope, fear, disappointment, relief. The fundamental role of mental 'anticipation' in the origin and nature of mind is stressed.

Cognitive Anatomy of Expectations

'Expectations' are not just 'Predictions'; they are not fully synonyms. And we do not want to use 'expectations' (like in the literature) just to mean 'predictions', that is, epistemic representations about the future. We consider, in particular, a 'forecast' [3] [4] as a mere belief about a future state of the world and we distinguish it from a simple 'hypothesis'. The difference is in terms of degree of certainty: a hypothesis may involve the belief that future p is possible while in a forecast the belief that future p is probable. A forecast implies that the chance threshold has been exceeded (domain of probability).

Putting aside the degree of confidence (we need a general term covering weak and strong predictions), for us 'expectations' have a more restricted meaning (and this is why computer can produce weather 'predictions' or 'forecasts' but do not have 'expectations'). In 'expectations'

- (i) the prediction is *relevant* for the predictor; he is *concerned, interested*, and that is why
- (ii) he is 'expecting', that is the prediction is aimed at being verified; he is *waiting* in order to know whether the prediction is true or not.[1]

Expectation is a suspended state *after* the formulation of a prediction[2]. If there is an expectation then there is a prediction, but not the other way around.

Epistemic Goals and Activity.

First of all, X has the Goal of knowing whether the predicted event or state really happens (epistemic goal). She is 'waiting for' this; at least for curiosity. This concept of 'waiting for' and of 'looking for' is necessarily related to the notion of expecting and expectation, but not to the notion of prediction.

Either X is actively monitoring what is happening and comparing the incoming information (for example perception) to the internal mental representation; or X is doing this cyclically and regularly; or X will in any case at the moment of the future event or state compare what happens with her prediction (epistemic actions) [14] [15]. Because in any case she has the Goal to know whether the world actually is as anticipated, and if the prediction was correct. Schematically [3]:

Expectation x p => Bel x at t' that p at t'' (where t'' > t') & Goal x from t' to t''' KnowWhether x p or Not p at t'' (t''' ≥ t''). This really is 'expecting' and the true 'expectation'.

Content Goals.

This Epistemic/monitoring Goal is combined with *Goal that p*: the agent's need, desire, or 'intention that' the world should realize [5] [6]. The Goal that p is true (that is the Goal that p) or the Goal that Not p. This is really why and in which sense X is 'concerned' and not indifferent, and also why she is monitoring the world. She is an agent with interests, desires, needs, objectives on the world, not just a predictor. This is also why computers, that already make predictions, do not have expectations.

When the agent has a goal opposite to her prediction, she has a 'negative expectation'; when the agent has a

[1] Notice that the first two meanings of 'to expect' in an English dictionary are the following ones:
1 to believe with confidence, or think it likely, that an event will happen in the future
2. to wait for, or look forward to, something that you believe is going to happen or arrive
While the definition of 'to forecast' is as follows:
1. to predict or work out something that is likely to happen, for example, the weather conditions for the days ahead
(Encarta® World English Dictionary © 1999 Microsoft Corporation).
Notice, the second component of 'expecting' meaning (absent in 'forecasting'): *wait for, or look forward to.* But also the idea that there is some 'confidence' in expectation: the agent *counts on* that.

[2] 'Prediction' is the result of the action of predicting; but 'expectation' is not the result of the action of expecting; it is that action or the outcome of a prediction relevant to goals, basis of such an action.

[3] We will not use here a logical formalization; we will just use a self-explanatory and synthetic notation, useful for a schematic characterization of different combinations of beliefs and goals. For a real formalization of some of these mental attitudes see [4].

goal equal to her prediction she has a 'positive expectation' (see § 3.1). [4]

In sum, Expectations (Exp) are axiological anticipatory mental representations, endowed with *Valence*: they are positive or negative or ambivalent or neutral; but in any case they are *evaluated against some concern, drive, motive, goal of the agent.* In Exp we have to distinguish two components:

- On the one side, there is a mental anticipatory representation, the belief about a future state or event, the "mental anticipation" of the fact, what we might also call the pre-vision (to for-see).

The format of this belief or pre-vision can be either propositional or imagery (or mental model of); this does not matter. Here just the function is pertinent.

- On the other side, as we just argued, there is a co-referent Goal (wish, desire, intention, or any other motivational explicit representation).

Given the resulting *amalgam* these representations of the future are charged of value, their intention or content has a 'valence': it is positive, or negative. [5] More precisely, Exp s can be:

- **positive** (goal conformable): $(\text{Bel x } p^{t'})^{t<t'}$ & $(\text{Goal x } p^{t'})$ [or $(\text{Bel x } \neg p^{t'})^{t<t'}$ & $(\text{Goal x } \neg p^{t'})$]
- **negative** (goal opposite): $(\text{Bel x } p^{t'})^{t<t'}$ & $(\text{Goal x } \neg p^{t'})$ [or $(\text{Bel x } \neg p^{t'})^{t<t'}$ & $(\text{Goal x } p^{t'})$]
- **neutral**: $(\text{Bel x } p^{t'})^{t<t'}$ & $\neg(\text{Goal x } p^{t'})$ & $\neg(\text{Goal x } \neg p^{t'})$ [or $(\text{Bel x } \neg p^{t'})^{t<t'}$ & $\neg(\text{Goal x } p^{t'})$ & $\neg(\text{Goal x } \neg p^{t'})$]
- **ambivalent**: $(\text{Bel x } p^{t'})^{t<t'}$ & $(\text{Goal x } p^{t'})$ & $(\text{Goal x } \neg p^{t'})$ [or $(\text{Bel x } \neg p^{t'})^{t<t'}$ & $(\text{Goal x } p^{t'})$ & $(\text{Goal x } \neg p^{t'})$]

The quantitative aspects of mental attitudes

Decomposing in terms of beliefs and goals is not enough. We need 'quantitative' parameters. Frustration and pain have an *intensity*, can be more or less severe; the same holds for surprise, disappointment, relief, hope, joy, ... Since they are clearly related with what the agent believes, expects, likes, pursues, can we account for those dimensions on the basis of our (de)composition of those mental states, and of the basic epistemic and motivational representations? We claim so.

Given the two basic ingredients of any Exp (defined as different from simple forecast or prediction) Beliefs + Goals, we postulate that:

P1: Beliefs & Goals have specific quantitative dimensions; which are basically independent from each other.

Beliefs have strength, a degree of subjective certainty; the subject is more or less sure and committed about their content [25]. *Goals have a value, a subjective importance for the agent.*

To simplify, we may have very important goals combined with uncertain predictions; pretty sure forecasts for not very relevant objectives; etc.

Thus, we should explicitly represent these dimensions of Goals and Beliefs: $\text{Bel}^{\%}$ x p^t; $\text{Goal}^{\%}$ x p^t

Where % in Goals represents their subjective importance or value; while in Beliefs % represents their subjective credibility, their certainty.

An Exp (putting aside the Epistemic Goal) will be like this:

$$\boxed{\text{Bel}^{\%} \text{ x } p^t \text{ \& Goal}^{\%} \text{ x } [\neg]\, p^t}$$

The subjective *quality* of those "configurations" or macro-attitudes will be very different precisely depending on those parameters. Also the effects of the invalidation of an Exp are very different depending on: (i) the positive or negative character of the Exp ; (ii) the strengths of the components. (See § 6.)

We also postulate that:

P2: The dynamics and the degree of the emergent configuration, of the Macro-attitude are strictly a function of the dynamics and strength of its micro-components.

For example anxiety will probably be greater when the goal is very important and the uncertainty high, than when the

[4] To be true a Goal equal to the prediction in Expectation is always there, although frequently quite weak and secondary relatively to the main concern. In fact, when X predicts that p and monitors the world to know whether actually p, she has also the Goal that p, just in order to not disconfirm her prediction, and to confirm to be a good predictor, to feel that the world is predictable and have a sense of 'control'. (see § 3.2). We are referring to *predictability*, that is, the cognitive component of self-efficacy [16]: the need to anticipate future events and the consequent need to find such an anticipation validated by facts. This need for prediction is functional in humans in order to avoid anxiety, disorientation and distress. Cooper and Fazio [17] have experimentally proved that people act in order to find their forecasts (predictions) validated by facts and feel distressed by invalidation.

[5] • Either, the expectation entails a cognitive evaluation [18]. In fact, since the realization of p is coinciding with a goal, it is "good"; while if the belief is the opposite of the goal, it implies a belief that the outcome of the world will be 'bad'.

• Or the expectation produces an implicit, intuitive appraisal, simply by activating associated affective responses or somatic markers [18]; or both;

• Or the expected result will produce a *reward* for the agent, and – although not strictly driving its behavior, it is positive for it since it will satisfy a drive and reinforce the behavior.

We analyze here only the Expectations in a strong sense, with an explicit Goal; but we mentioned Expectations in those forms of reactive, rule-based behaviors, first in order to stress how the notion of Expectation always involves the idea of a *valence* and of the agent being concerned and monitoring the world; second, to give an idea of more elementary and forerunner forms of this construct. It is in fact the case of proto-expectations or expectations in 'Anticipatory-Classifiers' based behaviors, strictly conceived as reactive (not really goal-driven) behaviors, but based on anticipatory representation of the outcomes [1] [2] [7] [13].

goal is not so crucial or the certainty is high.. Let us characterize a bit some of these emergent macro-attitudes.

Hope and Fear.

'Hope' is in our account [3] [4] a peculiar kind of 'positive Exp' where the goal is rather relevant for the subject while the Exp (more precisely the prediction) is not sure at all but rather weak and uncertain.[6]

$$\boxed{Bel^{low}\ x\ p^t\ \&\ Goal^{high}\ x\ p^t}$$

Correspondingly one might characterize being afraid, 'fear', as an Exp of something bad, i.e. against our wishes:

$$\boxed{Bel^{\%}\ x\ p^t\ \&\ Goal^{\%}\ x\ \neg p^t}$$

but it seems that there can be 'fear' at any degree of certainty and of importance.[7]

Of course, these representations are seriously incomplete. We are ignoring their 'affective' and 'felt' component, which is definitely crucial. We are just providing their cognitive skeleton [26].

The Implicit Counterpart of Expectations

Since we introduce a quantification of the degree of subjective certainty and reliability of Belief about the future (the forecast) we get a hidden, strange but nice consequence. There are other implicit opposite beliefs and thus implicit Exp s. For "implicit" belief we mean here a belief that is not 'written', is not contained in any 'data base' (short term, working, or long term memory) but is only potentially known by the subject since it can be simply derived from actual beliefs. For example, while my knowledge that Buenos Aires is the capital of Argentina is an explicit belief that I have in some memory and I have just to retrieve it, on the contrary my knowledge that Buenos Aires is not the capital of Greece (or of Italy, or of India, or of ...) is not in any memory, but can just be derived (when needed) from what I explicitly know. Until it remains implicit, merely potential, until is not derived, it has *no effect* in my mind; for example, I cannot perceive possible contradictions: my mind is only potentially contradictory if I believe that p, I believe that q, and p implies Not q, but I didn't derive that Not q.

Now, a belief that "70% it is the case that p", implies a belief that "30% it is the case that Not p"[8]. This has interesting consequences on Exps and related emotions. The Positive Exp that p, entails an implicit (but sometime even explicit and compatible) Negative Exp:

$$\boxed{\begin{array}{l} Bel^{\%}\ x\ p^t \\ \& \\ Goal^{\%}\ x\ p^t \end{array}} \rightarrow \boxed{\begin{array}{l} Bel^{\%}\ x\ \neg p^t \\ \& \\ Goal^{\%}\ x\ p^t \end{array}}$$

This means that a hope implicitly contains some fear, and that any worry implicitly preserves some hope. But also means that when one gets a 'relief' because a serious threat strongly expected is not arrived and the world is conforming to her desires, she also gets (or can get) some exultance. It depends on her focus of attention and framing: is she focused on her worry and evanished threat, or on the unexpected achievement? Vice versa when one is satisfied for the actual expected realization of an important goal, she also can get some measure of relief while focusing on the implicit previous worry. Not necessarily at the very moment that one feels a given emotion (for example fear) she also feels the complementary emotion (hope) in a sort of oscillation or ambivalence and affective mixture. Only when the belief is explicitly represented and one can focus – at least for a moment – her attention on it, it can generate the corresponding emotion.

Analytical Decomposition and the Gestalt Character of Mental Attitudes

Moreover, a hard problem for symbolic (and analytic) cognitive science deserves to be underlined: *the mental Gestalt problem.* Disappointment, expectation, relief, etc. seem to be unitary subjective experiences, typical and recognizable "mental states"; they have a global character; although made up of (more) atomic components, they form a *gestalt.* To use again the metaphor of molecules vs. atoms, the molecule (like 'water') has emergent and specific properties that its atoms (H & O) do not have. How can we account for this gestalt property in our analytic, symbolic, (de)composition framework? We have implicitly pointed out some possible solutions to this problem. For example:

- A higher-level predicate exists (like 'EXPECT') and one can assume that although decomposable in and implying specific beliefs and goals, this molecular predicate is used by mental operations and rules.
- Or one might assume that the left part of a given rule for the activation of a specific goal is just the combined pattern: belief + goal; for example, an avoidance goal and behavior would be

[6] We may also have – it is true - 'strong hope' but we explicitly call it 'strong' precisely because usually 'hope' implies *low* confidence and some anxiety and worry. In any case, 'hope' (like explicit 'trust') can never really be subjectively 'certain' and absolutely confident. Hope implies uncertainty.

[7] To characterize *fear* another component would be very relevant: the goal of avoiding the foreseen danger; that is, the goal of *doing* something such that Not p. This is a goal activated while feeling fear; fear 'conative' and 'impulsive' aspect. But it is also a component of a complete fear mental state, not just a follower or a consequence of fear. This goal can be a quite specified action (motor reaction) (a cry; the impulse to escape; etc.); or a generic goal 'doing something' ("my God!! What can I do?!") [27]. The more intense the felt fear, the more important the activate goal of avoidance [26].

[8] We are simplifying the argument. In fact it is possible that there is an interval of ignorance, some lack of evidences; that is that I estimate with a probability of 45% that p and with a probability of 30% Not p, while having a gap of 25% neither in favor of p nor of Not p [29] [30].

elicited by a serious negative *Exp* (and the associated 'fear'), not by the simple prediction of an event.

- One might assume that we "recognize" - or better "individuate" (and "construct")- our own mental state (thanks to this complex predicate or some complex rule) and that this "awareness" is part of the mental state: since we have a complex category or pattern of "expectation" or of "disappointment" we recognize and *have* (and feel) this complex mental state.

This would create some sort of "molecular" causal level. However, this might seem not enough in order to account for the gestaltic subjective experience, and reasonably something additional should be found in the direction of some typical "feeling" related to those cognitive configurations. Here we deal with the limits of any disembodied mind (and model).

Expectation: An 'Emergent' Mental Object and its Functions

Exps are new mental entities; they play a role as such, as global representations, as a gestalt, not just on the basis of their atomic components: beliefs and goals. Since in fact what matters is also the specific structure or relation (between Bel and G) which makes an Exp, makes it 'positive', 'negative', 'ambivalent', or 'indifferent', and makes its 'strength' which is neither reducible to the value of the goal, nor reducible to the certainty of the prediction.

Let us consider some of the main functions of Exp as a unitary mental representation:

Choices are Expectation-Based

A goal has a 'motivating force', which predicts the probability of its being pursued against costs and efforts, or chosen against other possible attractive goals (its 'priority'). However, the priority of the goal, its *motivating force* is not only due the subjective *'value'* of the goal: how important it is for us, how much it promises to us.

The *'value'* of a goal (desire, intention, objective, purpose, ...) either

- is not derived but just given if it is a 'terminal' (non instrumental) goal, an aim/end, the 'motive', ('given' of course for a given person in a given moment on the basis of its age, condition, personality, gender, culture, experience,..; or
- is derived from the value of the higher goals (to whom it is instrumental), and from the value of the goals that one has to sacrifice for achieving it: *costs* (invested resources) and *renounces*.

The motivating force and the priority of a goal is due both to its 'value' and to its estimated possibility. In other terms, what really matters in deliberation, what really prevails in choice, is not merely the goal with its 'value', but the 'Exp', with its new emergent metrics, which is the resultant of the goal-value and of the certainty of the prediction. It would be stupid (irrational) to give priority, to choose always the most valuable goal independently from its low possibility to be achieved; or - vice versa - to choose always the most probable result, independently from its marginal value. The right strategy is choosing the most valuable goal among the most probable goals; or the most probable goal among the most valuable ones.

Economic theory has proposed the SEU (the multiplication of Utility per Probability). This is a good mathematical solution for economics; but both the notion of 'utility' and the notion of 'probability' have serious problems for psychological theories.

Independently from the precise mathematical function (one might also think of several possible context-dependent heuristics for the choice) what matters here is the idea that the objects of a choice/deliberation, what is taken into account, are not Goals but complex and global Exps. We compare two Exps (not just two goals), and we are motivated in our intentional action by the Exp (the value and the likelihood).

This makes also more integrated and homogeneous the background of the candidate (or chosen) 'intention': the beliefs supporting and justifying on the one side the 'value' ascribed to the goal, and on the other side, the credibility/strength of the belief (prediction). Also because some beliefs might affect both of them. For example, the belief that the achievement of the goal is very close, on the one side increases the *certainty* of the Exp, but - on the other side - increases its importance (the *value* of the goal).

There are – of course - other situations or processes where not the Exp but just the *value* of the goal counts. This is the case, for example, in the degree of the 'frustration' (and consequent 'suffering'): the greater the value of the frustrated goal, the greater the pain. There might even be no Exp at all but just a (inactive) goal. (This of course does not mean that Exps do not play their own role in suffering; see later). In general, Goals have their specific and separated functions. For example goals, (mere goals not joined with any prediction) are used for ***evaluating*** the current state of the world (the match/mismatch step in cybernetic regulation of purposive behavior). The goal can be realized, not to be pursue, and thus without any anticipatory character, but it remains a goal (what one wishes, likes, wants, desires, …) while evaluating the world as 'good', as 'satisfying'; or while evaluating a 'success' (on the action). In that very moment only the goal counts; there is no probability estimation about its future achievement.

Analogously, mere predictions (beliefs about future states or events) can have their own specific functions, without any combination with motivational stuff, 'duties' or 'desires'. Like forecasts that we make for the others, even ignoring their specific goal.

Intentions are Expectation Driven

Decisions about future actions and the resulting 'intentions', and intentional actions presuppose an explicit Exp about the result. In order to decide to pursue and to pursue an intention the agent has to believe a lot of things

[8] [9] [20] [24] that the goal is not yet realized, that it is not impossible (it can be realized), that it is not self-realizing (by a natural process or by the forthcoming action of another agent), thus that it should be realized by the agent itself (it depends on the agent). Moreover, he has to believe that he knows the right action (plan), that is, that there is an action/plan producing that specific outcome, that he is able to correctly perform that action, and that there are the external conditions for a successful execution of that action. Only when/if the agent believes so he decides to pursue the goal by doing that action. But this obviously means that he *believes that the goal will be realized by his performing the action*; that is, the subject while intentionally acting (not just subjectively 'attempting') has *positive Exps* about the performance of the action and the realization of the goal. He is not intentionally 'trying', he is intentionally 'doing'.

We can formulate intentions only because we are able to build predictions relevant for and related to our desires.

Attempting

Another function of Exp (not of a mere goal) is the subjective 'attempt'. While, from the point of view of the observer, any intentional (or at least purposive) action actually is an 'attempt', since and until it cannot be sure that it will succeed, not any intentional action is an 'attempt' from the subjective point of view of its agent. Subjectively speaking one 'does' something, doesn't 'attempt to do' something. Or better, it is different when one 'does' something and when one just 'attempts' to do something [31]. It is different if one subjectively is 'paying' or 'closing the door' and when he is subjectively 'attempting to pay or to close the door'.

In order to subjectively just 'attempting to do' it is necessary that the agent explicitly conceives and takes into account the possibility of failure. He is not sure about the achievement of the goal of the action.

As we know (§3) any positive Exp (since is about the future and cannot really be 100% certain) logically/necessarily implies a negative one; and vice versa. However, we know that such complementary Exp can be merely 'implicit', 'potential', not really mentally formulated. In other terms, the subject can find satisfactory a belief that P with 80% of certainty, and fill fully certain, without considering (generating) at all the fact that there is a 20% possibility that Not p. Failure remains in his mind just a potential knowledge. Moreover, the subject can formulate for a moment the idea of a possible Not P (20%), but nevertheless he can put this aside, and do not take into account at all this eventuality in his reasoning and decision. This is why subjectively speaking not all our actions are 'attempts' and when we intentionally do something we are not 'trying'.

However, sometimes we really subjectively 'attempt' to do something. In this case, our mental representation is precisely the idea of the possibility of a failure. We have both a positive and a negative Exp; we are explicitly uncertain about the result. Thus, an attempt necessarily entails an Exp in the agent: an Exp not so sure about the positive result, implying some represented Exp of failure [9].

There also are attempts or better trials not really aimed at succeeding (but with some doubt), but just or mainly aimed at learning. The agent acts in order to see whether (the door is open or not) or to discover how (the door opens). The epistemic function, which is present in any action and especially (consciously) in any attempted action, here is dominant or is the only real goal.

Sustaining Persistence (Waiting for Rewards)

Another interesting function of the Exp as such, as a whole, is the fact that entertaining an Exp in mind (especially a sensory-motor representation of a desirable, pleasant state) seems to be useful for our capacity of *delaying* the realization of our the desires (Freud), although *persisting* in a prolonged activity, and paying costs, or persisting just in waiting for something, without receiving rewards (except from our imagination) (DESIRES). Long term planning is a fundamental capacity of humans, and the needed persistence and coherence [23] is neither due simply to 'predictions' (belief) per se', nor just to the goal. The goal without the Exp cannot *sustain* and *support* the effort toward the future; the belief per se' has no motivational power.

This is also why one of the worst forms (and causes) of suffering [28] is not just that our goal is frustrated but that this is 'forever'. That is when we do not only see our goal destroyed, but also wasted any possible 'hope'. We cannot have any (although weak) Exp about a future realization of our goal; the world doesn't simply answer "No!", it answers "Never!".

We can cope with a failure or a loss also thanks to the 'consolation' that at least one day it will be possible (again)

Expectations and Suffering

Expect can make suffering worst. If not only a given goal is frustrated, but there also was a joined prediction which is invalidated (was wrong), in other words, if there was not simply a goal but a full Expt, then the sufferance is worst (given the same value of the goal). To the frustration it is added the 'disappointment' (see later), which is an additional dimension of sufferance; either because you were already enjoying the desired result (it was already 'yours'), and you perceive this failure more as a loss than as a simple missed gain; or because also the meta-goal of being a good predictor is frustrated; or because not only you do not get the price, but you also get *less than* expected.

Emotional Response to Expectation is Specific: the Strength of Disappointment

As we said, also the effects of the *invalidation* of an expectation are very different depending on: a) the positive or negative character of the expectation; b) the strengths of the components. Given the fact that X has previous expectations, how this changes her evaluation of and reaction to a given event?

Invalidated Expectations

We call invalidated expectation, an expectation that results to be wrong: i.e. while expecting that p at time t', X now beliefs that NOT p at time t'.

$(\text{Bel } x\ p^{t'})^{t<t'} <==> \mathbf{(Bel\ x\ \neg p^{t'})^{t''>t}}$

This crucial belief is *the 'invalidating'* belief.

- Relative to the goal component it represents "frustration", "goal-failure" (is the *frustrating* belief): I desire, wish, want that p but I know that not p.

FRUSTRATION: $(\text{Goal } x\ p^{t'})$ & $(\text{Bel } x\ \neg p^{t'})$

- Relative to the prediction belief, it represents 'falsification', 'prediction-failure':

INVALIDATION: $(\text{Bel } x\ p^{t'})^{t<t'}$ & $(\text{Bel } x\ \neg p^{t'})^{t''>t}$

$(\text{Bel } x\ p^{t'})^{t<t'}$ represents the former illusion or delusion (X illusorily believed at time t that at t' p would be true).

This configuration provides also the cognitive basis and the components of "**surprise**": *the more certain the prediction the more intense the surprise.* [10] [11] Given positive and negative Expectations and the answer of the world, that is the *frustrating* or *gratifying* belief, we have:

	P	**¬P**
Bel x p & Goal x p	No surprise + achievement	*surprise* + *frustration* ***disappointment***
Bel x ¬p & Goal x p	*surprise* + *non-frustration* ***relief***	no surprise + frustration

Disappointment. Relative to the whole mental state of "positively expecting" that p, the *invalidating&frustrating* belief produces "disappointment" that is based on this basic configuration (plus the affective and cognitive reaction to it):

$$\text{DISAPPOINTMENT: } (\text{Goal}^{\%}\ x\ p^{t'})^{t\ \&t'}\ \&$$

$$(\text{Bel}^{\%}\ x\ p^{t'})^{t}\ \&\ (\text{Bel}^{\%}\ x\ \neg p^{t'})^{t'}$$

At t X believes that at t' (later) p will be true; but now – at t' – she knows that Not p, while she continues to want that p. Disappointment contains goal-frustration and forecast failure, surprise. It entails a greater *sufferance* than simple frustration [28] for several reasons: (i) for the additional failure; (ii) for the fact that this impact also on the self-esteem as epistemic agent (Badura's "predictability" and related "controllability") and is disorienting; (iii) for the fact that losses of a pre-existing fortune are worst than missed gains (see below), and long expected and surely expected desired situation are so familiar and "sure" that we feel a sense of loss.

The stronger and well-grounded the belief the more disorienting and restructuring is the *surprise* (and the stronger the consequences on our sense of predictability). The more important the goal the more *frustrated* the subject.

In Disappointment these effects are combined: *the more sure the subject is about the outcome & the more important the outcome is for her, the more disappointed the subject will be.*

- Te degree of disappointment seems to be a function of both dimensions and components [9]. It seems to be felt as a unitary effect.

"How much are you disappointed?" "I'm very disappointed: I was sure to succeed"

"How much are you disappointed?" "I'm very disappointed: it was very important for me"

"How much are you disappointed?" "Not at all: it was not important for me"

"How much are you disappointed?" "Not at all: I have just tried; I was expecting a failure".

Obviously, worst disappointments are those with great value of the goal and high degree of certainty. However, the *surprise* component and the *frustration* component remain perceivable and a function of their specific variables.

Relief. Relief is based on a 'negative' expectation that results to be wrong. The prediction is invalidated but the goal is realized. There is no frustration but surprise. In a sense relief is the opposite of disappointment: the subject was "down" while expecting something bad, and now feel much better because this expectation is invalidated.

RELIEF: $(\text{Goal } x\ \neg p^{t'})$ & $(\text{Bel } x\ p^{t'})$ & $(\text{Bel } x\ \neg p^{t'})$[10]

- *The harder the expected harm and the more sure the expectation (i.e. the more serious the subjective threat) the more intense the 'relief'.*

More precisely: the higher the worry, the threat, and the stronger the relief. The worry is already a function of the value of the harm and its certainty.

Analogously, **joy** seems to be more intense depending on the value of the goal, but also on how *unexpected* it is.

A more systematic analysis should distinguish between different kinds of surprise (based on different monitoring activities and on explicit vs. implicit beliefs), and different kinds of disappointment and relief due to the distinction between 'maintenance' situations and 'change/achievement' situations.

More precisely (making constant the value of the Goal) the case of loss is usually worst than simple non-achievement. This is coherent with the theory of psychic suffering [28] that claims that pain is greater when there is not only frustration but disappointment (that is a previous Exp), and

---

[9] As a first approximation of the degree of Disappointment one might assume some sort of multiplication of the two factors: Goal-value * Belief-certainty. Similarly to 'Subjective Expected Utility': the greater the SEU the more intense the Disappointment.

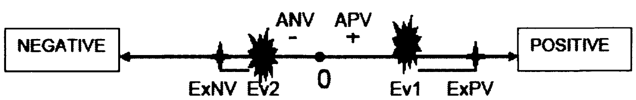
[10] Or – obviously - (Goal x pt') & (Bel x ¬pt') & (Bel x pt').

when there is 'loss', not just 'missed gains', that is when the frustrated goal is a maintenance goal not an achievement goal. However, the presence of Exps makes this even more complicated.

Level of expectation: how to be unhappy with positive results

The *level of Exp* also plays a very important role. In fact after having Exps (with a given expected *value*) the appreciation of the outcome is no longer *absolute*: good or bad, achieved goal vs. frustrated goal, failure (or at most the evaluation of the degree of the achievement/frustration: fully vs. partially achieved). The appreciation of the outcome becomes *relative* to the expected outcome.[11] This also has not so nice consequences, like the possibility to find unsatisfactory even good results (if inferior to the Exps).

What matters in fact is not only if the outcome is positive or negative, but if it better or worst of the desired and predicted level. Suppose a polarity of good (pleasant)/bad (unpleasant) results; and suppose now that we have a given positive Exp (Expected positive value – ExPV) or a given negative Exp (Expected negative value – ExNV). Given this and given the positive result of Event 1 (Ev1) or the negative result of Event 2 (Ev2), we get both an absolute Actual positive (APV) or Actual negative value (ANV) of Ev1 and Ev2 (relative to the 0 point), but also a relative value of Ev1 or Ev2 relatively to their Exp levels.



The interval (ExPV – APV) gives us the measure of the 'disappointment', 'discontent' even with a positive result. The APV can give us the measure of a possible 'consolation' ("nevertheless the result is quite good").

On the negative side, the interval (ExNV – ANV) gives us the level of the 'relief,' even with a negative event. The APV gives us the level of absolute frustration, but (ExNV – ANV) (if ANV is less than ExNV) can give a sort of 'consolation': "It might have been much worst!").

Of course, the APV can be greater/better than the ExPV; and in this case there is surprise and joy; while the ANV can be greater/worst than the ExNV; and in this case we get a higher degree of frustration than just due to ANV; the ANV is made worst by the fact that it is even worst than expected. Pessimistic Exps in part protect us from frustration and disappointment; while too optimistic Exps can expose to frustration even with good results.

In other words, we have to cross two dimensions of evaluation of the results: on the one side if they are good or bad (realized goals or frustrated goals), on the other side if they are better or worst than expected.

RESULTS	GOOD	BAD
< Expectation	Disappointment	Relief
> Expectation	?? "Whoow!"	?? "Not so bad"

Relief: less bad than expected; *Disappointment*: less good than expected; *Consolation*: Although bad, at least something good.[12]

All this is due not just to our goals and their values, but to the fact that we *expect* certain outcomes

Concluding remarks

In conclusion, Exps are composite and hybrid mental representations with an epistemic component or attitude and a motivational component and attitude about the same content. But in fact this mental representation is a new unitary mental object with its own specific uses and functions, and gives rise to typical and new psychic phenomena. Like the activity of 'expecting for', the possibility of complex rational decisions based on the comparison not simply between two goals with their importance, or specific emotional states (hope, trust, fear, worries, …) and emotional reactions due to the pre-existence of such a state (relief, disappointment,..).

Expectations play a major role in the pressure for the origin of mind with its crucial anticipatory nature (cit), and primitive expectations - related to actions as anticipated rewarding perceptual inputs for monitoring and learning - are a fundamental step towards the evolution of true goal-directed (purposive) systems.

One of the aims of this contribution is to show that there is room for some sort of 'theoretical psychology', where an analytical and formal modeling is supposed to provide important insights and predictions, and produces indications and interpretations for empirical research. In other terms, we attempt to modestly follow the old arrogant program of

[11] We might call the Goal within a positive Exp 'aspiration' and 'aspiration level' its expected degree of realization combined with its subjective Value.

[12] The theory of 'relief' and of 'disappointment' is even more complicated. They are in fact 'counterfactual' emotions. They are based on the idea (imagination) of what *might have been/happen.* Relief is when what actually is now (what has happened) is better of what could have been; Disappointment is when the actual situation is worst than the possible one. Relief and disappointment due to a previous (bad/good) Exp are just sub-cases of this. In fact, if X expects/forecasts that P, this implies that he was considering P possible, probable; also after that P didn't in fact become true. Moreover, at least 'relief' is also possible simply at the end of an actual and present pain or sufferance. This is coherent with this analysis; since a current experienced sufferance – when finished – entails the (implicit) belief, the a-posteriori Exp, that it could have continued, and the relief is due to this possible but falsified continuation.

Spinoza about emotions (and mind) "more geometrico demostrata".

Acknowledgments

European Project *MindRaces* N°511931- EC's 6th Framework Programme - Unit: Cognitive Systems; & European HUMAINE network. I would like to tank Maria Miceli, Emiliano Lorini which have developed with me the general theory of expectations and of related emotional states, and their formalization; Giovanni Pezzulo for our view of anticipation and mind; Luca Tummolini with whom I have developed the social and normative aspects of expectations; Michele Piunti which has developed a computational model of this and of surprise and cautiousness.

References

Bandura A., (1990). Self-efficacy mechanism in human agency. *American Psychologist*, 37, pp. 122-147, 1990.

Bratman M. E., (1988). Intentions, plans, and practical reason, Cambridge, MA: Harvard University Press, 1988.

Butz, M.V. & Hoffman, J. (2002) Anticipations control behavior: Animal behavior in an anticipatory learning classifier system. *Adaptive Behavior*, 10, 75-96.

Butz, M.V. (2002) *Anticipatory learning classifier system* Boston, MA: Kluwer Academic Publisher.

Castelfranchi C., (1996). Reasons: Belief Support and Goal Dynamics. *Mathware & Soft Computing, 3*. pp. 233-47, 1996.

Castelfranchi, C. (2005) Mind as an Anticipatory Device: For a Theory of Expectations, *Brain, Vision and Artificial Intelligence, 1st International Symposium* (BV&AI 2005), eds. M. De Gregorio, V. Di Maio, M. Frucci and C. Musio, Springer-Verlag, Berlin, 2005, pp. 258-276.

Castelfranchi, C., Lorini E., (2003). Cognitive Anatomy and Functions of Expectations. In *Proceedings of IJCAI'03 Workshop on Cognitive Modeling of Agents and Multi-Agent Interactions*, Acapulco, Mexico, August 9-11, 2003.

Castelfranchi, C., Paglieri, F. (2007). "The role of beliefs in goal dynamics: Prolegomena to a constructive theory of intentions". *Synthese*, in press. (DOI: 10.1007/s11229-006-9156-3).

Castelfranchi, C., Tummolini, L. and Pezzulo, G. (2005) From Reaction to Goals – AAAI Ws on From Reaction to Anticipation,. 2005

Cohen, P. R., Levesque H. J., (1990). Intention is choice with commitment. *Artificial Intelligence, 42*, pp. 213-261, 1990.

Cooper, J., Fazio R. H., (1984). A new look at dissonance theory. In L. Berkovitz (Ed.), *Advances in experimental social psychology, Vol. 17*, pp. 229-266, San Diego, CA: Academic Press, 1984.

Corrêa, M., Coelho, E., Agent´s programming from a mental states framework. In Proceedings of the 14th Brazilian Symposium on Artificial Intelligence (SBIA98), Lecture Notes in AI 1515, pp. 31-39, Springer-Verlag, 1998.

Drescher, G. (1991) *Made-up minds: A constructivist approach to artificial intelligence*. MIT Pres.

Galliers, J.R. (1991). Modelling Autonomous Belief Revision in Dialogue, In *Decentralized AI-2*, Y. Demazeau, J.P. Mueller (eds), 231-43. Armsterdam: Elsevier.

Jones, O. R. (1983). Trying. *Mind*, XCII(367):368-385.

Kahneman, D., Miller D. T., (1986). Norm Theory: Comparing reality to its alternatives. *Psychological Review*, 93, pag. 136-153, 1986.

Kirsh, D., Maglio. P., On distinguishing epistemic from pragmatic action. *Cognitive Science, 18*, pp. 513-549, 1994.

Lorini, E. and Castelfranchi, C. (2007). The cognitive structure of surprise: looking for basic principles. *Topoi: an International Review of Philosophy*, (forthcoming).

Lorini, E., Castelfranchi C., (2004). The role of epistemic actions in expectations. In *Proceedings of Second Workshop of Anticipatory Behavior in Adaptive Learning Systems 2004 (ABIALS 2004)*, Los Angeles, 17 July 2004.

Lorini, E., Herzig, A., and Castelfranchi, C. (2006). Introducing Attempt in a modal logic of intentional action. Michael Fisher, Wiebe van der Hoek (Eds.), 10th European Conference on Logics in AI (JELIA06), Springer-Verlag, LNAI, p. 280-292.

Miceli, M. & Castelfranchi, C. (1997). Basic principles of psychic suffering: A prelimirary account. *Theory & Psychology*, *7*, 769-798.

Miceli, M. & Castelfranchi, C. (2005). Anxiety as an "epistemic" emotion: An uncerainty theory of anxiety. *Anxiety, Stress, and Coping*, *18*, 291-319.

Miceli, M. & Castelfranchi, C. (2006) Hope: The power of wish and possibility. (Submitted).

Miceli, M. & Castelfranchi, C. (2000). The role of evaluation in cognition and social interaction. In K. Dautenhahn (Ed.), *Human cognition and agent technology*. Amsterdam: Benjamins, 225-61.

Miceli, M., Castelfranchi. C., (2002) The Mind and the Future. The (Negative) Power of Expectations. *Theory & Psychology*, 12(3), pp. 335-366, 2002.

Miller, G., Galanter E., and Pribram. K. H., (1960) Plans and the structure of the behavior. Rinehart & Winston, New York.

Ortony, A., Partridge. O., (1987) Surprisingness and expectation failure: What's the difference? In *Proceedings of the 10th International Joint Conference on Artificial Intelligence*, pp. 106-108, Los Altos, CA: Morgan Kaufmann, 1987.

Pezzulo, G., Lorini, E., Calvi G. (2004). How do I know how much I don't know? A cognitive approach about Uncertainty and Ignorance. In *Proceedings of 26th Annual Meeting of the Cognitive Science Society (CogSci 2004)*, Chicago, USA, 5-7 August, 2004

Rao, A.S., Georgeff M.P., (1992). An abstract architecture for rational agents. In Proceedings of the Third International Conference on Principles of Knowledge Representation and Reasoning, C. Rich, W. Swartout, and B. Nebel (Eds.), pp. 439-449, Morgan Kaufmann Publishers, San Mateo, CA, 1992.

Rosenblueth, A., Wiener N., and Bigelow. J., (1960) Behavior, Purpose, and Teleology. In W. Buckley (Ed.), Modern Systems Research for the Behavioral Scientist, Aldine, Chicago.

Shafer G., (1976). *A mathematical theory of evidence*. Princeton University Press, Cambridge, 1976.

Language Activity as a Representational Activity: Typological Approaches to Comparison

Catherine Fuchs (catherine.fuchs@ens.fr)
Lattice, ENS, 1 rue Maurice Arnoux, 92120 Montrouge, France

Abstract

Within the field of cognitive science, linguistics has to account both for the diversity of semantic representations construed by various languages and for their unity (as regards their mapping into universal conceptual representations). This issue is illustrated here by examples from typological studies on the diversity of linguistic expressions of comparison.

Introduction

Language ability is part of the human nature; being somehow rooted in the human brain, it is an object of science for neurobiologists and experimental psychologists. Linguists for their part are interested in languages: by analyzing languages (*"les langues"*), they try to enlighten general properties of language (*"le langage"*). In other words, the universal faculty of language is, no doubt, an innate property of human species, but linguists are faced with language diversity, not with language universality.

Now, if we assume that *(a)* natural language is a mechanism that connects mental or conceptual representations to syntactic (and ultimately phonetic) forms, and *(b)* the conceptual representations that underlie non-linguistic thinking are universal, being part of our biological endowment, how then can we describe the connection between universality and diversity? Various answers have been put forward, ranging from Chomsky's 'universal grammar' (focusing on syntactic structures) and Fodor's 'language of mind', to different types of 'cognitive grammars' (mainly concerned with semantic structures).

Different Levels of Representation

The notion of 'representation' can be misleading: different levels of representation must be distinguished (Culioli, 1995).

Conceptual Representations. The conceptualization of reality (objects and events) deriving from our perceptions, tastes, dislikes, collective representations, *etc.*, is part of human cognitive activity: it is the level of conceptual representations (CRs), to which we have no direct access other than through our actions, including our language activity.

Semantic Representations. Through language activity, utterances (the only observable language phenomena) are produced in communication acts. Utterances are concatenations of 'markers' (signs) - the relationship between the *signifiants* and the *signifiés* being specific to each particular language. Consequently, utterances give birth to linguistic (semantic) representations (SRs) that stand for mental representations, but do not code them univocally in term for term relationships: SRs and CRs are not isomorphic (see also Levinson, 1997).

Thus the task of linguistics is to elaborate metalinguistic representations of the SRs of particular languages; such metalinguistic representations consist of rules and operations, which should be subject to generalization. For simplicity's sake, I will not concern myself with the formal aspect of metalinguistic representations, which will be accounted for in terms of simple glosses.

Linguistics within the Field of Cognitive Science

In my view (Fuchs, 1999), the main task of linguistics within the field of cognitive science is to provide a description of semantic representations that underlie linguistic meaning in various languages, and to account both for their diversity (since they make different choices among various conceptualizations of situations) and for their unity (since they are supposed to correspond ultimately to unique mental structures). For, if languages were direct codes of non-linguistic thinking, they would all be similar: language diversity implies the existence of an intermediate level of variable (semantic) representations between the level of mental representations and that of superficial structures.

Such investigations have been carried out at large in a number of lexical domains (colours, numbers, …) and grammatical domains (space, temporality, possession, actancy, …). But other domains, which obviously do not fall within the scope of elementary categories, may also be of interest: such is the case of 'comparison', which I will take as an illustration of the methodological and theoretical problems faced by linguistics in *(1)* studying the representational activity at work in SRs of particular languages (*i.e.* looking for linguistic operations with cognitive impact), and *(2)* trying to pinpoint similarities (as well as differences) between SRs of various languages (*i.e.* looking for cross-linguistic invariants), in order to *(3)* discover some general properties of language itself, concerning the link between SRs and CRs.

Looking for Linguistic Operations in Particular Languages

As regards the grammatical category of comparison, cognitively significant observations made on single instances of language, like English or French (but undoubtedly subject to further generalizations) may be summarized as follows.

Prototypical Representation and Basic Structures

Quantification vs. Qualification. Two main types of comparison are to be distinguished, namely 'quantitative' and 'qualitative' comparison. The former implies grading (two items being graded against each other, regarding a given property, ex: *The tower is higher than the house*), as opposed to the latter (two items being compared as to their manner of performing a given action, or of verifying a given property, ex: *He swims like a duck*).

The Prototypical Representation (Stassen, 1985) is concerned with comparison between two objects or individuals (typically expressed in the form of NPs) — named the 'comparee' and the 'standard' — with respect to a given 'parameter' (typically expressed in the form of a verbal or adjectival predicate). This representation presumably reflects some kind of cognitively-based process: assigning a graded position on a predicative scale (for instance, the scale of *heaviness*) respectively to two objects (for instance, two *stones*), in order to make a relevant choice before performing a given action; or evaluating the similarity of two objects (*e.g.* two *stones*) with respect to a given capacity (*e.g.* their capacity of *hurting*) or quality (*e.g.* their *roundness*).

Identity vs. Non-Identity. There appear to be only two types of relations (Rivara, 1975, 1995): either identity (=) or non-identity (≠). Equality (*i.e.* quantitative identity) and inequality (*i.e.* quantitative non-identity) operate on subjectively oriented scales: either on the scale of great quantities (*as = 'as much, as many'; more / -er = 'more much, more many'*) or on the scale of small quantities (*as little, as few; less = 'less little, less few'*) — which means that, contrary to logical relations, semantic relations of quantitative comparison are not symmetrical: *John is more friendly than Peter = 'John surpasses Peter in friendliness on the scale of great quantities' ≠ Peter is less friendly than John = 'Peter surpasses John in friendliness on the scale of small quantities'.*

Basic Structures. All possible constructions are not equally frequent: inequalities and equalities operating on the 'great quantities' scale are more frequently expressed than those operating on the opposite scale; and, concerning qualitative comparison, similarity (*i.e.* qualitative identity, expressing sameness of manner) is more frequent than dissimilarity (*i.e.* qualitative non-identity, expressing difference of manner). Consequently, the three basic structures — probably the most cognitively salient — are the following:

(i) inequality on the scale of great quantities (known as 'superiority'): *Mary is prettier than Jane*
(ii) equality on the scale of great quantities: *Mary is as pretty as Jane*
(iii) similarity: *Mary sings like a nightingale.*

Non-Basic Structures

Now, if one looks more closely into the system of widely described languages like English (or French), one can easily notice various extensions of these basic structures, leading to more abstract and complex types of representations that could hardly be conceivable without the help of language.

The Parameter can be extended to a 'secondary quality' — e.g. an adverb qualifying the main predicate (*Mary sings louder than Jane; Mary sings as loudly as Jane*) or to a 'secondary predicate' (*Being prettier than Jane, Mary won; Being as expensive as John's, Peter's car is likely to be stolen*).

The Compared Items can be other than two objects or individuals. Comparison markers are also used to express comparisons between two circumstances (*It's colder today than yesterday; The weather is as cold in Paris as it is in London*), two properties (*Mary is more cunning than intelligent; Jane is as pretty as she is intelligent*), two modalities (*Mary is prettier than I thought; The sky is as sunny as I hoped*), two events (*Jane loves her son, more than you do yours; Mary bought a flat, as you did a house*), or even two 'enunciations' (*"P", as they say).*

Marker-Operation Relationships

Just like any detailed study of a given grammatical category, the study of comparison must take into account the non-univocal relationships between markers and operations: that is, polysemy on one side, and paraphrase on the other. Let's take the example of the French marker *comme*. It is a polysemous item, which covers a large range of meanings (Fuchs & Le Goffic, 2005): similarity (*chanter comme un rossignol: "to sing like a nightingale"*), temporal simultaneity (*Il arriva comme je partais: "He arrived just as I was leaving"*), a kind of inference (*Comme je ne suis pas pressé, je vais attendre: "Since I am in no hurry, I will wait"*), exclamation (*Comme elle est jolie!: "How pretty she is!"*), *etc.* And as a marker of comparison, it has a number of 'quasi-synonyms' (Fuchs, 2007): *ainsi que, de même que, à la façon de, à la manière de, etc.* Such multiple correlations, which are specific to a given language, speak for the relative autonomy of SRs as procedures of meaning construction.

Looking for Cross-Linguistic Invariants

This is where typologists come in. Working on data from extensive samples of historically unrelated languages, they classify languages (on a structural basis, not on a genetic one) and formulate generalizations supplementing the regularities discovered in the study of single instances of language. The pioneering research in linguistic typology was devoted to word-order and morpho-syntax.

For many years, comparison was considered so central that the word order patterning of a specific language was supposed to be determined by the order of elements in a comparative construction — *cf.* the so-called 'implicational' (or relational) 'universal 22' in Greenberg (1966); see also Lehmann (1972). Although the validity of that claim has been questioned later on (Andersen, 1983), it seems that the prototypical inequality comparative is "the most secure of constructions" regarding word-order, since "it is never changed for poetic effect" (Lehmann, 1973). This stability in word-order within each particular language gives evidence that the comparative SRs – and especially the SRs of inequality - are deeply rooted in languages, and meant to express some fundamental cognitive processes.

More recent typological studies have been concerned with the semantics of comparison. Most of them confine themselves to describing the basic structures where two objects or individuals are being compared, for the grammars of their sample languages generally do not provide sufficiently reliable data on other more complex constructions.

Semantic Variations

The methodology consists, first in observing the cross-linguistic variations (lexical, syntactic and semantic variations) from a 'semasiological' point of view (*i.e.* from forms to meanings), and gathering the various representations into several 'types'; and only then in trying to recover invariants: "*In order to understand the grammar of comparative constructions, not much is gained in looking for one uniform universal structure; rather what is required is that the entire pool of possible conceptual sources be considered*" (Heine, 1997). Actually, there are numerous surface constructions available across languages, which can be reduced to a small number of representation types. The markers involved in these representations are generally used for a number of different grammatical categories and not exclusively for expressing comparison – which indicates that they represent an inventory of the possible sources to choose from.

The three main typological works I am referring to are: Stassen (1985) - who studied inequality in 110 different languages, Haspelmath & Buchholz (1998) - who examined equative and similative constructions in 47 European plus several non-European languages, and Henkelmann (2006) – who studied equative constructions in 25 languages all over the world. To summarize briefly the results of these works: comparatives turn out not to be independent, autonomous construction types, but to derive from more basic representations. For instance, it should be noticed that many languages do not resort to relative degree words (such as French *plus* or English *more*) to encode quantitative comparison – in these languages, the notion of grading results from other types of semantic operations.

Main Types of Inequality SRs. A limited number of types (underlying SRs) can be determined, which correspond to various semantic 'strategies', *i.e.* to different choices made by languages among elementary conceptual sources. The main 'schemas' that happen to be used for encoding inequality – equivalent to English *A is bigger than B* (Stassen, 1985; Heine, 1997) – are the following:

(a) 'Action Schema', glossed: *A big EXCEEDS B* (surface variants: *A is big surpasses B / A is big to surpass B / B is big (but) A exceeds / A surpasses B (at) bigness*)

(b) 'Spatial Schemas', which can be subdivided into three subtypes:

- 'source (or separative) schema', glossed: *A big FROM B* (by far the most common spatial schema expressing comparison in the languages of the world; surface variants: ablative or genitive adverbial phrases)
- 'locative schema', glossed: *A big AT B* (second best spatial schema expressing comparison; surface variants: *A is big on / above / in / by / ... B*)
- 'allative (or goal) schema', glossed: *A big TO B* (rather rare; surface variants: allative / benefactive / dative / ... adverbial phrases)

(c) 'Conjoined Schemas', which can be subdivided into two subtypes:

- 'polarity schema', glossed: *A big (AND) B not big* (either positive/negative polarity: *A is big and/but/while B is not big* – variant *B is not big and/but/while A is big* - or antonymy: *A is big and/but/while B is small* – variant *B is small and/but/while A is big*)
- 'sequence (temporal) schema', glossed: *A big THEN B* (surface variants: *A is big and/and then/thereafter/... B*)

It should be noted, incidentally, that from a typological point of view, French constructions (*A est plus grand que B*) and English ones (*A is bigger than B*), are difficult to classify, for they "*have been grammaticalized to such an extent that the cognitive schema underlying them is not readily reconstructible*" (Heine, 1997). Some typologists consider that such constructions correspond to a specific 'particle schema'; other tentatively analyze them as resulting from a process of syntactization, which could have led to the transformation of a coordinate clause (corresponding to a 'conjoined schema' – whether a 'polarity' or a 'sequence' one) into a subordinate clause.

Main Types of Equality SRs. A limited number of underlying schemas has also been listed for the encoding of equality (Henkelmann, 2006):

(a)'Extents Schema', glossed: *A (as) big (as/like) B* (equality relation is established between the relative extents of the quality that is being attributed to the entities being compared)

(b) 'Entities Schema', glossed: *A big EQUALS B* (equality relation is established between the entities directly by means of an equative predicate; *cf.* inequality 'action schema' *supra*)

(c) 'Possessive Schema', glossed: *A's bigness is B's bigness* (the extent of the quality that is said to be equal is represented as a possession of the entities being compared)

(d) 'Representative Schema', glossed: *A is (of) B's bigness* (the entities being compared are represented as instances or representatives of an equal quality appearing as an entity)

(e) 'Implicit Schema', glossed: *A big, B big* (two predications in the positive degree are juxtaposed without any explicit encoding of the notion equality; *cf.* inequality 'conjoined schema' *supra*).

Invariants

The first observation to be made is that underlying comparative SRs are neither indefinite nor random: they are in a relatively small number and seem to derive from cognitively motivated basic schemas. The different ways of encoding quantitative comparison thus illustrate different choices made by languages within a limited 'repertoire' of possible conceptual sources.

From SRs to CRs. As has been pointed out by Levinson (1997), the relation between SRs and CRs is a matter for empirical investigation. There seem to be good psycholinguistic and neurolinguistic reasons to assume that on the CRs level, the cognitive structure of comparison is in the form of a spatial global configuration: the parameter being pictured as an axis, marked for positive-negative polarity; the two objects being positioned on the axis so that their positions define extents, which represent the degree to which the compared objects possess the quality at issue - in other words, relative degrees of intensity with respect to a certain quality being represented in terms of relative distance on the axis. According to Stassen, the mapping of this CR into various SRs could be seen as a transition of such a spatially modelled configuration into a configuration which is modelled on 'temporal chaining' (*i.e.* temporal ordering between two events): strategies 'read off' various bits of information from the CR and codify them in the form of a sequence of two propositions. Thus, the three types of cognitive strategies distinguished by Stassen could fit inequality as well as equality schemas :

(a) The 'Independent Strategy', where the axis is taken as the salient feature of the CR, so that the compared items A and B are associated with the opposite sides of the axis – thus leading to the 'polarity schema' of inequality and to the 'implicit schema' of equality.

(b) The 'Ordered Strategy', where the salient feature is provided by the extents demarcated on the axis, so that A and B are both associated with the positive side of the axis – thus leading to the 'action schema' for inequality and to the 'extents schema' and 'entities

(c) the 'Relative Strategy', where the spatial relation between A and B is the salient feature, so that only one of them is associated with the axis – thus leading to the 'spatial schemas' and 'sequence schema' for inequality and to the 'possessive schema' and 'representative schema' for equality.

General Grammar. The aim of linguistic typology is precisely to sketch out a 'general grammar' (or 'typological universal grammar'), at the intermediate level between the universal level of mental representations – conceptualization of the world - and that of particular languages, *i.e.* to discover the 'menu' of techniques and grammatical categories where individual languages make their choice (Seiler, 2000; Lazard, 2000).

While some typologists balk at the idea of semantic 'universals', they all agree to pinpoint 'invariants', *i.e.* general regularities which impose limits to the variations and govern the possible relationships between markers and underlying operations.

Besides, it should be noted that typological investigations have showed the role of areal forces that are largely responsible for the choice of SRs types made by languages (Heine, 1997; Haspelmath & Buchholz, 1998).

Looking for General Properties of Language

Within the field of cognitive science, linguistics – defined as the science of language apprehended through the diversity of languages (Culioli, 1995) – has a specific role to play, by proposing metalinguistic representations of SRs in various languages, and relating them both to CRs and to surface structures. From that cross-linguistic point of view, only local theories can be constructed for the time being. Several local theories, which are more or less disjoined, are being produced on limited domains of grammatical systems or lexical fields, but only 'local maps' (so to speak) are available, and we still lack a general overview model.

Linguistic Relativity

Limited though they may be, 'local maps' drawn by typologists cannot but contribute to throw light on representational activity at work in languages. In the first place, they give evidence against any assimilation of SRs to CRs., and tend to favor some kind of 'constructivist' approach to linguistic meaning: language activity means producing and recognizing significant configurations (or 'forms' in the abstract sense of the term).

This, in turn, could contribute to the question of 'linguistic relativity', which has been widely debated for years. The clear-cut opposition between so-called 'Whorfian' positions (*i.e.* variant SRs mapping into variant CRs) and 'anti-Whorfian' positions (*i.e.* SRs mapping directly into universal CRs) has probably been overestimated. As advocated by cognitive grammars, semantic variability manifested in SRs does not involve conceptual variability in non-linguistic thinking and does not go against psychic unity of mankind. But, as far as it concerns 'thinking for speaking' (Slobin, 1996), it illustrates the relative autonomy of languages towards mental representations.

Cognitive Impact of SRs. Languages enforce obligatory grammatical and semantic distinctions, which reorganize mental representations; cross-linguistic studies thus shed a light on the specific ways in which small sectors of cognition are being structured in order to be represented by languages: they allow us to discover, so to speak, the topography of those sectors. Regularities can be observed in the 'grammatical slicing' operated by languages (Lazard, 2004): there seems to be invariant notions around which the grammatical categories of individual languages tend to take form. Some regions of what may be called the 'conceptual space' are such that most languages construct grammatical tools there: for instance temporality, space, or possession. In turn, these grammatical tools serve to express more complex notions (such as comparison for instance).

Language activity also contributes to reorganizing mental representations by forcing a linearization of thought and the taking of perspective (Levinson, 1997). Furthermore, while some SRs, which are prototypical, clearly derive from cognitively based conceptualizations, other SRs turn out to be secondary complexes that seem to be made possible by language activity itself (for instance comparisons where the compared items are no longer objects or individuals, but abstract constructs – such as events, modal or propositional contents). In other words, language activity has probably been a facilitating factor for the development of reflexive thinking.

A unified Approach to Language? Neurobiology has recently made great progress in the exploration of the brain, offering very precise insights into normal and pathological functioning of language. So has experimental psychology concerning language acquisition and linguistic performances of humans. But mutual contributions with linguistics (which is concerned with the very nature of linguistic systems and their variability) remain somewhat limited. Since linguistic phenomena are far from being known as they should, the gap between neuropsychological research on language faculty and linguistic studies on invariants is not likely to be filled in the near future. For only when linguistic invariants have been discovered at large and confirmed by adequate investigations, time will come to look for their possible psychological and neurological roots.

References

Andersen, P. K. (1983). *Word order typology and comparative constructions.* Amsterdam/Philadelphia: John Benjamins.

Culioli, A. (1995). *Cognition and representation in linguistic theory*. Amsterdam/Philadelphia: John Benjamins.

Fuchs, C. (1999). Diversity in linguistic representations: a challenge for cognition. In C. Fuchs & S. Robert (Eds.), *Language diversity and cognitive representations*. Amsterdam/Philadelphia: John Benjamins.

Fuchs, C. (2007). Relations de synonymie entre polysèmes: le réseau 'comme-manière-façon'. *Le Français moderne, LXXV:1.*

Fuchs, C., & Le Goffic, P. (2005). La polysémie de 'comme'. In O. Soutet (Ed.), *La polysémie.* Paris: Presses de l'Université de Paris-Sorbonne.

Greenberg, J. (1966). *Language universals.* The Hague: Mouton.

Haspelmath, M., & O. Buchholz (1998). Equative and similative constructions in the languages of Europe. In J. van der Auwera (Ed.), *Adverbial constructions in the languages of Europe.* Berlin/New-York: Mouton-de Gruyter.

Heine, B. (1997). *Cognitive foundations of grammar.* Oxford/New-York: Oxford University Press.

Henkelmann, P. (2006). Constructions of equative comparison. *Sprachtypol. Univ. Forsch., 59:4,* 370-398.

Lazard, G. (2000). Two-level relationships between language typology and cognitive linguistics. *Proceedings of the international conference on Cognitive Typology (Antwerpen).*

Lazard, G. (2004). On the status of linguistics with particular regard to typology. *The Linguistic Review, 21,* 389-411.

Lehmann, W. (1972). Contemporary linguistics and IE studies. *Publications of the Modern Language Association of America, 87,* 976-993.

Lehmann, W. (1973). *Historical linguistics : an introduction.* 2nd ed. New-York: Holt, Rinehart & Winston.

Levinson, S. (1997). From outer to inner space: categories and non-linguistic thinking. In E. Pederson & J. Nuyts (Eds.). *Language and conceptualization* Cambridge: Cambridge University Press.

Rivara, R. (1975). How many comparatives are there?. *Linguistics, 163,* 35-51.

Rivara, R. (1995). Pourquoi il n'y a que deux relations de comparaison. *Faits de langues, 5,* 19-39.

Seiler, H-J. (2000). *Language universals research: a synthesis.* Tübingen: Gunter Narr Verlag.

Slobin, D. (1996). From 'thought and language' to 'thinking for speaking'. In J. Gumperz & S. Levinson (Eds.). *Rethinking language relativity.* Cambridge: Cambridge University Press.

Stassen, L. (1985). *Comparison and universal grammar.* Oxford: Basil Blackwell.

Against Darwinism

Jerry Fodor (fodor@ruccs.rutgers.edu)
Department of Philosophy – Rutgers University, Davison Hall, Douglass Campus
New Brunswick, NJ 08903, New Jersey, USA

This started out to be a paper about why I don't like EP (i.e. the evolutionary theory of prepositional attitudes, hence of intentional states). But then it occurred to me that what the paper was really about wasn't the tension between Darwinism and theories that are intentional (with a \`t'), but the tension between Darwinism and theories that are intensional (with an \`s\`) [1]. The latter is more worrying since Darwinism, or anyhow adaptationism, is itself committed to intensionally individuated processes like \`selection for.' So the claim turned out to be that there is something seriously wrong with adaptationism per se. Having arrived at that, I could have rewritten this as straight-forwardly a paper about adaptationism, but I decided not to do so. It seems to me of interest to chart a route from being suspicious of Evolutionary Psychology to having one's doubts about the whole adaptationist enterprise. Hence what follows.

The central claim of Evolutionary Psychology (EP) is that heritable properties of psychological phenotypes are typically adaptations; which is to say that they are typically explained by their histories of selection. In particular, this is claimed on behalf of heritable phenotypic properties that involve intentional states like believing, desiring, and acting (or being disposed to act) in one way or another. It is reasonable to hold that the evidence for this claim, so far at least, is underwhelming. Be that as it may; in the first part of this paper I want to argue for something much stronger: that the whole idea of an evolutionary psychology is very likely ill-conceived. Much of the main line of argument I'll pursue is already to be found in the philosophical literature, especially the literature on evolutionary semantics. So my strategy is to start by reminding you of some of the morals of that discussion and to contend that they apply quite generally to selectionist accounts of the cognitive psychological phenotype.

The Edifying Fable of the Frogs and the Flies

Frogs snap at flies; having caught one, they then ingest it. It is in the interest of frogs to do so since, all else equal, the fitness of a frog that eats flies (and hence the likelihood of its contributing to the local gene pool) exceeds the fitness of a frog that doesn't. It is likewise plausible that the frogs' penchant for catching flies is an adaptation; which is to say that it was established in their behavioral phenotype by a process of natural selection. If so, then perhaps it follows that the function of the behavior (and/or of the physiological mechanisms by which it the behavior is implemented), is precisely to mediate the catching of the flies by the frogs. Maybe, that's to say, some selectionist story about the phylogeny of fly-snapping can provide, at the same time, an account of the teleology of that response. I don't believe much of that, but never mind; let's assume for now that it's all true.

I suppose it is likewise plausible that frogs catch flies with the intention of doing so. (If you are unprepared to swallow the attribution of intentions to frogs, please feel free to proceed up the phylogenetic ladder until you find a kind of creature to which such attributions are, in your view, permissible.) Now, intentions-to-act have intentional objects, which may serve to distinguish among them. A frog's intention to catch a fly, for example, is an intention to catch a fly, and is ipso facto distinct from, say, the frog's intention to sun itself on the leaf of a lily. This consideration may encourage the following speculation: the fact about the teleology of the frog's fly catching mechanisms and the fact about the intentional object of its snaps both reduce to the fact that the frog's behavior is an evolutionary adaptation selected for catching and eating flies; which is, in turn, a fact that a selectionist account of the behavior's phylogeny may be supposed to entail. If that's right, then the transition from an adaptationist theory that explains the frog's behavior in terms of its effects on fitness, to a functional theory that explains the frog's behavior in terms of its teleology, to a psychological theory that explains the frog's behavior in terms of the content of its propositional attitudes amounts, in effect, to a reduction of intentionality to selection. This line of thought is not without its partisans, either in philosophy or in cognitive science at large.

But for every ointment there is a fly. The problem is that nothing about content or about teleology appears to follow directly from the assumption that fly-catching is an evolutionary adaptation in frogs. At a minimum, such inferences require the further, stronger, assumption that fly-catching behavior is an adaptation for catching flies; (i.e that catching flies is what the behavior was selected for). But \`adaptation for…', \`selection for…' and the like are themselves intensional contexts (just like \`belief that…' and \`intention to…'.). A mechanism that's selected for catching flies is not ipso facto a mechanism that's selected for catching ambient black nuisances; not even if, either in this part of the woods or in general, all and only the ambient black nuisances are flies. This logical quirk distinguishes \`selection for' from mere selection. If you select a mechanism that catches Xs, and if the Xs are Ys, then you thereby, select a mechanism that catches Ys. Selection is an extensional process, so it can't, as it were, \`see' the difference between intentional states that are extensionally equivalent. But the analogous point doesn't hold if the topic is \`selection for…' If you are selecting for Bs and Bs are Cs, it doesn't follow (and it needn't be true) that you are selecting for Cs. \`Select' doesn't distinguish among extensionally identical states, but `select for...' does.

So the situation is this: either natural selection is a species of `selection for...', and is thus itself a kind of intensional process; or natural selection is a species of selection tout court, and therefore cannot distinguish between coextensive mental states. In the former case it may, but in the latter case it doesn't, provide an explanation either of the teleology or of the intentional content of the frogs' snapping.

In the literature on philosophical semantics, the present point is often formulated as the `disjunction problem'. In the actual world, where ambient black dots are flies, it is in a frog's interest to snap at flies. But, in such a world, it is equally in the frog's interest to snap at flies-or-ambient-black-dots. Snap for snap, snaps at the one will net you exactly as many flies to eat as snaps at the other. Snaps of which the intentional objects are flies and snaps whose intentional objects are black dots both affect a frog's fitness in exactly the same way and to exactly the same extent. Hence the disjunction problem: what is a frog snapping at when it snaps at a fly?

Thus far: It's plausible that natural selection can account for (heritable) intentional properties of a creature's phenotype only if it can distinguish selection of creatures that have such properties from selection of creatures for having such properties. If that's right, we can take the line of thought a step further. It would seem that the relevant difference between mere selection and selection for has to do with the status of certain counterfactuals. For example, according to this suggestion, to claim that frogs were selected for snapping at flies is to say (first) that in this world, where the ambient black- dots-or-flies are generally flies, frogs that snap at them are selected; and (second) that such frogs would not be selected in (nearby) counterfactual worlds where the ambient flies-or-black dots generally aren't flies (perhaps they're bee bees) So, now: can natural selection settle the issue between these counterfactuals?

I can think of two ways in which it might be supposed to do so. Both crop up, more or less explicitly, in the adaptationist literature, but I'm going to argue that neither of them has a prayer of working. I haven't heard of other alternatives and I can't prove that there are none. But I do rather think that these two exhaust the options. I am even prepared to wager moderate sums that they do.

First Option: Mother Nature

There's a sort of analogy between what natural selection does when it culls a population and what breeders do when they select from a population those members that they encourage to reproduce. This analogy was noticed by Darwin himself, and it has been influential in the popularizing adaptationist literature ever since. Suppose Granny breeds zinnias, with the intention of selling them on Market Day. Then Granny is selecting zinnias for their value on the market, and not, say, for the elaboration of their root-systems. This is so even if, as a matter of fact, it's precisely zinnias with elaborate root-systems that sell at the best prices. Likewise, the fact about her intentional psychology that explains which zinnias Granny chooses when she sorts them is that she is interested in selling them, and not that she is interested in their having lots of roots. (Granny may not even know about the connection between market values and root systems. Or, if she knows, she may not care.) In short, since Granny is in it for the money and not for the roots, there is a matter of fact about what she selects for when she selects some of the zinnias and rejects the others. What Granny selects for is: whatever it is that she has in mind when she does her selecting.

So, then, perhaps we should take the analogy between natural selection and selective breeding at its face value. Perhaps we should say of natural selection just what we said of Granny: that what it selects for is whatever it has in mind in selecting? The counterfactuals fall out accordingly: If Granny is interested in high market value rather than big roots, that decides what she would do in a world where the salable zinnias are the ones with short roots, or no roots, or green roots with yellow polka dots, or whatever. Likewise, if natural selection has it in mind that there should be lots of frogs that eat flies, then, in the actual world, where the flies or bee bees are mostly flies, it favors both frogs that snap at flies and frogs that snap at bee bees. But in the counterfactual world where the flies-or-bee-bees are mostly bee bees, natural selection will favor only the frogs that snap at flies.

That, surely, is the thought that explains the prominence of anthropomorphized avatars of natural selection in the EP literature: Mother Nature, The Blind Watchmaker, The Selfish Gene or, for that matter, God. All of these are supposed to be (as one says); `intentional systems': they have intentions in light of which they act. So, if one construes natural selection on the model of selection by an intentional system, one thereby makes room for a distinction between selection that has it in mind to propagate frogs that snap at flies and selection that has it in mind to propagate frogs that snap at flies-or-bee-bees; which is, I'm supposing, precisely the sort of distinction that you need to make room for if you are going to make sense of selection for beliefs, desires, goals and the like.

When it's put that baldly, however, it's perfectly obvious what's wrong with this line of thought: natural selection doesn't have a mind; a fortiori, it has nothing in mind when it selects among frogs. Likewise, if genes were intentional systems, there would be an answer to, for example, the question whether natural selection favors creatures that really do care about the flourishing of their children or creatures that really care only for the propagation of their genotypes. All you have to do, if you want to know, is find out which phenotype their genes prefer.

So, if genes are themselves intentional systems, or if there is a Mother Nature who selects with ends in view, then which creatures are selected can after all determine which traits they are selected for. That's the good news. The bad news is that, unlike natural selection, Mother Nature is a fiction, and fictions can't select things, however hard they try. Nothing cramps one's causal powers like not existing. Likewise, mutatis mutandis, the genes that make you cause your children to flourish (if, indeed, there are such genes) couldn't care less about why you want your children to do so. They couldn't care less about that because they don't care at all about anything.

Only agents have minds, and only agents act out of their intentions, and natural selection isn't an agent. To the contrary, it's an important part of the advertising for adaptationism that its way of explaining why the selection of phenotypes generally tends towards increasing fitness doesn't require attributions of agency. Because that's so (and assuming that it's true), adaptationism can legitimately claim to advance the scientific program of naturalizing nature.

You may think the preceding speaks without charity; that I am, in fact, shooting in a barrel that contains no fish. Surely, you may say, nobody could really hold that genes are literally concerned to replicate themselves? Or that natural selection literally has goals in mind when it selects? Or that it's literally run by an intentional system? Maybe.[2] But, before you deny that anybody could claim any of that, please do have an unprejudiced read through the EP literature. Meanwhile, I propose to consider a different way of arguing that adaptationism, because it can support the counterfactuals that distinguish mere selection from selection for, can likewise distinguish fly-snapping frogs from fly-or-bee-bee snapping frogs; thereby providing a paradigm for selectionist accounts of the content (and the teleology) of intentional states.

Second Option: Laws of Selection

Laws can support counterfactuals. Arguably, that's what makes laws different from mere true empirical generalizations. So, then, suppose there is a law from which it follows that t1s are selected in competitions with t2s. It's truistic that, if there is such a law, then it holds in all nomologically possible situations; which is to say that it determines the outcome of any nomologically possible t1 v. t2 competition. That includes competitions that are merely counterfactual, so long as they are nomologically possible. So then, if there's a law that connects the property of being a t1 and the property of competing successfully with t2s, and if the distinction between selection of t1s and selection for being t1s turns on the corresponding counterfactuals, then laws of selection might vindicate the selection/selection for distinction.

Well, are there such laws? I think it's most unlikely.

It's a thing about laws that they aspire to generality. In the paradigm cases, a law about Fs is supposed to apply to instances of F per se. Conversely, to the extent that a generalization applies not to Fs per se, but only to Fs-in-such-and-such circumstances, it's correspondingly unlikely that the generalization is a law (or, if it is a law, it's correspondingly unlikely that it's a law about Fs per se.) I take that to be common ground; but if it's right, then quite likely there aren't laws of selection. That's because who wins a t1 v. t2 competition is massively context sensitive. (Equivalently, it's massively context sensitive whether a certain phenotypic trait is conducive to a creature's fitness.) There are a number of respects in which this is true, some obvious some less so.

For example, it's obvious that no trait could be adaptive across the board. Rather, the adaptivity of a trait depends on the ecology in which its bearer is embedded. In principle, if a trait is maladaptive in a certain context, you can fix that either by changing the trait or by changing the context. Is a creature's being green good for its fitness? That depends on whether the creature's background is green too. Is being the same color as its background good for a creature's fitness? That depends on whether camouflage that makes the creature hard for predators to find also makes it hard for the creature to find a mate. Is it good for a creature's fitness to be big? Well, being big can make it hard to flee from predators. Is it good for a creature to be small? Perhaps not if its predators are big. Is it good for a creature to be smart? Ask Hamlet (and bear in mind that, when it's all over and evolution has finished doing its thing, it's more than likely, that the cockroach will inherit the earth). Whether a trait militates for a creature's fitness is the same question as whether there's an `ecological niche' for creatures that have the trait to occupy; and that always depends on what else is going on in the neighborhood. Is it good to be a square peg? Not if the local holes are mostly round.

I want to emphasize that my point isn't just that, if there are laws about which traits win which competitions, they must be `ceteris paribus' laws. To the contrary, I take it to be true quite generally that special science laws hold only `all else equal'. If that's so, it's not a complaint against the putative laws of selection that they do too. I think, however, that the present considerations go much deeper.

To a first approximation, the claim that, ceteris paribus Fs cause Gs says something like: `given independently justified idealizations, Fs cause Gs.'[3] The intuition is that, underlying the observed variance, there is a bona fide, reliable, counterfactual-supporting connection between being F and causing Gs, the operation of which is obscured by the effects of unsystematic, interacting variables; the underlying generalization comes into view when the appropriate idealizations are enforced. By contrast (so I claim) there just aren't any reliable generalizations about which traits win competitions with which others. It simply isn't true, for example, that being big is in general better for fitness than being small except when there are effects of interacting variables; or that flying slow and high is in general better for fitness than flying fast and low except when there are effects of interacting variables; or that being monogamous is in general better for fitness than being polygamous except when there are effects of interacting variables etc. It's not that the underlying generalizations are there but imperceptible in the ambient noise. It's rather that there's just nothing to choose between (e.g.) the generalization that being small is better for fitness than being big and the generalization that being big is better for fitness than being small. Witness the fact that the world contains vastly many creatures of both kinds. I don't doubt that there are explanations of why competitions between creatures with different traits come out the way they do; but such explanations don't work by subsuming the facts they explain under general laws about the relative fitness of the traits. (I'll say something, pretty soon now, about how I think they actually do work.)

Nor is that by any means the whole story about the context dependence of being a trait that's selected for. In fact, strictly speaking, traits don't get selected at all; traits don't either win competitions or loose them. What wins or looses competitions are the creatures that have the traits. That's to say that what's selected is whole phenotypes; and, quite possibly, whether a certain trait is fitness-enhancing depends a lot on what phenotype it's embedded in. That too is practically a truism; but it's one that game-theoretic models of evolution (for example) have a bad habit of ignoring. `What would happen if a population of ts were to invade a population of not-ts?' That depends a lot on what other differences distinguish the ts from the not-ts. `Yes, but all else equal what will happen if a population of ts invades a population of not-ts?' Since `all else' never is equal, the question doesn't seriously arise. Unlike a scientist in a laboratory, natural selection can't control for confounding variables; it has no access to the method of differences.

As we're about to see, these sorts of considerations apply to adaptationist explanations across-the-board; but they apply in spades when what's at issue is selection for intentional states. That's because, unlike any others, intentional states invariably have unintended consequences, and natural selection can't see the difference between a consequence that is intended and a consequence that isn't. `Jack and Jill/ Went up the hill/ To fetch a pail of water/ Jack fell down/ And broke his crown/ And thus decreased his fitness.' We can see that what was detrimental to Jack's fitness was neither his intention to fetch water, nor his intention to climb a hill in order to do so. It was the falling down that was bad for him, and that wasn't part of the intention on which he acted. Since we can see all that, we're prepared to conclude that, although Jack's action brought him to grief, evolution shouldn't count its having done so as a reason for selecting against mental states whose intentional objects are climbings of hills or fetchings of water. Jack's climbing the hill eventuated in damage to his crown; but it wasn't, as one says, `intentional under that description.'

To suppose that the processes of evolution can see that the actual outcome of Jack's action was incidental to its intentional object is precisely to beg the questions that are now at issue. We can understand what went wrong with Jack because we have the concept of `the maxim of an act' (the description under which the act was intended), and it's clear to us that the maxim of Jack's act was something like `when thirsty, fetch water' and nothing at all like `when thirsty, fall down/ and break your crown'. But, recall that (putting aside the loose talk about what evolution can `see`) the adaptationist's aim was to explain how the fitness of an intentional state varies as a function of its content. So, if he's to avoid circularity, he can't take for granted either that intentional states with distinct effects on fitness are ipso facto distinct in content or that intentional states that are distinct in content are ipso facto distinct in their effects on fitness. Jack's crown got broken and Jill's didn't. It remains entirely possible that they both acted with the very same end in view.

I hope it's clear that I've thus far been running two kinds of arguments in tandem; two kinds of arguments that happen to converge in the case of issues about evolutionary psychology. The first concerns the goals of evolutionary psychology in particular; it's that data about effects on fitness can't, even in principle, distinguish the selection of any given intentional state from the selection of any other intentional state with the same actual outcomes. What's making the trouble here is the intensionality of the mental: Beliefs, desires and the like are individuated not by the consequences of having them but by their contents, and these two come apart whenever (or to the extent that) the actual effects of being in such a state are not the effects intended. But while the kind of worry we've just been discussing arises because of the intensionality of mental traits, there is another and independent kind of worry that derives from the intensionality of the notion of selecting for a trait, mental or otherwise. Once again, which trait a phenotype was actually selected for depends on which phenotype would have been selected-tout-court in appropriate counterfactual situations. And, once again, natural selection has no access to the counterfactuals; it can only 'see' the actual outcomes of phenotypic variations.

Because 'selection for' is intensional, so too are a galaxy of other adaptationist concepts that are defined in terms of it including, notably, that of a 'problem of adaptation' (aka an `ecological problem'); the very same configuration of the environment may present a problem of adaptation to one kind of creature but not to another even though the creatures live side by side. And, just as one would expect, the intensionality of ecological problems makes their individuation deeply obscure.

In familiar cases, solutions are defined by the problems that they solve. Thus the order of metaphysical dependence is that keys solve the problem of finding something to open locks, not that locks solve the problem of finding something for keys to open. In adaptationist theory, by contrast, there's a sort of topsy-turvy: Which problem a creature had depends on which of its traits were selected for solving it. But that there are spiders, who would have guessed that how to spin webs to catch flies is an ecological problem? Or that there are creatures whose fitness is a consequence of their having solved it? Because selection for is intentional, a range of questions to which a theory of adaptation ought to be responsive are, in fact, answered entirely post hoc.

The long and short is that the intensionality of the attitudes and the intensionality of selection for both raise problems of individuation, but the first kind of problem is much less of a worry than the second. A reasonable biologist might be willing to live without a selectionist evolutionary psychology so long as there's no implied threat to adaptationism per se. So, when the weather gets rough, there's an entirely understandable temptation to lighten the ship by throwing the psychologists overboard. But, in fact, to do so wouldn't help; the intentionality of selection for makes trouble for adaptationism as such, and it would continue to do so even if, in our panic, we were to adopt some sort of behaviorism or neurological reductionism, thereby making intentional psychology disappear.

Exasperation may now urge the following response. `Why shouldn't I think that that was all just epistemology pretending to be metaphysics? `What you've offered isn't grounds for claiming that there are no laws of selection. At most it's grounds for claiming that, if there are such laws, then, because of their context dependence, they must be very complicated; perhaps, even, they're not within our capacity to formulate. But nothing of metaphysical interest follows from that. In particular, nothing follows as to the status of counterfactuals about which phenotypes would, and which ones wouldn't, be selected in possible worlds other than our own. Laws that are too complex for us to formulate can support counterfactuals all the same.

"After all, do you really want to say that adaptationist explanations aren't ever any good; that selection histories never explain phenotypic traits, psychological or otherwise? Surely you're aware that the textbooks simply team with examples to the contrary. These textbook explanations purport to, and often clearly do, give reasons why phenotypes are the way they are; why there are lots of populations of T1s, but few or no populations of T2s. Well, there can't be such explanations unless there are laws about the relative fitness of various traits. Since you can't have the explanations without the laws, and since the illumination that the explanations often provide isn't subject to serious doubt, it would seem that if you don't like laws of selection, you will have to learn to lump them."

Thus the voice of exasperation, and I think there's a lot in what it says. Certainly I have no objection to the form of its argument: If there are few or no examples of laws of selection on offer, that could be either because there are few or no such laws; or because there are many such but we aren't smart enough to find them out. And I, for one, disapprove, vehemently, of arguments that purport to draw metaphysical conclusions from epistemological premises. Still more vehemently do I disapprove of ignoring what otherwise seems to be successful science on the grounds of merely philosophical scruples.

On the other hand, it's crucial in the present case not just that there are bona fide successful adaptationist explanations, but also that such explanations invoke laws of selection. If they don't, then the success of the explanations is not a reason to think that there are such laws. In fact, I'm inclined to think that the premises invoked in explaining phenotypes by reference to selection histories generally aren't nomological, and that they don't claim or even aspire to be. What they are is precisely what they look like on the face of them; they're historical explanations. Very roughly, historical explanations offer not covering laws but plausible narratives; narratives which (purport to) articulate the causal chain of events leading to the event that is to be explained. Whereas covering law explanations are about (necessary) relations among properties, historical narratives are about (causal) relations among events. That's why the former support counterfactuals, but the latter don't.

Historical explanations are as far as I know, often perfectly ok. Certainly they are sometimes thoroughly persuasive, so perhaps they are sometimes true. But, prima facie at least, historical explanations don't seek to subsume events under laws. `She fell because she slipped on a banana peel.' Very likely she did; but there's no law ---there's not even a statistical law--- that has `banana peel' in its antecedent and `slipped and fell' in its consequent. Likewise, Napoleon lost at Waterloo because it had been raining for days, and the ground was too muddy for cavalry to charge. So, anyhow, I'm told; and who am I to say otherwise? But it doesn't begin to follow that there are laws that connect the amount of mud on the ground with the outcomes of battles.

Metaphysical naturalists have to say, I suppose, that the effect of the mud on the outcome of the battle at Waterloo must have fallen under some covering laws or other. No doubt, for example, it instantiated laws of the mechanics of middle-sized objects. But it doesn't follow that there are laws about mud so described, or about battles so described, still less about causal connections between them so described; which is what would be required if `he lost because of the mud' is to be an instance of a covering-law explanation. It likewise doesn't follow, and it isn't remotely plausible, that whatever explains why Napoleon lost at Waterloo likewise explains why Nelson won at Trafalgar; i.e. that there are laws about battles as such, of which Nelson's victory and Wellington's are both instances. `Is a battle' doesn't pick out a natural kind; it's not (in Nelson Goodman's illuminating term) `projectible`.

It's of a piece with their not appealing to covering laws that historical-narrative explanations so often seem to be post hoc. The reason they so often seem to be is that they usually are. Given that we already know who won, we can tell a pretty plausible story (of the too-much-mud-on-the-ground variety) about why it wasn't Napoleon. But, what with their being no covering law to cite, I doubt that Napoleon or Wellington or anybody else could have predicted the outcome prior to the event. The trouble is that there would have been a plausible story to explain the outcome whoever had won; prediction and retrodiction are famous for exhibiting this asymmetry. That being so, there are generally lots of reasonable historical accounts of the same event, and there need be nothing much to choose between them. Did Wellington really win because of the mud? Or was it because the Prussian mercenaries turned up just in the nick of time? Or was it simply that Napoleon had lost his touch? (And while you're at it, what, exactly, caused the Reformation?)

It's not in dispute that competitions between creatures with different phenotypes often differ in their outcomes; and, of course, there must be, in each case, some explanation or other of why the winner won and the looser didn't. But there's no reason at all to suppose that such explanations typically invoke laws that apply to the creatures in virtue of their phenotypic traits. That being so, there need be nothing to choose between claims about the corresponding counterfactuals. Small mammals won their competition with large dinosaurs. But did they do so because of their smallness? That depends (inter alia) on whether they would have won even if there hadn't been a meteor. I can tell you a plausible story about why they might have won: Small animals are able to snitch dinosaur eggs to eat when the dinosaurs aren't looking, (which is bad for the dinosaurs' fitness.) On the other hand, I can tell you a plausible story about why they might not have won: lacking the meteor, there wouldn't have been selection for tolerance to climate change, which mammals had but dinosaurs didn't. (Notice that, according to the latter story, it wasn't the smallness or quickness of the mammals that was selected for, but the range of temperatures they were able to tolerate.) So which of the counterfactuals do our evolutionary narratives about the extinction of dinosaurs support? Neither? Both? And, likewise, what intentional content did evolution select for when it selected creatures that protect their young? Was it an altruistic interest in their offspring or a selfish interest in their genes?

The moral, so far, is that a phenotype's having been selected doesn't determine which (if any) of its traits a creature was selected for. Quite generally, if you want to infer from the one to the other, you have two choices (and, as far as I can see, only two.) You can try attributing intentions to the agent of selection (hence Mother Nature); or you can try to find a covering law that connects its having some phenotypic trait with a creature's having been selected. The former tactic is hopeless; there simply isn't any Mother Nature, and natural selection has nothing in mind when it prefers some creatures to others; natural selection has nothing in mind at all. But the second tactic seems hopeless too, given the extreme context sensitivity of selection processes. Whether a trait is conducive to fitness appears to be just about arbitrarily dependent on which sort of creature it's a trait of and what sort of ecology the creature inhabits. If that's so, then there can't be laws of selection, and `selected for' can't be a projectible predicate.

There is, however, a model of adaptationist explanation that seems to fit the facts pretty well. If it's otherwise viable, it suggests that such explanations, though they aren't nomic, have perfectly respectable precedents. Adaptationist explanations are species of historical narratives. If so, then everything can be saved from the wreckage except the notion of selection for: Since historical narratives don't support counterfactuals, it seems selection for very likely can't be salvaged. That's all right; much spilled ink to the contrary notwithstanding, there very likely isn't any such thing.

I want to close by suggesting an analogy; it is, I think, a very close analogy. For each rich person, there must be something or other that explains his being so: heredity, inheritance, cupidity, acuity, mendacity, grinding the faces of the poor, being in the right place at the right time, having friends in high places, sheer brute luck, highway robbery or whatever. Which things conduce to getting rich is, of course, highly context dependent: It's because of differences in context that none of us now has a chance of getting rich in (for example) the way that Genghis Khan did; or in the (not dissimilar) way that Andrew Carnegie did; or in the (quite different) way that Andrew Carnegie's heirs did; or in the (again quite different) way that Liberacie did; and so forth. Likewise (and not withstanding all those how-to-get-rich books) the extreme context sensitivity of the traits that eventuate in getting rich make it most unlikely that there could be a theory of getting rich per se. In particular, it's most unlikely that there are generalizations that are lawful (hence counterfactual supporting, not ad hoc, and not vacuous) that specify the various situations in which it is possible to get rich and the properties in virtue of which, if one had them, one would get rich in those situations. This is, please notice, fully compatible with there being convincing stories that explain, case by case, what it was about a guy in virtue of which he got as rich as he did in the circumstances that prevailed when and where he did so.

Well, I think adaptationist explanations of the evolution of heritable traits are sort of like that. When they work it's because they provide plausible historical narratives, not because they cite covering laws. In particular, pace Darwinists, adaptionism doesn't articulate the mechanisms of the selection of heritable phenotypic traits; that's because there aren't any mechanisms of the selection of heritable phenotypic traits as such. All there are is the many, many different kinds of ways in which various kinds of creatures manage to flourish in the many, many environmental situations in which the do so.

None of this should lighten the heart of anybody in Kansas; not even a little. In particular, I've provided not the slightest reason to doubt the central Darwinist thesis of the mutability of species. Nor have I offered the slightest reason to doubt that we and chimpanzees had (relatively) recent common ancestors. Nor do I suppose that the intentions of a designer, intelligent or otherwise, are among the causally sufficient conditions that good historical narratives would appeal to in order to explain why a kind of creature has the phenotypic traits it does (saving, of course, cases like Granny and her zinnias.) It is, in short, one thing to wonder whether evolution happens; it's quite another thing to wonder whether adaptation is the mechanism by which evolution happens. Well, evolution happens; the evidence that it does is overwhelming. I blush to have to say that so late in the day; but these are bitter times.

Footnotes

[1] It's hard to imagine a less fortunate terminology than the philosopher's `intention/intension' distinction. But I suppose there's nothing can be done at this late date. In what follows, an intensional context is one in which the substitution of coextensive expressions isn't valid. Intentional states are just the familiar beliefs, desires, intentions and so forth that populate theories of cognition and the integration of behavior. I assume, following the tradition, that expressions that refer to propositional attitudes typically establish intensional contexts, so that one can believe that Venus is the Morning Star and yet not believe that Venus is the Evening star, despite the fact that... etc.

[2] Admittedly, the tactic of resorting to scare quotes when push comes to shove (as in `what natural selection `prefers,'' `what Mother Nature `designs,'' `what the selfish genes `want'' and so forth) can make it hard to tell just what is being claimed in some of the canonical texts. Still, there are plenty of apparently unequivocal passages. Thus Pinker (1997, p.93): "Was the human mind ultimately designed to create beauty? To discover truth? To love and to work? To harmonize with other human beings and with nature? The logic of natural selection gives the answer. The ultimate goal that the mind was designed to attain is maximizing the number of copies of the genes that created it. Natural selection cares only about the long-term fate of entities that replicate..." Fiddlesticks. The human mind wasn't created, and it wasn't designed and there is nothing that natural selection cares about; it just happens. This isn't Kansas, Toto.

[3] It's crucial that the idealizations are independently justified; otherwise `ceteris paribus Fs cause Gs' collapses into `Fs cause Gs except when they don't.'

Invited Symposia

Analogy-Making: Resolved and Unresolved Mysteries

Boicho Kokinov (bkokinov@nbu.bg)
Central and East European Center for Cognitive Science, Department of Cognitive Science and Psychology,
New Bulgarian University, 21, Montevideo Str., Sofia 1635, Bulgaria

Dedre Gentner (gentner@northwestern.edu)
Department of Psychology, School of Education and Social Policy Northwestern University
2029 Sheridan Road; Evanston, IL 60208, USA

Researchers agree that analogy-making is about relational processing, and that the main components of analogy-making are representation building, memory retrieval, mapping, transfer, evaluation, and learning. There is also agreement in that analogy is a major contributor to human learning and reasoning. In this symposium, we would like to explore the extent of this agreement and understand where disagreements start. The contributors are researchers from various theoretical traditions, all engaged in original studies in various directions exploring new and still unresolved issues. They will explore the degree of convergence (or divergence) as to which territory to explore and how to best approach the issues.

Learning New Relations

Dedre Gentner & Stella Christie, Northwestern University

A hallmark of human cognition is the use of relational concepts—concepts like *monotonicity* and *reciprocity*. How do such concepts get formed? This is a challenging question, especially given that young children tend to focus on objects rather than relations. Our studies show that aligning two examples during learning dramatically increase children's relational insight, as compared to seeing the examples sequentially. A model of this phenomenon using SME (the structure-mapping Engine) suggests that comparison promotes relational insight: (1) by promoting a focus on connected relational structure; and (2) by inviting relational re-representation. Implications for purely Bayesian accounts of learning are discussed.

A Theory of Relation Discovery and Predication

Leonidas A. A. Doumas, Indiana University

The ability to think and reason about relations is a central component of human cognition. While we understand much about the mechanisms of relational thinking and analogy, little is known about how children and adults acquire relational concepts and represent them in a form that is useful for the purposes of relational thought (i.e., as structures that can be dynamically bound to arguments). We present DORA, a computational theory of relation discovery. DORA is a neurally-plausible cognitive architecture that learns relational concepts from examples and represents these concepts as explicit structures (predicates) that can take arguments.

Categories Based on Analogies

Kenneth Kurtz, State University of New York at Binghamton

The discovery of commonalities among the representational elements used to encode examples is critical to schema abstraction and category formation. In the categorization literature, similarity-based abstraction is traditionally understood in terms of overlapping or intercorrelated sets of features. In the comparison literature, similarity-based abstraction is based on alignable sets of explicitly-coded relations between objects/features. As part of an effort to bring these two views of learning and abstraction into closer contact, my talk will focus on behavioral data and theoretical claims addressing the acquisition, structure, and function of categories that cohere (like analogies) around relational content.

Analogy Programs that Learn

Robert French, CNRS and University of Burgundy

In this presentation we will concentrate on what we believe to be the challenges facing the new generation of analogy-making programs. These will focus on developing analogy-programs that learn and some of the challenges in developing these programs. In particular, we will discuss what will be required to break free from tradition of doing analogy-making with hand-coded problems. We will present a list of criteria for determining the success (or failure) of these new programs.

Modular or Interactive, Sequential or Parallel Processing in Analogy-Making

Boicho Kokinov, New Bulgarian University

The interactions between perception (representation building), memory retrieval (memory construction), mapping, transfer, and re-representation will be explored. Various theoretical possibilities, such as encapsulated modules vs. fully interactionist systems, sequential vs. parallel processing views, will be discussed together with their ability to generate testable predictions. Experimental designs to test some of these predictions will be suggested and the results of some initial experiments will be presented.

Language and Brain

Tatiana Chernigovskaya (Tatiana@TC3839.spb.edu)
Department of General Linguistics, Philological Faculty, St. Petersburg State University,
University Åmb.11, 199034 St. Petersburg, Russia

The cerebral basis for language is a central problem within cognitive science and can be studied only in a multidisciplinary anthropological perspective. Among other questions this symposium considers the importance of cross-language results, the diversity of behavioral, neuro-imaging and clinical data, and their interpretation. It also discusses the role of attention and error detection, the importance of species specific factors and input characteristics for language acquisition and processing. The results are observed in the framework of neurobiological localisation vs. network models and in connection with the debate on single vs. dual mechanism of mental lexicon organization.

On the Relationship between Functional and Structural Differences in the Brain

Kenneth Hugdahl University of Bergen

Brain asymmetry relates to both functional and structural differences between the two cerebral hemispheres. Despite all research devoted to functional asymmetry, very little is known of corresponding structural, or anatomical, differences. One notable exception is the larger left planum temporale area in the upper posterior part of the temporal lobe, and the relation of this to functional differences for speech perception. In my talk I will review recent data on asymmetry of speech sound perception, from both a basic and clinical perspective. In particular I will make an argument that auditory hallucinations in schizophrenia may be instances of speech sound mis-perceptions, caused by pathology at the neuronal level in the left planum temporale area. I will review data using both behavioural and psychophysiological measures, focusing on fMRI studies of neuronal activation during dichotic listening to simple speech sound syllables.

Speech Perception and Linguistic Experience

Inger Moen University of Oslo

Speech perception involves the identification of different types of structural information such as the identification of particular phonemes, lexical items, suprasegmentals, or the various phonetic properties associated with syntactic boundaries or discourse units. This is an ability which develops on the basis of linguistic experience. The details of this process are not fully understood. What is clear is that for continuous speech, perception does not depend solely on cues present in the acoustic waveform. The paper will present data from two Norwegian investigations which indicate that frequency of use is an important feature in speech perception.

Brain Mechanisms of Error Detection

Svyatoslav Medvedev Institute of the Human Brain, Russian Academy of Sciences

The brain mechanisms of error detection (ED), was firstly described by Bechtereva and Gretchin in 1968, as a physiological reaction to erroneous performance. The error detection system is a stabilizing mechanism of brain functioning. The basic functional principle of ED is the dissociation between the reality and the model in relevant memory matrix. Malfunctioning of ED is a physiological basis of some mental and language disorders. The data to be presented show that ED is functioning even when error (deception) is profitable for the performed task. There is a special mechanism that activates even in an intention to make the wrong action. ED forms a brain basis of conscience and language acquisition.

Normal and Deviant Processing of Inflected Nouns

Jussi Niemi University of Joensuu

Psycholinguistic studies of Finnish morphology show that in this language – excluding the very high frequent end of the continuum – inflected nouns like sauna-ssa 'sauna'-inessive, i.e. 'in a/the sauna' are morphologically parsed into their components in perception. In our presentation we will discuss the normal ontogenetic path of processing of these types of lexical items and especially deviant processing by speakers exhibiting Specific/Familial Language Impairment (SLI/FLI). Our extensive analysis of surface and lemma frequencies of Finnish noun paradigms and the expected use of frequency by normals and the unexpected insignificance of frequency in SLI/FLI are – we claim – in line with our previous findings regarding the present SLI/FLI individuals in fMRI and in dichotic listening, in which it was shown that these speakers have abnormal attentional patterns

Computational Psycholinguistics beyond Words

Hedderik van Rijn (D.H.van.Rijn@rug.nl)
Artificial Intelligence, University of Groningen, The Netherlands

Stefan L. Frank (S.Frank@nici.ru.nl)
NICI, Radboud University Nijmegen, The Netherlands

Research on language production and perception aims for the precise description and quantification of the cognitive processing steps involved in language use. While numerous detailed computational models have been presented that involve single word production and perception, only recently has research become more prominent that focusses specifically on modeling language above the level of single words. This symposium highlights this development by presenting four lines of research from different fields all involved in "Computational Psycholinguistics beyond Words".

A Probabilistic, Corpus-based Model of Syntactic Parallelism

Amit Dubey, University of Edinburgh
Although most research on sentence processing involves modeling ambiguity resolution, certain processes actually work to remove ambiguity from language. Two such phenomena are syntactic priming (Bock, 1986), a general process, and the parallelism effect, which Frazier et al. (2000) claim is due to a mechanism distinct from priming. After describing these two effects, we then introduce a sentence processing model which helps answer a basic theoretical question: is parallelism really distinct, or an instance of the more general priming effect? The model accounts for syntactic priming by making the novel assumption that a probabilistic sentence processor (Jurafsky, 1996) operates within the ACT-R cognitive architecture. Overall, we find that evidence against a general mechanism is not as strong as previously thought.

Modelling Sentence Comprehension as Situation Construction

Stefan Frank, Radboud University Nijmegen
There is increasing consensus among discourse-comprehension researchers that understanding a text involves the construction of a mental representation of the described situation, which depends more on the reader's world knowledge and experience than on the text's linguistic and propositional form (Zwaan, 1999). Yet, nearly all computational models of sentence comprehension represent meaning propositionally, that is, as a combination of arbitrary symbols forming predicate and arguments (e.g., Mayberry et al., 2006). In contrast, I present a connectionist model that constructs non-propositional and non-symbolic 'situational representations' (Frank et al., 2003) given sentences describing situations in a simple microworld. Also, its ability to account for experimental data is discussed.

Connectionist Models of Sentence Comprehension in Context

Marshall R. Mayberry, III, Saarland University
According to the coordinated interplay account (CIA), there is a dynamic interaction between language and visual context, in which comprehension of a situated utterance rapidly guides attention to objects and events in a scene and, in turn, the attended region of the scene tightly constrains and influences comprehension (Knoeferle & Crocker, 2006). We present an architecture, CIANet, that directly models the CIA by means of an explicit attentional mechanism to select the event in the scene most relevant to the utterance (Mayberry et al., 2006). We show how this mechanism enables predictions of how people resolve conflicting information, as well as how argument information from the scene can be used in the face of initial structural ambiguity.

Producing Time

Hedderik van Rijn, University of Groningen
Simone Sprenger, Radboud University Nijmegen
Performance in relative clock time naming (e.g., pronouncing 3:50 as "ten to four") has been described as depending on three factors: reference hour determination, minute transformation, and an additional distance component (Meeuwissen et al., 2003). However, this account does not specify the cognitive operations that are responsible for the distance effect. We present a computational model that explicates these cognitive operations, and provide support for this model by sets of regression models of speech onset latencies.

References

Bock, J.K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, *18*, 355–387.
Frank, S.L., Koppen, M., Noordman, L.G.M., & Vonk, W. (2003). Modeling knowledge-based inference in story comprehension. *Cognitive Science*, *27*, 875–910.
Frazier, L., Munn, A., & Clifton, C. (2000). Processing coordinate structures. *J. of Psycholinguistic Res.*, *29*, 343–370.
Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, *20*, 137–194.
Knoeferle, P., & Crocker, M.W. (2006). The coordinated interplay of scene, utterance, and world knowledge: evidence from eye-tracking. *Cognitive Science*, *30*, 481–529.
Mayberry, M.R., Crocker, M.W, & Knoeferle, P. (2006). A connectionist model of the coordinated interplay of scene, utterance, and world knowledge. *Proc. of CogSci Conf.*
Meeuwissen, M., Roelofs, A., Levelt, W.J.M. (2003). Planning levels in naming and reading complex numerals. *Memory & Cognition*, *31*, 1238–1248.
Zwaan, R.A. (1999). Embodied cognition, perceptual symbols, and situation models. *Discourse Processes*, *28*, 81–88.

Cooperation and the Evolution of Cognition

Peter Gardenfors (Peter.Gardenfors@lucs.lu.se)
Cognitive Science, Lund University, Kunghuset, Lundagård, S-222 22
Lund Sweden

In the debate concerning the evolution of cognition, much focus has been put on deception, in term of so called Machiavellian intelligence. The symposium will instead focus on the role of cooperation as a selective force for primate and hominid cognition. Different forms of cooperation will be related to the role of a theory of mind (intersubjectivity) and different communication systems.

On why Humans are the Only Animals who Have Developed A Symbolic Communication System

Peter Gardenfors, Cognitive Science, Lund University

This talk proposes an ecologically based answer to why humans are the only animals who have developed a symbolic communication system. The overall thesis is that there has been a co-evolution of anticipation, co-operation and communication. The first part of the argument claims that the Oldowan culture generated selective forces that lead to the evolution of anticipatory cognition, that is, the ability to mentally represent future needs and events. It is argued that anticipatory planning opened up for new forms of cooperation about future goals that were beneficial for hominid societies. Symbolic communication then emerged as the most efficient way of solving problems concerning cooperation about future non-existent goals. For example, the evolution of indirect reciprocity, which seems to be a uniquely human form of co-operation, depends on a symbolic communicative system and shared beliefs concerning the ”reputation” of individuals.

An Embodied and Distributed Approach to Primate Cognitive Evolution.

Louise Barrett, University of Lethbridge, Alberta, Canada

Living in a group is a cooperative act, which requires a delicate balance between individual and group level costs and benefits. Among primates, the social intelligence hypothesis has tended to focus attention on the means by which (Machiavellian) individuals cope with, and overcome, the costs that group-living imposes so that they may reap the associated benefits. These have been argued to involve highly cognitive strategies designed to track, monitor, cooperate with, and potentially outwit, other individuals. This, in turn, stems from a Cartesian view of the mind and cognition, and also from the kinds of evolutionary models used to predict and explain cooperative behaviour. This is problematic, however, as it has created a view of primate social complexity that is congenial to our view of ourselves, rather than one that is representative of primate social worlds. Here, drawing on work in cognitive science, including robotics, as well as neurobiology, I argue for a more embodied and distributed approach to primate cognitive evolution. Such an approach, when incorporated into evolutionary theories of multi-level selection and niche construction, presents us with the opportunity to explore primate cooperation and social complexity in ways that allow the animals to speak with their own voice, and not merely echo our own anthropocentric concerns.

The Cooperative Problem-Solving Abilities of Chimpanzees and Bonobos.

Brian Hare, Max Planck Institute for Evolutionary Anthropology

We compared the cooperative problem-solving abilities of chimpanzees and bonobos. When two subjects were confronted with a tray of out-of-reach, sharable food, both species were skillful at spontaneously pulling a rope simultaneously to obtain the food. When two subjects were again placed in the same situation, except the food was no longer sharable, bonobos showed more skill at solving the task. These results support the hypothesis that flexibility in cooperative problem solving is relative to different levels of emotional reactivity and suggests that the flexibility seen in our own species cooperative skills may have only evolved after the evolution of our unique human temperament.

Discussant

Cristiano Castelfranchi, CNR Rome, Italy

Consciousness and the Brain

Konstantinos Moutoussis (k.moutoussis@ucl.ac.uk)
Department of Philosophy and History of Science, University of Athens
Panepistimioupolis, GR. 157 71, Athens, Greece

One of the most exciting questions in Philosophy, Psychology & Neuroscience is that of consciousness and its relation to matter. The symposium will focus on a very basic form of consciousness, that of sensory awareness, present in both humans and animals. How can brain activation result in a subjective experience of the world, through the senses? Three invited speakers will give their own account with respect to visual perception.

The Many Consciousnesses of the Brain

Semir Zeki, Wellcome Laboratory of Neurobiology, University College London

The visual brain is now known to consist of many distinct visual areas, specialized to process and perceive different attributes of the visual scene. Contrary to common assumption, we do not perceive different attributes simultaneously. Instead we perceive some attributes before others; for example we perceive colour by about 100ms before we perceive motion, and we perceive simple forms after we perceive colour. Since perceiving an attribute is being conscious of it, and the perception of these different attributes is the result of activity in geographically distinct visual areas, it follows that visual consciousness is distributed in space. Equally, since we become conscious of different attributes at different times, visual consciousness is also distributed in time. It follows therefore that there is not a single unified visual consciousness, but there are instead many visual micro-consciousnesses. Moreover, binding of activity in these different areas occurs post-consciously.

Investigation of Subjective Motion Perception

Lars Muckli, Max-Planck Institute for Brain Research, & Department of Psychology, University of Glasgow

Bistable perception is ideally suited for the investigation of the cortical correlates of conscious perception. We used various apparent motion illusions (apparent motion breakdown, motion quartet) to induce bistable perception, and we investigated the cortical activation as a function of these subjectively perceived switches in perception. Cortical activation was measured by using functional magnetic resonance imaging (fMRI), and electroencephalography (EEG) and revealed how a cortical network comprising of motion specialized regions (including V5) interacts with retinotopic visual areas (including V1). We found fMRI activation and perceptual interference along the path of the apparent motion illusion. I like to speculate, that the perceived gist of a visual scene is the most important content of visual consciousness and, moreover, that the visual system uses the perceived gist to predict the near future and to provide effective filters for expected information – affecting also primary sensory areas.

How are Conscious and Unconscious Mental States Encoded in the Human Brain?

John-Dylan Haynes, Bernstein Center for Computational Neuroscience Berlin and Charité - Universitätsmedizin, Berlin & Max-Planck Institute for Human Cognitive and Brain Sciences

Accurate prediction of the conscious experience of an individual based only on measurements of their brain activity would provide strong evidence for a close link between brain and mind. Recent empirical and methodological advances in such 'brain reading' have yielded promising findings, particularly in the domain of visual perception. Here we show that conscious awareness of both simple features and complex object categories can be predicted from characteristic, distributed activity patterns in human visual cortex. Furthermore, feature-selective processing can also be demonstrated for stimuli of which the subject is completely unaware. Such an approach can also extend to decoding of other types of mental states, such as a subject's current focus of attention, their current intentions and even unconscious determinants of their behaviour. Taken together, this novel line of research helps reveal the way in which individual experiences are encoded in the human brain, and how they may be practically decoded.

Discussant

Fotini Stylianopoulou, University of Athens, Greece

Can Design Research Contribute to Bridge the Gap between Theory and Educational Practice?

Erik De Corte (erik.decorte@ped.kuleuven.be)
Centr. Instructiepsychol.&technologie, University of Leuven,
Andreas Vesaliusstraat 2, bus 3770, 3000 Leuven, Belgium

Stella Vosniadou (svosniad@phs.uoa.gr)
Department of Philosophy and History of Science, University of Athens
University Town, 157 71, Athens, Greece

Applied cognitive science research in education has a dual goal: (1) to develop theories and the impact of instructional interventions on learning and development, and (2) to contribute to the improvement of educational/classroom practices. The history of the field shows that there has always been a tension between these two goals. The aim of this symposium is to unravel and discuss from different perspectives the promises, the potential, but also the possible pitfalls of design-based research in the light of the dual goal mentioned above.

Design Experiments for Improving Thinking Skills through the Content of Learning

Benő Csapó, University of Szeged, Hungary

The most promising methods for developing general thinking skills are those that use the content of teaching itself for composing specific training exercises (often called enrichment, infusion or embedding). Experimenting with these methods is especially difficult. Furthermore, the limitations and the specific settings that are determined by the nature of the teaching materials challenge the generalizability of the results. This paper presents examples from the teaching of a variety of reasoning skills (e.g. operational, analogical and inductive reasoning) and school subjects in order to demonstrate how such difficulties may be overcome by further increasing the complexity of the design.

The Potential of Design-Based Research for Bridging the Theory – Practice Gap Relating to Education

Erik De Corte, University of Leuven, Belgium

Design-based research aims at the simultaneous pursuit of the advancement of our understanding of the processes of learning and instruction, on the one hand, and at the improvement of classroom practices, on the other hand. However, recently some authors (e.g., Phillips & Dolle, 2006) have disputed the potential of design experiments to achieve both goals simultaneously. Moreover, design experiments are criticized from a methodological perspective for lack of control and randomization resulting in confounding of variables. It will be argued, and illustrated, in this presentation that under certain conditions design-based research can accommodate both of these objections.

Bridging the Gap between Basic and Applied Research by an Integrative Research Approach (IRA)

Heinz Mandl & Robin Stark, University of Munich & Saarland University, Germany

The main goal of IRA is to bridge the methodological gap between basic and applied research on learning environments. IRA was developed and realized in the context of a special priority program of Deutsche Forschungsgemeinschaft (DFG) to investigate and foster processes of learning and teaching in the context of initial training in business administration. IRA combines laboratory with field methodology in a special way. Theory-based lab experiments are conducted with use questions in mind: Results of the experiments should be applied later on in practice. Field studies on the other side are conducted as “controlled” as possible. All characteristics of IRA will be exemplified by our research projects in the domain of economics and medicine.

The Problem of Knowledge in the Design of Learning Environments

Stella Vosniadou, University of Athens, Greece

The *cognitive apprenticeship* metaphor has played an important role in guiding the engineering approach to the design of learning environments. I will argue that the cognitive apprenticeship metaphor has provided important guidelines for teaching cognitive skills but that it has downgraded the problem of knowledge. Learning and the development of expertise in curricular domains require the construction of significant domain knowledge. Furthermore, the acquisition of *learning how to learn* skills cannot be divorced from substantial knowledge building.

Discussant

Naomi Miyake, Chukyo University, Japan

Symposia

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

NEUROCOGNITION OF HUMAN SPATIAL MEMORY: NEW APPROACHES, TESTS AND TECHNOLOGIES

Panagiota Panagiotaki (panagiota.panagiotaki@college-de-france.fr)

LPPA, Collège de France, Paris, France

Laboratory of Physiology of Perception and Action, CNRS- Collège de France, 11, place Marcelin Berthelot, 75005, Paris, France

Introduction

Spatial memory is a necessary cognitive property for human life. Previous studies mainly focused on static aspects of spatial encoding, such as object-location binding, or Euclidean positional memory. Nevertheless, the study of spatial tasks like navigation and reorientation into the 3-dimensional space demands to take under serious consideration their dynamic properties.

The purpose of this symposium is to present and discuss novel methodological approaches that contribute to the exploration of the cognitive mechanisms involved in the dynamic aspects of spatial encoding and of the neural networks underlying this processing. New paradigms focusing on this problematic are described: new neuropsychological tests administered in healthy controls and different categories of neurological patients, assessing different components of spatial memory; novel experiments of cognitive psychology investigating distinct cognitive strategies of encoding and storing spatial information; and finally studies on immersive virtual reality that test the function of cognitive strategies in humans that are present in lower levels of the phylogenetic continuum.

Cecilia Guariglia (cecilia.guariglia@uniroma1.it)

Università di Roma "La Sapienza"-IRCCS Fondazione Santa Lucia, Rome, Italy

Visuo-spatial exploration and memory for object arrays in the peripersonal space and for environmental information are often been considered equivalent in the neuropsychology. Nevertheless, there are some evidences that inability in coding and memorizing visuo-spatial information in the peripersonal space not ever corresponds to an inability in coding and memorizing information for environmental navigational purposes. A set of data showing that the system coding environmental information for building mental maps of the environment to be stored in long term memory could be double dissociated from the process involved in coding visuo-spatial information in the peripersonal space. In each experiment two tests were used: Corsi test and an experimental version of Corsi test on large scale. Two neuropsychological studies on neurological impaired subjects are also presented. The results of both studies underline the presence of dissociations in the two kinds of tests strongly supporting the hypothesis of separate memory systems for different types of visuo-spatial information.

Jan Wiener (jan.wiener@college-de-france.fr)

MPI for Biological Cybernetics, Tübingen, Germany & LPPA, Collège de France, Paris, France

A series of three experiments is presented that studied human path planning performance as well as the cognitive strategies and processes involved. 25 places were arranged on a regular grid in a large room. Each place was marked by a unique colored symbol. Subjects were repeatedly asked to solve traveling salesman problems (TSP), i.e. to find the shortest closed loop connecting a given start place with a number of target places. To specifically test for the relevance of spatial working memory (SWM) and spatial long-term memory (LTM) for path planning, the number of target places (ranging from 4 to 9 targets) as well as the mode of presenting targets was varied. Path planning performance systematically decreased with increasing TSP size. Furthermore, performance between the three experiments differed systematically. The results suggest the usage of different path planning strategies according to the specific memory demands.

Panagiota Panagiotaki

We present a study, where we explore the accuracy of human visual path integration (vPI) in complex and long paths (significantly longer than the simple triangle tasks) to the encoding and recalling of navigated paths. During this experiment, participants navigated semi-actively in an immersive virtual environment, performing an outbound searching trajectory in order to reach a specific goal. Once their goal reached, participants had to return immediately to their base (inbound journey), either based on visual cues or by performing vPI. Our results indicate that human vPI appears very deteriorated and it cannot be a reliable strategy for homing after complex and long trajectories.

Discussant: Alain Berthoz (alain.berthoz@college-de-france.fr)

Laboratory of Physiology of Perception and Action, CNRS- Collège de France,11, place Marcelin Berthelot, 75005, Paris, France

Causal Reasoning in Practice – How Causal Assumptions affect Explanations, Inferences and Decision Making

York Hagmayer (yhagmay@gwdg.de)
Department of Psychology, University of Göttingen,
Gosslerstr. 14
37073 Göttingen, Germany

Denis Hilton (hilton@univ-tlse2.fr)
Laboratoire Dynamiques Sociocognitives et Vie Politique,
Universite de Toulouse II - le Mirail,
Allees Antonio Machado 5
31058 Toulouse Cedex 9, France

Motivation

Causal learning and reasoning has been investigated intensively in the last decade, crucial insights were made and new theories were developed to explain both the acquisition and the use of causal knowledge. However, this research mostly focused on causal reasoning itself. Only recently researchers on causality became interested in how causal reasoning affects other areas of thought. This symposium aims to show that causal learning and reasoning is not distinct from other areas but permeates reasoning, judgment and decision making throughout. Examples from cognitive, social, clinical and applied psychology will be presented. In addition, it will be shown how theories of causal reasoning provide new insights and offer new theoretical accounts.

York Hagmayer – Decision Making

Causal considerations must be relevant to making good decisions. Nevertheless, most current decision making theories do not explicitly take causality into account. Experimental evidence will be presented showing that people are sensitive to the causal structure underlying a given decision problem. They tend to prefer the option that has the causal power to increase the probability of the desired outcome the most. A descriptive model of decision making integrating causal considerations will be introduced.

Dave Lagnado – Legal Decision Making

Causal models are critical to the evaluation of evidence in legal contexts, both from a normative and descriptive perspective. A series of experiments investigates how mock jurors combine several items of evidence to make judgments of guilt or innocence. The main findings are that people's reasoning is well-captured by qualitative Bayesian network models, but they are susceptible to biases due to the integrative encoding of information. A game-theoretic approach to modelling the informational interactions in the legal process is also proposed.

Clare Walsh – Judgments of Responsibility

People tend to ascribe more responsibility for actions and failures to act when both lead to the same unwanted outcome. We consider two alternative explanations for this difference. Counterfactual theorists propose that it is easier to imagine an alternative to an action than a failure to act. Other theorists propose that an action involves a process of transmission from cause to effect whereas failures to act do not. We tested these theories by examining judgments of responsibility for actions that change the outcome but do not involve a process of transmission.

Denis Hilton – Selection of Explanations

Causal explanations are often selected from complex causal chains. I examine what principles govern selections from *unfolding causal chains* and present data that suggest that: a) the proximal abnormal condition is selected, unless b) a voluntary deliberate action precedes it, and c) these selections are in part predicted by surprise value, i.e.: perceived (un)predictability of consequents given antecedents.

John McClure – Causal Attributions for Brain Injuries

I report studies that show how attributions about brain injuries are shaped by the visibility of markers of the injury and information about the normality of the behaviour for the individual and for the culture. Causal assumptions on these two parameters shape judgments about treatment and interventions for the persons, as well as court decisions about liability for the outcomes of the injury.

JF Bonnefon, R Da Silva Neves, D Dubois, & H Prade – Causal Transitivity

If A caused B, and if B in turn caused C, is it true that A caused C? Despite its critical importance, the issue of transitivity has been neglected in experimental research. Drawing onto a formal qualitative model from AI, we identify a key condition for accepting the transitivity property: A will be perceived to have caused C when B is generally diagnostic of A, that is, when there are few conceivable ways for B to occur in the absence of A. We present results supporting that prediction, and discuss them in relation to other theories of causation.

Discussant Fintan Costello

University College Dublin

Thinking With and Without Language

Does the language we speak shape the thoughts we think? We observe children who have not been exposed to an accessible model of language and ask whether they can perform spatial and number tasks previously hypothesized to depend on language. Two papers examine deaf children who have not acquired spoken language and who have not been exposed to sign language. While these children invent homesigns to communicate, their systems lack signs for many of the notions central to understanding spatial and numerical relations. Can these children perform non-linguistic tasks involving these relations? The third paper examines spatial tasks in four closely related genera, humans and three non-human primates who do not possess language. The symposium thus examines which tasks can – and cannot – be performed without human language.

Language and Numerical Cognition: The Case of Nicaraguan Homesigners

Elizabet Spaepen (liesje@uchicago.edu), Marie Coppola
Dept. of Psychology, University of Chicago, Chicago IL
Susan Carey, Elizabeth Spelke
Dept. of Psychology, Harvard University, Cambridge, MA
Susan Goldin-Meadow
Dept. of Psychology, University of Chicago, Chicago IL

Evidence from animals and pre-linguistic infants suggests that exact representations of small sets and approximate representations of the cardinal values of large sets can be developed without linguistic input. By contrast, cross-linguistic and cross-cultural evidence suggests that representations of large exact numbers may require access to a conventional count list. We tested six child and two adult homesigners, who do not have access to a conventional count list, on a variety of verbal (gestural) and nonverbal numerical cognition tasks. Some, but not all, of the homesigners created and used gestures for number. The results supported the hypothesis that learning a conventionalized count list is critical to representing the natural number system. Further research is needed to determine the nature of the homesigners' representations of numbers.

Spatial Language Potentiates Spatial Cognition: Turkish Homesigners

Dedre Gentner (gentner@northwestern.edu),
Dept. of Psychology, Northwestern University, Evanston, IL
Asli Ozyürek
Max Planck Institute for Psycholinguistics, Nijmegen, NL
Susan Goldin-Meadow
Dept. of Psychology, University of Chicago, Chicago IL
Ozge Gurcanli
Dept. of Psychology, Johns Hopkins University, MD

Previous work suggests that spatial language aids in spatial cognition, based on findings that preschool children performed better on a spatial mapping task when spatial relational terms (such as *top, middle, bottom)* were used. However, these children were learning English, and may have internally invoked spatial terms. Here we present more definitive evidence, by giving the same task to two groups in Turkey: homesigners, deaf children who are *not* learning a conventional language, and whose homesigns did not include gestures for spatial relations, and hearing children. The homesigners performed far worse than the hearing children. These findings suggest that spatial relational language may play a central role in promoting spatial reasoning.

Cognitive Inheritance and Cultural Override in Human Spatial Cognition

Daniel B.M. Haun (haun@eva.mpg.de)
Max Planck Inst. for Evolutionary Anthropology, Leipzig
Christian J. Rapold
Max Planck Institute for Psycholinguistics, Nijmegen, NL
Josep Call
Max Planck Inst. for Evolutionary Anthropology, Leipzig
Gabriele Janzen
Donders Centre for Cognitive Neuroimaging, Nijmegen, NL
Stephen C. Levinson
Max Planck Institute for Psycholinguistics, Nijmegen, NL

In order to investigate language impact on non-linguistic cognition we first need to know the structure of non-linguistic cognition. The present study systematically extends prior work with non-human animals by investigating the spatial relations of our closest phylogenetic relatives (*Pongo, Gorilla,* and *Pan*). If all great ape genera share a particular cognitive preference, it has most likely been passed on from the common ancestor, and is therefore part of the inherited defaults of human cognition. We found that all three non-human great ape genera prefer to process spatial relations based on environmental cues and not self. We also compared human children and adults from two linguistic communities with different dominant spatial relation representations. The data show a correlation between the linguistic representation and the preferred cognitive strategy both children and adults used to process spatial relations. The model for human cognition that we propose then, has a rich, inherited primate basis, which may be overlaid by language and culture.

Acknowledgments

Funding and support provided by NSF-ROLE grant #REC-0196471; NIH grant #5 R01 DC0049; NSF-ROLE grant # 21002/REC-0087516; and the Max Planck Institutes in Nijmegen and Leipzig.

The body: from experience to representation

Andrea Serino (serino@cnc.psice.unibo.it)
Faculty of Psychology, Centro studi e ricerche in Neuroscienze Cognitive, University of Bologna, Italy

Manos Tsakiris (e.tsakiris@ucl.ac.uk)
Department of Psychology
University College London, UK. (Symposium Organiser)

Katerina Fotopoulou (a.fotopoulou@iop.kcl.ac.uk)
Institute of Psychiatry, King's College, UK

Frederique de Vignemont (fvignemont@isc.cnrs.fr)
Institut des Sciences Cognitives, Lyon, France

Much of our mental life relates to our physical body. Despite the pervasive role of the body-representation concept in almost every aspect of psychology, it is yet unclear how the versatile nature of body-experiences, from multisensory perception and body-ownership to voluntary control of the body and tool-use, is represented in the brain. The symposium will focus on the behavioral, functional and neural correlates of body-specific processing from the perspectives of cognitive neurosciences cognitive sciences, clinical neuropsychology, and philosophy of mind.

A. Serino: "Extending the body space by long term tool-use experience"

Representation of body space extends to the space that is reachable by an arm movement. However, in everyday life we use tools to interact with objects placed in far space, suggesting that tool-use experience might extend the representation of space surrounding the body. Audio-tactile integration was studied in the space around the hand and in far space in blind and sighted subjects who used a cane to navigate, before and after a short training with the cane. In sighted subjects, before tool-use, auditory peripersonal space was limited around the hand, then it expanded after tool-use and contracted backwards after a resting period. On the contrary, in blind subjects, peri-hand space was immediately expanded when holding the cane but limited around the hand when holding a short handle.

K. Fotopoulou: "Is this My Hand is I see before me? Body Representation and Ownership in Anosognosia for Hemiplegia"

Anosognosia for hemiplegia (AHP) refers to the apparent unawareness of paralysis in neurological patients. The neurological and neuropsychological profile of six patients with severe AHP following right-hemisphere (RH) stroke will be compared to control patients showing similar RH lesions, hemiparesis, visuospatial neglect but no AHP. The experimental investigations addressed two main questions: a. Does motor intention influence the awareness of movement? b. Are there implicit emotional biases in AHP? The results suggest that motor intention has a profound influence on the on-line representation of one's actions, and emotional factors may have a top-down influence on one's body-representation. The relation of these findings to critical determinants of bodily representation and awareness will be discussed.

M. Tsakiris: "The bodily self: signatures of body-ownership in the brain"

We constantly feel, see and move our body, and have no doubt that it is our own. This sense of 'body-ownership' is a basic form of self-consciousness. Consistent psychophysical results suggest that body-ownership arises as an interaction between multisensory perception and representations of the body's permanent structure: current sensory integration is modulated by top-down processes reflecting a pre-existing reference of the postural and visual features of one's body. This functional interaction has identifiable neural signatures in the right hemisphere. The right temporo-parietal junction modulates the assimilation of sensory signals to a body-reference, while the subjective experience of body-ownership is correlated to activity in the right posterior insula. These structures may form a network that plays a fundamental role in self-consciousness

F. de Vignemont: "How many representations of the body?"

The body can be viewed from many different perspectives (e.g. semantic, emotional, spatial, motor, tactile, visual, proprioceptive, etc.) and described with many pairs of opposing properties: conscious/unconscious, conceptual/ non-conceptual, dynamic/stable, innate/acquired, personal/generic, spatial/non-spatial. How many body representations do we really have? One representation, integrating all the different types of information into a unified neuromatrix? Two representations, based on the model of visual perception, distinguishing the body image for recognition and the body schema for action? Three representations, for a more fine-grained distinction within the body image, disentangling the visuo-spatial bodily map and the body semantics?

Discussion:

"Somatognosia: scientific and philosophical perspectives" (Coordinator: K Fotopoulou)

Metacognition, Mindreading and Self-consciousness

Cristiano Castelfranchi
ISTC-CNR, Roma
cristiano.castelfranchi@istc.cnr.it

Joëlle Proust
Institut Jean Nicod, Paris
jproust@ehess.fr

Tjeerd Jellema
University of Hull
t.jellema@hull.ac.uk

Fabio Paglieri
ISTC-CNR, Roma
fabio.paglieri@istc.cnr.it

This symposium will bring together leading scholars from four large-scale ongoing research projects on the study of metacognition and consciousness, to discuss the following research challenges:

- *continuity between mindreading, metacognition and self-consciousness* in evolution and development;
- conceptual differences between *metacognition* and *metarepresentation*;
- the role of *sensory-motor capacities* in enabling both mindreading and metacognition.

Each of the three speakers will have a commentator and will reply to his critical remarks on the presentation. The three invited discussants for this event will be dr. Julian Kiverstein (University of Edinburgh), prof. Andreas Roepstorff (University of Aarhus), and prof. Jerome Dokic (Institut Jean Nicod, Paris).

In order to allow for in depth debate on each of the presentations, the symposium will have a duration of 3 hours, with a 15-minutes break after the first relation.

The *European Science Foundation* is financing the event, and a representative of the ESF, dr. Eva Hoogland, will open the symposium by briefly outlining the new EuroCORES Program "CNCC – Consciousness in a Natural and Cultural Context" (http://www.esf.org/cncc).

Joëlle Proust

Metacognition without metarepresentation

Commentator: Julian Kiverstein, University of Edinburgh

Metacognition is often defined as thinking about thinking. It is exemplified in all the activities through which one tries to predict and evaluate one's own mental dispositions, states and properties for their cognitive adequacy. This talk will discuss the view that metacognition has metarepresentational structure. Properties such as causal contiguity, epistemic transparency and procedural reflexivity are present in metacognition but missing in metarepresentation, while open-ended recursivity and inferential promiscuity only occur in metarepresentation. It is concluded that, although metarepresentations can redescribe metacognitive contents, metacognition and metarepresentation are functionally distinct.

Tjeerd Jellema

Developmental and neuroscientific perspectives on mindreading and action understanding

Commentator: Andreas Roepstorff, University of Aarhus

My talk will be on the neural basis of social cognition and ToM from an evolutionary perspective. A basic question is what are the neural structures and computational/cognitive processes that enable us to infer an individual's goal or intention, or to attribute a mental state (ToM), when all we see are their bodily postures, implied and actual articulated actions, gaze direction and facial expression. A working hypothesis is that differential but overlapping neural substrates exist for the forming of descriptions of social behaviour: (a) in terms of the 'mechanics', i.e. the physical causes, action sequences and consequences of actions allowing prediction of the most likely next behaviour ('behaviour reading'), and (b) in terms of the 'mentalistic' underpinnings of the behaviour ('mind reading'). Specific questions that will be looked at are: What are the necessary building blocks, or precursors, that lead to the forming of ToM, to what extent are they already present in non-human primates, how can we 'measure' their operations in humans, and to what extent is the automation of social cue evaluation compromised in autism?

Fabio Paglieri

From mindreading to mindchanging: Metacognition, social influence, and self-consciousness

Commentator: Jerome Dokic, Institut Jean Nicod, Paris

Regardless heated debate on the proper format of our capacities for understanding each other mental states, proponents of both simulationist and theoretical accounts of mindreading tend to agree on what are the main adaptive functions that are served by such alleged capacities: *interpreting*, *predicting*, and *coordinating* with the behaviour of others. In contrast, *influencing* the conduct of other agents through an adequate appreciation of their inner states does not usually figure among the main functions of mindreading. In this talk, I will argue that, on the contrary, influencing each other conduct is one of the most relevant powers engendered by sophisticated mindreading skills, and that this crucial capacity is intertwined with other functions of mindreading.

In particular, the presentation will aim to:

- outline a more fine-grained *theory of interference*;
- define *influence* as a specific type of interference;
- relate *types of interference with different cognitive capacities*, in particular influence and mindreading;
- show the *mutual interactions between different functions of mindreading*: explanation, prediction, coordination, and interference;
- discuss the import of the influence function of mindreading for a theory of *self-consciousness* and *strength of will*.

Papers

Elementary student nascent abilities for scientific argumentation

Loucas Louca (Louca.L@cytanet.com.cy)
Learning in Science Group, University of Cyprus
P.O. Box 20537, Nicosia 1678, Cyprus

David Hammer (davidham@umd.edu)
Department of Physics & Department of Curriculum and Instruction, University of Maryland
2226 Benjamin Building, College Park, MD 20742 USA

Abstract

We analyze a 5-6th grade class conversation about what will happen to a pendulum if it is released at the highest point of its swing. Using two frameworks for analyzing the quality of argumentation, we argue that prior to any formal instruction the students showed abilities for scientific argumentation. This and other evidence in the literature supports the contention that children have resources for argumentation from their everyday experience, which suggests a shift in orientation for instruction and research: Rather than work to instruct students in scientific argumentation, educators should focus first on recognizing and cultivating the abilities they already have.

Introduction

There has been substantial growth of research in science education on student argumentation (Kelly, et al,, 1998; Newton et al., 1999; Driver, et al., 2000; Jimenez-Aleixandre, Rodriguez & Duschl, 2000; Felton & Kuhn, 2001; Kelly & Takao 2002; Kuhn & Udell, 2003; Erduran et al., 2004). This work is motivated by old and widely subscribed views of a need for increased emphasis on students learning to engage in scientific inquiry (NSES, 1996; CSMEE, 2000; Minstrell & van Zee, 2000).

Although researchers have not formalized a single definition of argumentation, all share a basic view of it as the advancement and supporting of claims with logic and evidence, and of sophisticated argumentation as involving attention and response to others' reasoning (Erduran et al., 2004; Kuhn, 1993; Toulmin, 1958, Osborne et al., 2004; Jimenez-Aleixander et al., 2000; Driver et al., 2000). The literature is not so consistent, however, with respect to what educators should *expect* to see in students. By large, it either argues or assumes that prior to explicit instruction, students lack abilities for argumentation.

Our purpose in this paper is to argue for this interpretation, that children bring resources for argumentation with them to the classroom from their everyday experience, and that the central role of instruction at the outset should be to elicit, recognize, and promote those abilities (Hammer & Elby, 2003; Hammer, 2004). We make this case by using data from a 45-minute conversation in a combined 5th-6th grade science class of student who had never had formal instruction in scientific argumentation. We show that, analyzed using existing coding schemes (Erduran et al., 2004; Felton & Kuhn, 2001), there is ample evidence of students' nascent abilities.

Views about student abilities for argumentation

A large body of the recent literature on argumentation (Kelly et al., 1998; Jimenez-Aleixandre et al, 2000; Erduran et al., 2004; Osborne et al., 2004) draws on Toulmin's (1958) "argument pattern" (TAP), defining argumentation in terms of *claims*, *data* that support the claim, *warrants* for linking the data to the claim, *backings* for the warrants in the form of generalizations from a body of experience, and *rebuttals* that attempt to "undermine the force of the supporting arguments" (Toulmin, in Erduran el al., p. 918).

Erduran et al. discussed methodological challenges in applying TAP to analyzing student argumentation, such as the ambiguity of distinguishing between data and warrants or between warrants and backings. Following earlier work (Kelly et al., 1998), they simplified the scheme to focus on *claims*, *justifications*, and *rebuttals*, where justifications may be any use of data, warrants or backings to support the claim, and rebuttals are attempts to identify flaws in the justifications for an opposing view.

Kuhn and her colleagues (Felton & Kuhn, 2001; Kuhn & Udell, 2003) draw on Walton's (1989) work, rather than Toulmin's (1958), but their scheme has similar features. They quote Walton's description of skilled argumentation as having two goals, one to "secure commitments from the opponent that can be used to support one's own argument," and the other "to undermine the opponent's position by identifying and challenging weaknesses in his or her argument." This supports Kuhn and her colleagues to see sophisticated argumentation as involving recognizing and responding to the opinions, reasons, and evidence of an opponent's position. Felton & Kuhn (2001) identify "challenge type" discourse including "Counter-A" and "Counter-C," defined as disagreements accompanied by an alternative argument (A) or a critique (C).

The two schemes both depict sophisticated argumentation to involve not only supporting one's own views with reasoning but also attending and responding to others' reasoning. In this way, they represent a consensus in the literature for what educators should hope to see in students' argumentation.

Expectations for student abilities

The literature is not so consistent, however, with respect to student abilities for scientific argumentation. Kuhn and her colleagues have pursued a developmental view, documenting limitations of abilities for coordinating theory and evidence in children and scientifically naïve adults (Kuhn, 1989, 1993; Felton & Kuhn, 2001; Kuhn & Udell, 2003; Kuhn et al., 2004). Kuhn & Udell suggest that there are serious weaknesses in the skills of adolescent and young adults. Although children encounter evidence from early ages, the coordination of evidence does not take place at a level of conscious awareness or control until much later.

Kuhn's findings of developmental limitations has been the subject of much debate. Metz (1995) argued that developmental accounts have systematically underestimated children's abilities in science. Koslowski (1996) criticized Kuhn's work for placing too much emphasis during the interviews on covariation of evidence and not enough on mechanism. When presented with simpler tasks, even first and second graders can distinguish between theory and evidence (i.e., Samarapungavan, 1992).

A number of studies give evidence of students' abilities for a "natural form of argumentation" (Jimenez-Aleixandre et al., 2000). Given appropriate opportunities, these studies suggest, students will draw on these abilities (Gallas, 1995; Hammer, 2004). In fact, this view of student abilities has some support in the findings of most studies. Erduran et al.'s (2004) analysis records instances of high-level argumentation in 12-14 year old students pre-instruction, almost as many as it records for the students post-instruction. Kuhn & Udell (2003) report increases in "challenge type" codings of argumentation as a result of instruction, but there are instances at the outset. Similarly, McNeill et al.'s (2006) analysis identified pre-instructional instances of students using evidence to support claims.

There is evidence of argumentation in many case studies of children's reasoning (Gallas, 1995; Hammer and van Zee, 2006). But it is not always easy to discern, which raises an important possibility: Educators may think that abilities for argumentation do not come naturally because the abilities are difficult to recognize in their natural forms. If so, it is a crucial agenda for research to learn better to identify nascent abilities, to identify circumstances in which they tend to appear, and to make progress in understanding how they develop into expert abilities. As well, if the developmental issue involves meta-level awareness and control, as Kuhn and her colleagues suggested, it is important to identify the "inchoate forms" of argumentation such that children recognize them as well.

Methodology

Like other recent efforts, we are studying argumentation as it occurs in a classroom setting rather than in clinical interviews, because we expect that children's abilities depend on context. In particular, we examine a 45-minute conversation among students in a combined fifth and sixth-grade class (17 fifth and 11 sixth grade, ages 10-12) at a public elementary school in Prince George's County, Maryland. It was a racially diverse class, reflecting the school population, with African Americans the majority.

Our purpose in this paper is to document children's pre-instructional abilities for scientific argumentation, as defined in the literature. To that end, we base our main analysis on the coding scheme presented in Erduran et al. (2004) focused on identifying rebuttals. Erduran et al.'s. (2004) scheme consists of a five-level framework, summarized in Table 1, for coding episodes of opposition they identified in student conversations.

We present the results of our analyses at two levels of explication. Our main analysis consists of coding 16 minutes of continuous conversation (120 utterances) using Erdurna's at al (2004) coding scheme. Second, we apply a narrative analysis, to examine five minutes from the conversation in detail. We present the transcript of these five minutes, providing the results from (i) the Erduran et al. (2004) coding scheme, along with (ii) Felton & Kuhn's (2001) coding scheme, accompanied by further insights from the narrative analysis.

Adapting and applying an existing coding scheme

The units for analysis in Erduran et al.'s (2004) coding were *episodes* of opposition they identified in the data as a first step. We have modified their framework slightly to take *conversational turns* as the units of analysis, because we found it difficult to segment the transcript into episodes. That change obviated the difference between levels 4 and 5 in their framework. Rather than identify oppositional episodes, we coded 16 uninterrupted minutes of the conversation, coding each conversational turn in the transcript by the first 4 levels of Erduran et al.'s scheme.

Conducting the analysis below, we worked principally from the transcript of the conversation, consulting the video only in particular moments. Each of us coded 16 minutes of transcript (120 utterances) independently by our adaptation of Erduran et al.'s (2004) scheme. Our agreement was 80% (Cohen's Kappa: 71%), and we resolved disagreement through discussion.

To provide further support for our claim of nascent abilities, we chose a five-minute segment of the conversation for additional coding by a second scheme by Felton & Kuhn (2001). Because we were studying a conversation among more than two students, we adapted their codings to substitute "another student" for "partner." We also allowed that the preceding utterance might not be immediately preceding in the transcript, such as when a student waits several conversational turns for the chance to respond to a particular statement.

Narrative analysis of student discourse

For our second analysis we used narrative analysis seeking to examine the details and to make explicit our assessment of student argumentation. Like Kelly et al. (1998), we expect that the substantive details and context of the conversation are critical and need to be examined; unlike those authors we work from the whole of the conversation rather than try to isolate and code only specific utterances. This difference in approach stems in part from a difference in theoretical perspective: We expect that student abilities for argumentation may be activated or not in different contexts, including within the conversation itself (Louca, Hammer & Bell, 2002). In this analysis, then we attend closely to the conversational context.

Level 1 argumentation consists of arguments that are a simple claim vs. a counter-claim or a claim vs. a claim. **Level 2** argumentation has arguments consisting of a claim vs. a claim with either data, warrants, or backings but do not contain any rebuttals. **Level 3** argumentation has arguments with a series of claims or counter-claims with either data, warrants, or backings with the occasional weak rebuttal. **Level 4** argumentation shows arguments with a claim with a clearly identifiable rebuttal. **Level 5** argumentation displays an extended argument with more than one rebuttal.

Table 1: Analytical framework used for assessing the quality of argumentation (from Erduran et al., 2004)

Data synopsis

We begin with a brief description of the conversation as a whole, and in the next section provide findings.

The teacher started the conversation by showing students a pendulum made of a metal washer and a string, and asked the class to explain how the washer would move if the string were cut just when the pendulum had swung to its highest point. As we recount elsewhere (Louca et al., 2002), for the first several minutes of the conversation the students' contributions consisted almost entirely of claims for what would happen, with little or no justification, even with the teacher prompting. Two students sketched their answers on the board, as shown in Figure 1: Chris[1] drew answer 1; Ike answer 2.

Victoria (49)[2] then provided answer 3, in sharp contrast with Chris's and Ike's, that the washer would fall straight down, explaining that "the string is gonna kind of curve," but "gravity is gonna push it [the washer] down" (56).

Victoria's idea sparked a new level of energy in the class —with students competing for the floor. Jeff (60) seconded her answer; Brandon (61) disagreed, and Shadawn (62) said that "it depends on how fast it's going," that "if it's going really fast and you cut it, it's gonna fly somewhere," but if it's "going really slow" then "it's just gonna go straight down." Mathew (64) offered an analogy in support of Shadawn's idea, comparing the pendulum question to what happens when someone swings from a rope into a lake. Vanice (65) objected to that comparison, because "the washer is tied to the string" but the person who swings from the rope lets go of it.

The conversation continued, lasting about 30 minutes in all. Much of it concerned the students discussing how the motion of the bob may depend on its weight, how quickly it is moving when it is released, and where it is released in the swing. In many instances, they explain their reasoning using comparisons to other situations or objects, including swinging keys, heavy weights on cranes, or a pencil.

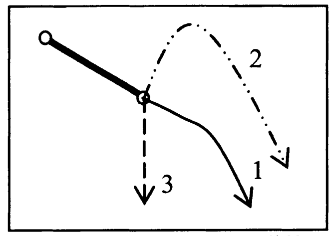


Figure 1: Summary of the three different ideas

Analysis and findings

A segment of conversation

64. Mathew: Can I say something? I, I agree with, um, Shadawn because it's kind of like, you, you have a little, like you know how some times on movies and things and real life, they have lakes or swimming pools and you have a little rope and you run and grab on to the rope and then fly and then let go and you go flying over to the side? That's just like that, the washer. It depends on how much force is on it.
65. Vanice: Not exactly Mathew, because the pendulum is, I mean the washer is tied to the string so it won't go to the other side.
66. Mathew: But she is cutting it, or she'll let it go.
67. Vanice: I know but it still not going to go to the other side because it's hooked together, if it wasn't hooked together then yeah it might go to the other side, like the str, the string would still be in your hand but the pendulum, I mean the washer will go somewhere else.
68. Mathew: I know that. It's kind like, it's kind like, um, the person flying off of it letting go and then going into the water.
69. Vanice: I know but it's not connected I mean it's connected so that wouldn't work.
70. Grace: Well um, I agree with Chris because um it can't really go up more because like gravity doesn't go up. And like I don't think it can just go straight down because I think you're swinging it.
71. Mathew: I disagree with, I disagree with Grace because, because it's kind of like you throw a bucket or a ball up in the air, gravity is coming down forcing it to come down, but you still, it still going up.

Mathew and Vanice – Coding & narrative analysis

The segment starts with Mathew agreeing (64) with Shadawn that the answer to the question depends on "how fast it's going" (62), justifying his agreement with a

[1] We have obtained permission to use students' real first names.
[2] Numbers in parentheses refer to the conversational turn in the transcript. The portion we analyzed begins at with Victoria's comment, the 49th turn in the transcript.

comparison to the experience of swinging from a rope into water. Vanice took issue with that comparison (65), pointing out a difference: The string stays attached to the washer, but the rope does not stay attached to the swimmer.

Mathew's initial contribution (64) was a clear example of a level 2, since there was a justification but no rebuttal. We coded Vanice's response (65) as level 4, because she provided data against the comparison, undermining Mathew's justification for his claim.

Mathew's responses (66 & 68) to Vanice were more difficult to code. One of us originally coded them each as level 1, on the rationale that it was simply a claim; the other coded them as level 4, on the rationale that Mathew was taking issue with Vanice's reasoning in her rebuttal: She had said the washer is tied to the string, and here Mathew was countering that the string would be cut or released (66) and arguing that her reasoning actually supports his position (68). We coded them each as level 2, treating his statements as justifications for why the two situations are comparable.

Analyzing by Felton and Kuhn's (2001) scheme, we saw most of the students' contributions as Counter-C, because the speaker's strategy was to disagree with the preceding remark while providing a critique. For example, Vanice's objection to Mathew's comparison criticized it (65), and Mathew's response to that criticism (66) was a counter in itself. Mathew's argument (64) showed another strategy (Advance) in his elaboration of Shadawn's previous argument. We coded Vanice's elaboration of her concern (67) as Counter-A, because there she provided further argumentation for her position, and Mathew's response (68) to that as a Coopt, because his approach was to try to incorporate her explanation into his own view.

Assessed by either coding scheme, the students were showing the beginnings of abilities for argumentation. We see further strengths to what Mathew and Vanice were doing, beyond what the coding schemes recognize.

First, Mathew's argument (64) was the first reference in the conversation to another situation, articulating the connection between the pendulum question and familiar experiences. That is a strength in his justification: It invokes knowledge he expected others would share and find compelling. Second, Vanice's response (65) attended to the details of the comparison he offered, noticing a specific difference that was likely connected to her sense of the mechanism by which the washer would move: With the string attached, she seemed to be thinking, the washer would still be tethered, "so it won't go to the other side." We noted further that her critique focused specifically on the relevance of the experience that Mathew provided as grounds to his claim. She did not challenge Mathew's reasoning about what happens on rope swings; she challenged whether what happens on rope swings can be compared to the washer on the string.

We also note a weakness. Part of the reason for our ambivalence in coding Mathew's responses to Vanice's argument (66,68) was that he did not understand her point. He probably intended to be undermining the force of her argument (Counter-C), and he was providing justification to support the comparison he was making (level 2), but it is clear he missed her meaning.

Grace and Mathew - Coding & narrative analysis

When Grace agreed (70) with Chris, she supported his answer by arguing against the competing possibilities: "it can't really go up more because gravity doesn't go up," and it cannot "just go straight because … you were swinging it." Mathew (71) responded to her first argument, again comparing the situation of the washer to other, more familiar situations: Grace's first point cannot be correct, he argued, because it is clearly possible to "throw a bucket or a ball up in the air," even with "gravity coming down."

We coded Grace's contribution (70) as level 2, because she provided a claim, in her agreement with Chris, and she supported it with two lines of justification. Someone might argue that her respective reasons for supporting Chris were rebuttals of the alternative answers, but we did not see evidence that she was attending to the reasoning that supported those claims. We coded Mathew's contribution as level 4, since he provided a rebuttal to the justification Grace had provided. Had we isolated this pair of statements as an episode to code as a unit we would have counted it as level 4, rather than coding one turn at level 2 and the other at level 4. According to Felton and Kuhn's scheme, Grace was advancing Chris's prior claim, and Mathew's strategy was Counter-A, to disagree with Grace and provide an argument in support of his disagreement.

Again, Mathew was referring to a shared, everyday experience to make his counter-argument, and in the process he was drawing a clear distinction between the motion of an object and the influence of another causal agent. Part of what is impressive about his argument is its clarity in that distinction, which is generally considered to be a difficult one for students to make.

We suspect, however, that Grace did not fully explain her reasoning: She was probably arguing that the washer cannot go up more once it has reached the apex of its swing, without articulating a sense that, at that point, the washer's own upward motion had run out. For it go up higher, something would have to push it higher, but "gravity doesn't go up." Here, then, Mathew may have again misunderstood the reasoning on the other side.

Coding levels of rebuttals

Figure 2 shows the total number of conversational turns we coded at each level: 18 at level 1 (34%), 17 at level 2 (32%), 8 at level 3 (15%) and 10 at level 4 (19%). The codes at level 4 involved turns by five different students; the codes at level 3 involved 4 additional students. Thus nine different students showed they were capable of some level of rebuttal over the course of the 16 minute excerpt.

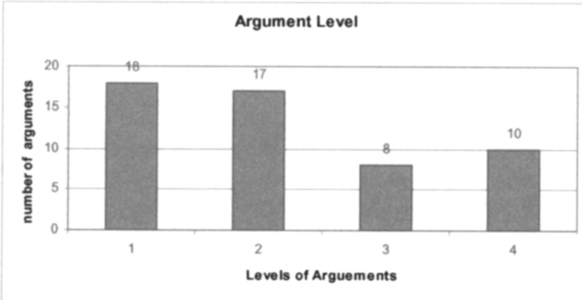


Figure 2: Summary of coding findings over 16 minutes of conversation

Our results are qualitatively comparable to Erduran et al.'s (2004): Across 43 discussion groups, pre-instruction they coded 28 episodes at level 3 or above (40% of coded arguments), post-instruction 36 (55%). In our case, analyzing a single 16-minute excerpt, we coded 18 utterances (34%) at level 3 or above. This supports our claim that children arrive pre-instruction with nascent abilities, including for generating rebuttals to others' arguments. We found further support to this claim from applying the Felton & Kuhn's (2001) scheme to the brief segment of data, which resulted in several codes of challenge-type discourse.

Quantitative comparison would not be meaningful for several reasons. First, as we noted above, we chose to code conversational turns rather than oppositional episodes, which resulted in many more codes than Erduran et al. (2004). Second, we understand the activation of student abilities for argumentation as sensitive to context, including fine-grained contexts within the course of a conversation. Different situations will trigger different argumentation.

Figure 3 includes coding results from the whole conversation. It displays the shift that followed Victoria's contribution (49), making apparent the increase of level 3 and the appearance of level 4 arguments. Students in the 16 minutes that followed Victoria's contribution do not engage in upper-level argumentation all the time, but they shift from level 1 through level 4 and vice versa. A topic for further study will be to understand what prompts these shifts during the conversation.

Conclusions and implications

Children's abilities for argumentation

In this paper we have documented 5th and 6th graders' making claims, supporting those claims with reasoning, and attending and responding to each others' claims and reasoning. The students had had no formal instruction in argumentation, and there was no particular teaching agenda for this discussion that they engage in argumentation. Still, they showed levels of argumentation comparable to the pre- or post-instructional levels reported in Erduran et al.'s (2004) study.

Of course, we do not claim that the students are already experts. There are several respects in which it is evident they are not, such as students misunderstanding each other's ideas, varying levels of articulateness, as well as the time it took the class to move past the initial phase of simply stating claims. That is, while there is evidence of students' abilities for argumentation, there is also evidence that they do not always use those abilities, as displayed in Figure 3.

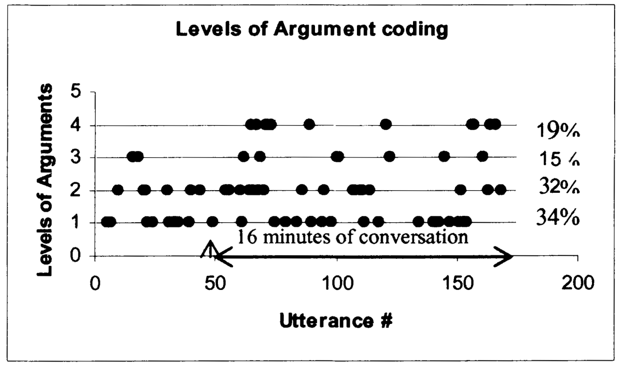


Figure 3: Tracking levels of argument coding over the course of the conversation

Our conclusion then is that children come to science instruction with nascent abilities for argumentation, that they may invoke spontaneously directly in line with the abilities educators have described as important to impart. In this we support the arguments in the literature that students have "natural" abilities (Jimenez-Aleixandre et al., 2000) for argumentation, as well as evidence from elsewhere, including Erduran et al.'s (2004) findings of high level episodes pre-instruction, Kuhn and Udell's (2003) of pre-instructional instances of counter-arguments. However, students' use of those abilities seems to be sensitive to context. That sensitivity provides an explanation for evidence in some studies (Kuhn, 1991, 1993) that show limitations: Depending on their design the contexts in different studies may not tap students' productive resources. This also provides an alternative account of the gains documented in various instructional studies (Erduran et al., 2004; Kuhn et al, 2006; Kuhn & Udell, 2003; McNeil et al., 2006): Improved performance may arise mainly from the activation of students' existing abilities, rather than from the acquisition of new abilities.

Implications for instruction

The first implication of this view for instruction is that the initial focus of attention should shift to children's current abilities, and the second is that teachers should be cautious in ascribing developmental limitations. That expectation should guide both the setting of objectives and the assessment of what takes place in class.

Rather than expect that "the use of valid argument does not come naturally" (Osborne et al, 2004, p. 996), educators should expect abilities for argumentation exist in "natural form" (Jimenez-Aleixandre et al. 2000), and that what comes naturally depends on the context. And so, rather than expect that argumentation must be "explicitly taught through suitable instruction, task structuring, and modelling" (Osborne et al, 2004, p. 997), educators should see the first step as helping children tap the resources they already have (Hammer and Elby, 2003). In other words, the initial emphasis should be on *eliciting and supporting* argumentation rather than on *instructing* students in it.

If the first step for instruction is to create contexts in which students are inclined to draw on their resources for argumentation, a subsequent step must involve helping them become aware of what they are doing, toward "enhanced metalevel awareness" (Kuhn and Udell, 2003). As the case study we presented illustrated, children may enter or leave productive argumentation, and so part of their development toward expertise should involve their developing more reliable access to those abilities.

Rather than suppose students lack abilities for argumentation pre-instruction, researchers might suppose they have abilities but do not use them reliably. On this view, some of the improvement documented in instructional studies (Kuhn & Udell, 2003; McNeill et al., 2006; Osborne et al., 2004) reflects increased use of abilities that students had already developed. This difference in interpretation would have substantial implications for research and instruction. Understanding children as lacking abilities, educators design instruction to help children form those abilities, and research focuses on assessing progress in that formation to identify factors that lead to the greatest improvement. In contrast, on the view that children already have abilities for argumentation, educators would design instruction to help children draw on those abilities, and the objective for research would be to understand their nature.

Acknowledgments

This study was supported by the NSF Award ESI-9986846.

References

Center for Science, Mathematics, & Engineering Education [CSMEE] (2000). *Inquiry and the National Science Education Standards a guide for teaching and learning*. Washington, DC: National Academy Press.

Driver, R., Newton, P. & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, 84(3): 287-312.

Erduran, S., Simon, S. & Osborne J. (2004). TAPping into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. *Science Education*, 88(6): 915-933.

Felton, M. & Kuhn D. (2001). The development of argumentive discourse skill. *Discourse Processes*, 32(2-3): 135-153.

Gallas, K. (1995). *Talking their way into science: hearing children's questions and theories, responding with curricula*. NY: Teachers College Press.

Hammer, D. & van Zee, E. H. (2006). *Seeing the science in children's thinking: Case studies of student inquiry in physical science*. Portsmouth, NH: Heinemann.

Hammer, D. & Elby, A. (2003). Tapping students' epistemological resources. *The Journal of the Learning Sciences*, 12 (1): 53-91.

Hammer, D. (2004). The variability of student reasoning, lectures 1-3. In E. Redish and M. Vicentini (Eds.), *Proceedings of the Enrico Fermi Summer School, Course CLVI* (pp. 279-340): Italian Physical Society.

Jimenez-Aleixandre, M. P., Rodriguez, A. B. & Duschl, R. A. (2000). "Doing the lesson" or "doing science": Argument in high school genetics. *Science Education* 84(6): 757-792.

Kelly, G. J. & Takao, A. (2002). Epistemic levels in argument: An analysis of university oceanography students' use of evidence. *Science Education* 86(3): 314-342.

Kelly, G. J., Druker, S. & Chen, C. (1998). Students' reasoning about electricity: combining performance assessments with argumentation analysis. *International Journal of Science Education*, 20(7): 849-871.

Koslowski, B. (1996). T*heory and evidence: the development of scientific reasoning*. Cambridge, Mass: MIT Press.

Kuhn, D. & Udell, W. (2003). The development of argument skills. *Child Development*, 74(5): 1245-1260.

Kuhn, D., Katz, J. B. & Drean Jr, D. (2004). Developing reason. *Thinking & Reasoning*,10(2):197-219.

Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review*, 96(4): 674-689.

Kuhn, D. (1993). Science as Argument: Implications for Teaching and Learning Scientific Thinking. *Science Education*, 77(3): 319-337.

McNeill, K. L., Lizotte, D. J., Krajcik, J., & Marx, R. W. (2006). Supporting Students' Construction of Scientific Explanations by Fading Scaffolds in Instructional Materials. *The Journal of the Learning Sciences*, 15(2): 153-191.

Metz, K. (1995). Reassessment of developmental constraints on children's science instruction. Review of Educational Research, 65(2): 93-127.

Minstrell, J. A. & Van Zee, E. H. (2000). *Inquiring into inquiry: learning and teaching in science.* Washington, D.C., American Association for the Advancement of Science.

Newton, P., Driver, R and Osborne, J. (1999). The place of argumentation in the pedagogy of school science. *International J. of Science Education*, 21(5): 553-576.

National Research Council [NRC]. (1996). National Science Education Standards. Washington DC: National Academy Press.

Osborne, J., Erduran, S. and Simon, S. (2004). Enhancing the quality of argumentation in school science. *J. of Research in Science Teaching*, 41(10): 994-1020.

Samarapungavan, A. (1992). Children's judgments in theory choice tasks: Scientific rationality in childhood. *Cognition*, 45: 1-32.

Toulmin, S. (1958). *The uses of argument*. Cambridge: Cambridge University Press.

Walton, D. N. (1989). Dialogue theory for critical thinking. *Argumentation*, 3: 169–184.

Nascent abilities for scientific inquiry in elementary science

Loucas Louca (Louca.L@cytanet.com.cy)
Learning in Science Group, Department of Educational Sciences, University of Cyprus
P.O. Box 20537, Nicosia 1678, Cyprus

Zacharia C. Zacharia (zach@ucy.ac.cy)
Learning in Science Group, Department of Educational Sciences, University of Cyprus
P.O. Box 20537, Nicosia 1678, Cyprus

Abstract

In this case study we analyze a series of student conversations about projectiles and relative motion in a combined 5^{th}-6^{th} grade science afternoon club to provide detailed descriptions of student inquiry, seeking to contribute to the development of a better understanding of nascent student inquiry in classroom settings. Prior to any formal instruction, we contend these students "have" a repertoire of abilities, e.g. for mechanistic or analogical reasoning and argumentation. With ambiguities regarding productive science inquiry, findings from this study reveal new insights with respect to the challenge of diagnosing student progress in the classroom. We also suggest that the role of instruction should be less on the direct teaching of elements of student inquiry and more on helping students develop reliable access to those abilities.

Introduction

Despite decades of calls for promoting inquiry in elementary grades, the agenda has yet to establish in instructional practice (Hawkins, 1974; NSES, 1996; Minstrell & van Zee, 2000; Osborne et al., 2004; Louca et al., 2002) for a number of reasons. First, while there is a consensus for the importance of inquiry in science learning, the education community has yet to agree on precisely what is important. For many, inquiry is a method for learning science "content," and it is important because it is more effective than other methods. Others consider it as part of science, and as an objective of itself. Second, there is not a consensus regarding what "productive" inquiry entails, especially in the early grades. For example, what should teachers be looking for and trying to cultivate? Answers have varied from Hawkins's (1974) general appeal for "messing about" to more specific targets for developing "concrete" abilities of observation and controlling variables in experiments (see Metz, 1995).

With these ambiguities, in contrast to tangible and seemingly straightforward objectives of traditional content, it is difficult to sustain instructional attention to student inquiry (Hammer, 1995). Regardless of any particular account of children's inquiry, there is the challenge of diagnosing student progress in any classroom situation.

To make progress in promoting student inquiry, science education needs to develop better understanding of student abilities for nascent inquiry in classroom settings. The purpose of this case study is to provide detailed descriptions of nascent student inquiry from the authentic learning context of the science classroom. We analyze a series of student conversations in science, identifying instances of productive student inquiry and looking for different elements of scientific inquiry that include mechanistic and analogical reasoning and argumentation. At the same time, we seek to speak to the debate about the development of student abilities for scientific inquiry, pointing to data that suggest that students without any formal instruction "have" the beginnings of those abilities. Our motivation comes from two directions. First, we seek to contribute to research aiming to help teachers understand how student inquiry looks in the science classroom and what they should be looking for to evaluate student progress regarding inquiry. Second, we seek to contribute to a growing body of research (e.g., Koslowski, 1996; Metz, 1995), suggesting that it is more productive to view students as "having" abilities for scientific inquiry and need to develop reliable access to.

Following current emphasis in science education for studying classroom-based scientific discourse, in this paper we adopt an analytic framework of recent research in science education about what constitutes student inquiry in the elementary science classroom, focusing on a number of different elements of student inquiry in science that have been highlighted in recent literature as central elements of scientific inquiry (Louca & Hammer, submitted).

Elements of scientific inquiry

Working from a variety of perspectives and intellectual traditions, the literature about elements of student abilities for scientific inquiry shows a general consensus with respect to the sorts of things we should value and try to promote in children's inquiry. That consensus, however, does not extend to the definition of what scientific inquiry looks like in the science classroom. It competes in particular with a widely-shared, if mostly tacit, sense of inquiry as a pedagogical strategy, a method for teaching the traditionally construed "content" of science. By this view, assessing the quality of children's inquiry is equivalent to assessing their progress toward the correct answers in the canon of accepted knowledge. Indeed, science educators have a much clearer sense of the canon of accepted knowledge than of what constitutes "good inquiry." Thus, while it is comparably straightforward to determine whether they are *correct*, inquiry-oriented objectives remain ambiguous.

To offer a working definition, we take inquiry to mean *the pursuit of causal, coherent explanations of natural phenomena* (Hammer, 2004). That pursuit may take many forms, both experimental and theoretical; in whatever form, the instructional agenda is to help students learn to engage in that pursuit for themselves. In this paper we suggest that evidence from classroom discourse shows that scientific inquiry includes a number of different elements that have been offered through recent literature about student inquiry in science. These elements may abilities for argumentation, mechanistic reasoning, and analogical reasoning and the list can go on to include abilities for modeling discourse, design and implement experiments and controlling variables. For the case study that we present below, students did not engaged in any experimentation, and thus we do not address issues related with those abilities. By this we do not suggest that these are the only elements of scientific inquiry nor do we believe that science education research community is anywhere close to a consensus about what student inquiry looks like in the science classroom.

Mechanistic reasoning

One area of research is calling attention to the scientific discourse that involves causal mechanism (Russ et al., submitted), following a number of studies that partly focused on student use of causation in science (i.e., Schauble, 1996; Koslowski, 1996). This suggests that "assessing when and how students seek causal mechanism in their understanding should be part of assessing their reasoning as inquiry" (Russ et al., submitted, p.1)

Using a framework derived from the philosophy of science, Russ et al. (submitted) propose a coding scheme of 7 major components of mechanistic reasoning that can be used to identify and assess student use of mechanistic reasoning. Those components include (i) descriptions of the target phenomenon (what we see happening), (ii) identification of the set-up conditions that are necessary for the phenomenon to happen, (iii) identification of entities (conceptual or real objects that play a particular role in the phenomenon, (iv) identification of (iv) the entities activities that cause changes in the surrounding entities, (v) the entities properties, and (vi) the entity organization (how entities are located, structured or oriented within the phenomenon), and (vii) chaining, that is using knowledge about causal structure to make claims about what has happened prior to a phenomenon and what will happen.

Analogical reasoning

Expert scientific inquiry also involves the generation, use and evaluation of analogies (May et al., 2006) because analogies can be valuable tools for constructing one's own understanding in a variety of contexts. Unlike most of the research about analogies (focused almost exclusively on their pedagogical value in curriculum materials and teacher explanations for promoting conceptual change in students (May et al., 2006)), we are interested in student abilities for analogical reasoning that includes (i) the generation of analogies (that includes a target case (unknown), a base case (known) and a relation that maps elements from one case to the other), (ii) the evaluation of the validity of an analogy and subsequent refinement (that includes identification of the key features of an analogy and its limitations), (iii) the use of analogies to create new knowledge by making new inferences about the target case and creating abstract generalizations and (iv) the use of analogies to communicate ideas in science to others (Clement, 1998; May et al., 2006).

Argumentation

Argumentation is one of the areas that research has made significant progress in understanding and defining it. Kuhn (1989; 1993) was the first to call attention to inquiry as an essential objective for science education, focusing specifically on abilities for coordinating theory and evidence. A number of recent efforts have focused on analyzing the sophistication of student arguments in science. Louca and Hammer (submitted) propose a framework for studying argumentation discourse in the science classroom, specifically focusing on students abilities to generate, use and evaluate arguments. Their framework consists of a modified coding scheme adopted from Erduran et al. (2004) looking for components of arguments that include (i) claims, (ii) grounds, (iii) counterclaims, and (iv) rebuttals.

Views about the development of student abilities for scientific inquiry

The disagreement about what educators should *expect* to see in children's inquiry in the science classroom also includes ongoing differences with respect to the development of abilities of scientific inquiry.

One view follows a developmental approach, focusing on the development of student abilities. Evidence from a number of studies (Kuhn, 1989; 1993) suggests that abilities for i.e. scientific argumentation increased with the subjects' age, suggesting that this ability may be part of general cognitive development providing evidence for a developmental trend in particular in argumentation (Kuhn & Udell 2003).

A second view has argued that developmental perspectives have systematically underestimated children's abilities providing differences in findings that reflect the contexts of the interviews and framing of the questions (Metz, 1995). These concerns are supported by evidence from psychology and education research regarding the universality of abilities and developmental stages (Feldman, 1994; Karmiloff-Smith, 1992; Koslwoski, 1996; Samarapungavan, 1992). On these accounts, human knowledge and reasoning is far more variable than traditional developmental schemes have indicated. Dunbar and Blanchette (2001) describe dramatic differences in the phenomenology of analogical reasoning between *in vitro* studies and their in *vivo* observations: Uses of analogy that are difficult to produce in the laboratory occur easily and spontaneously in naturalistic settings.

A third approach has argued that abilities for i.e. argumentation can and should be explicitly taught as early as in elementary school, including abilities for scientific argumentation (e.g., Erduran et al., 2004; Osborne et al., 2004). This view has motivated research to develop pedagogical practices that specifically support aspects of scientific inquiry (Osborne, Erduran et al., 2004), also suggesting that prior to any intervention students' skills are poor (Bugallo Rondriguez & Jimenez-Aleixandre, 1996).

We, on the other hand, suggest that children come to class already with abilities in engaging in scientific inquiry, and that teachers need to help them refine those abilities (not teach or develop) and most importantly to help them develop reliable access to those abilities for using them in the right context and time. We feel that the literature over-emphasizes the need to actually teach students how to construct, evaluate and respond to causal mechanism, analogies and arguments. Most of research in classroom-based argumentation discourse (Erduran et al, 2004; Kuhn & Udell, 2003) provide some evidence that children have at least *the beginnings* of abilities regarding argumentation. In this view, we will use our analysis below to suggest that we need to reconsider the fact that children may have already appropriated the beginnings of inquiry practices.

Methodology

This interpretive case study illustrates young children's nascent abilities for scientific inquiry. Data originate from a larger research study funded by the Cyprus Research Promotion Foundation aiming to develop case studies of student inquiry as professional development materials for science teachers.

This case study involves a group of 15 fifth and sixth grade students who volunteer for participating in an afternoon science club at their school. Data originate from 4 90-minute whole-class conversations about a combined projectile and relative motion that were facilitated by the club teacher. For the purpose of this paper, we focus on discourse-based data looking for different aspects of scientific inquiry that students use in the conversations. The conversations took place during March 2005, in the context of developing models of the phenomenon. Students had no prior formal instruction about any of the 3 elements of student inquiry that we are investigating.

We analyzed transcripts of student discussions using analysis of student conversation, following a current trend in research in science education focusing on classroom discourse in science and mathematics (e.g., Ball, 1993; Gallas, 1995) and shares the interest of the science education community in classroom discourse. In doing this we seek to describe the variability in students' scientific inquiry, as well as contextual possibilities that might have lead to different uses of different elements of scientific inquiry. This analysis uses transcribed conversations as a gateway to student thinking (Edwards & Mercer, 1995).

After transcribing all videos of whole-class conversations of this case, we skimmed the transcripts independently, identifying episodes in the conversations that fall under the three elements of scientific inquiry. After agreeing on 46 episodes, we characterized them independently identifying the components of each element based on the literature that we presented above. Our inter-coder agreement was 89% and we resolve disagreements over discussion. Below, we briefly present 5 short analyzed excerpts that are representative of the findings to support our claims.

The conversation that we present below started by the teacher by stating the question: "There is a boy standing on a moving hallway at a local airport. The boy is holding a ball in his hand. Suppose he throws the ball up in the air, where would the ball land?"

Findings

The presentation of findings below is structured following the temporal sequence of how things happened over time.

Asserting answers

At the outset, students simply described what they think would happen, disagreeing over two possibilities (the ball would fall in the boy's hand, or behind the boy) but doing little to justify their answers, without making any progress as to what causes the ball to fall either back to the boy's hand or behind it.

> 17. Myriani: Since his hand is open like that, when he'll throw the ball up this way, he'll move a little, and thus the ball will come back down and hit him on his head[1]. [...]
> 27. Dioni: It will fall behind the student.
> 28. Teacher: Why do you think that?
> 29. Sabina: As soon as he throws the ball, he moves. But the ball is going to fall back to the same point that was thrown initially. Therefore, it is going to fall behind him. [...]
> 32. Dioni: Well, the... the student is..., well, he is moving with the moving hallway, but the ball is going to fall back to the same point [that was initially thrown from], and therefore the ball will fall behind him[...]
> 61. Teacher: Ok. If you think the right answer is this [it will fall back in his hand], why do you think that the other answer [will fall back in his hand] is wrong?
> 62. Nasia: Well, if he's throwing it while he is moving, it [the ball] cannot fall back to his hand.

At the beginning of the conversation, students described the story of the physical system under study, by describing what would happen eventually, without providing any explanations as to why all these happen or how they happen. Although their answers seemed to have an underlying mechanism that could explain what they describe (17, 29, 32)[2], students neither articulated it nor addressed it, even when they disagreed with each other. In terms of mechanistic reasoning (Russ et al., submitted) the students

[1] Student conversations are translations from greek.
[2] Numbers represent utterance number from the transcript.

described the target phenomenon without any references to entities, their properties and their organization, which are more sophisticated elements of mechanistic reasoning. Even when the teacher prompted them to explain their ideas (28, 61), their answers were simply re-statements of their ideas (29, 62). In terms of argumentation students were simply constructing and providing claims without any justifications or grounds whatsoever.

As the conversation continued several students offered ideas about dependencies that affect the phenomenon.

> 37. Myriani: I think it depends on whether he would move a little or more. I mean, if he throws it and moves a little, then the ball might fall just in front of him. But if he moves a lot, then the ball would fall behind him.
> 38. Teacher: So, are you saying that it depends on the speed of the hallway?
> 39. Panayiotis: Can you tell us how much is the speed of the hallway? [...]
> 45. Erini: Well, it also depends on how high he throws the ball.
> 46. Kyrilos: That's exactly right! If he throws it high, he will move much further, but if he doesn't throw it high, the ball can even land on his head!

According to Russ et al.'s (submitted) scheme, in this mode of work students identified a number of possible set-up conditions that could affect the mechanism that produced the phenomenon. Myriani (37) was the first to suggest that the speed of the moving hallway can affect where the ball would land, indicating that if the boy moves only a little, the ball would fall in front of the boy, but if he moves a lot, then the ball would fall behind him. Myriani's idea may had sparked Erini's idea (45) that the higher the ball would thrown, the further back from the boy it would fall.

Beginnings of student inquiry

In line 131, the teacher decided to prompt students to bring in the conversation any relevant experiences to support their answers. Thus far, the use of experiences from everyday life was completely absent from the conversation, and the teacher thought that this could help students to make progress in the conversation. Students immediately started describing experiences, evaluating at the same time their relevance with the phenomenon under study.

> 138. Dioni: This is the same with throwing a ball in a moving car – but it has to have on open ceiling. Because there are some cars that have no ceiling.
> 139. Teacher: ok. [...]
> 146. Erini: Be in an airplane.
> 147. Teacher: So, when you are in an airplane or a car, like Dioni said, and you throw the ball up in the air what is going to happen to the ball?
> 148. Dioni: It is going to fall behind you.
> 149. Teacher: ok, it will fall behind you.
> 150. Panayiotis: It is going to fall in your hand! [...]
> 153. Merriam: It is going to fall behind you.
> 154. Sabine: Does the car have a ceiling or not?
> 155. Teacher: Does it matter?
> 156. Sabine: Of course! If it has a ceiling, then the ball is going to hit the ceiling and then return to the point that it was thrown in the first place.
> 157. Teacher: You mean in your hand? [...]
> 161. Sabine: Yea. [it will fall] In your hand, because it is going to hit the car's ceiling and then fall back down.

As soon as the teacher prompted for related experiences a couple of students provided some, related with cars and airplanes traveling with people sitting inside them, suggesting their similarities with the phenomenon. Dioni (138) suggested that the phenomenon under study was similar to throwing a ball within a car, highlighting at the same time that for the two situations to be comparable, the car should not have a ceiling, possibly thinking that if the ball touches the ceiling then the phenomenon would have different set-up conditions. Erini (143) talked about the example of a flying airplane, and for the first time Panayiotis introduced the idea that the ball would fall back in the boy's hand.

The teacher's prompt (131) sparked a new dynamic in the conversation. Students not only offered experiences as a justification of their ideas, but they also attempted to evaluate the relevance of these experiences with the phenomenon under study. At the same time, when Sabina (156, 161) for instance described what would happen to the ball when you throw it from within a moving car, she becomes quite specific, talking about the ball's motion "...the ball is going to hit the ceiling and then return to the point that it was thrown in the first place" making some references to a mechanism that could provide a partial explanation of the phenomenon.

In terms of analogical reasoning (May et al., 2006), the same conversational data suggest that students can generate analogies (by identifying a target (the phenomenon under discussion) and a base case (the car or the airplane example) and their relation), and they can also validate and evaluate the relevance those analogies by identifying their key features and their limitations (i.e., the car has to be without a ceiling). In terms of argumentation (Louca & Hammer, submitted), students' statements were now accompanied by some grounds, although they were doing much more (i.e. evaluating the relevance of experiences) that the argumentation coding scheme cannot capture.

> 204. Panayiotis: If the car has a ceiling, then the ball will fall back in his hand, but if the car's ceiling is open, then the ball will fall back.
> 205. Teacher: And why is that?
> 206. Panayiotis: Because when the ball gets outside the car, then it becomes a separate object from the car

which moves forward, whereas the ball falls straight down, after the car moved forward.

When Panayiotis re-iterated the car idea and its relationship with the phenomenon under study, he proposed the idea of independent systems. In an open-ceiling car, if the ball gets outside the car, then the ball becomes independent from the car and acts as a different object, whereas when the ball is within the car it acts like one object with the car. In this contribution, Panayiotis talked about set-up conditions (if the car has a ceiling… and if the car does not have a ceiling, if the ball gets outside of the car … and if the ball does not get outside of the car), different entities that play different roles in the phenomenon (the ball and the car), about the properties of these entities (the ball becomes a different object as soon as it gets outside the car), and those entities' activities (the ball would stay at the same stop, whereas the car will keep moving forward). Despite the wrong application of the idea of independent systems, all these suggest more sophisticated student inquiry in terms of mechanistic reasoning that was not evident thus far.

Hidden assumptions

Despite that progress, students did not seem to move towards analyzing the "story" of the physical system into smaller conceptual entities (Russ et al., submitted), their characteristics and their behavior, which is required to make progress in terms developing a mechanistic explanation about how the phenomenon happens. Apart for talking about physical entities, it is also important that students address conceptual entities (such as velocity in this case) that play important roles in the phenomenon. We are not suggesting that student do not have any ideas or cannot conceptualize those conceptual entities. In fact, the discourse suggests that their ideas had two underlying "hidden assumptions" (Hammer, personal communication) concerning the ball's horizontal velocity, which prevented them for making any progress: they thought that the ball had either no horizontal velocity or the horizontal velocity became 0 after leaving the boy's hand. For instance, when Dioni (32) stated her idea, she indicated that "…when the ball leaves the boy's hand, and because the hallway is moving forward, the ball is going to fall behind." Whereas Panayioitis indicated that starting from the point that the ball is released from the boy's hand "… the ball's velocity is slowly decreasing."

Analyze the story into conceptual entities

The problem was that although these ideas underlie their contributions about what would happen in the phenomenon, students did not address them directly. The teacher decided to help students realize and evaluate those hidden assumptions by prompting them to talk about the different velocities during the ball's motion. He decided to provide students with a video of the phenomenon and have a discussion specifically about the ball's motion. This happened over the next two meetings and students had the opportunity to watch a video about the phenomenon, and talk about why the ball falls back into the boy's hand. After that, students could clearly distinguish between the two velocities (the horizontal and the vertical one), and could talk about the result of the combination of those two velocities.

1013. Panayiotis: So, when the ball moves like that..., there is one velocity like that [his left hand shows the upwards velocity's direction] and there is another velocity that moves like that [his right hand shows the direction of the velocity due to the hallway's motion]. When you put these two [velocities] together, then they form this shape [shows the oval trajectory of the ball with his hand].

1014. Teacher: ok. Let's take them one-by-one. What do we know about the vertical velocity?

1015. Costas: When the ball leaves the hand, that velocity starts decreasing, until one point where it will become zero. Then, the ball will start falling down and its velocity will start increasing.

1016. Teacher: What about the other velocity?

1017. Myriani: That velocity is steady, and is the same with the hallway's velocity.

By analyzing the video about the phenomenon under study, students broke down the ball's story into smaller conceptual entities, sketching their relationships. They had now a more analytic understanding of the phenomenon, being able to describe the phenomenon both in small conceptual entities (i.e., the horizontal (1013) and vertical velocity (1015)) and as a whole (the result of the two velocities) (1013). In terms of mechanistic reasoning, students were able to talk about entities, their properties and organization and activities of these entities that produce change in the phenomenon, showing more sophisticated abilities for mechanistic reasoning.

Discussion

This case study is a demonstration of nascent student inquiry in classroom settings. Although we do not claim that our analysis covers the complete spectrum of classroom-based student inquiry, findings from this and other studies (Feldman, 1994; Karmiloff-Smith, 1992; Koslowski, 1996; Louca & Hammer, submitted; Metz, 1995), contend these students come in the classroom "having" some abilities for i.e., argumentation, mechanistic reasoning, and analogical reasoning. With ambiguities regarding productive science inquiry, findings from this study reveal insights with respect to the challenge of defining what student inquiry can look like in the classroom and what teachers should expect to see, speaking to the debate regarding what "productive" inquiry entails, especially in early grades.

Students in this class were able to use a number of different components of the 3 elements of student inquiry – some more sophisticated than others. We do not suggest that these students are experts in scientific inquiry, but rather that they have the beginnings of abilities for scientific inquiry. At the same time, these students seem not to use and apply those abilities in a systematic way. The use of those abilities seemed to vary probably due to a number of factors, possibly including the teacher's specific prompts and the micro-context of the conversation – in many cases, when a student entered a new "mode" of conversation that consisted of more sophisticated reasoning than before, other students followed this new mode of sophisticated inquiry.

Still, if we were to make an assessment of the students' abilities for scientific discourse from those first 46 conversational turns, we would have a very different sense than from what followed. Of course, the students have not developed new abilities in the five minutes since the beginning of the conversation. Rather, they are applying different abilities from their repertoire showing some sophistication in those abilities.

All these suggest that the emphasis of instruction should be on identifying the beginnings of abilities for scientific inquiry in children, focusing on abilities that they already have and possibly use in different contexts. We feel that the literature over-emphasizes the need to actually teach students about how to use arguments, analogies or even mechanistic reasoning. For instance, most of research in classroom-based argumentation discourse (Erduran et al., 2004; Kuhn & Udell, 2003) provide some evidence that children have at least *the beginnings* of abilities regarding argumentation. Instead of seeing children as capable of learning i.e. how to use and evaluate arguments, use analogies, or develop scientific explanations, we suggest that science educators need to help students refine (not teach or develop) abilities for scientific inquiry that they already have. Since is seems to be a matter of ability activation in the appropriate context, by refining we mean helping students develop reliable access to those abilities for using them in the right context and time.

Acknowledgments

This study was supported by the Cyprus Research Foundation Awards #ΕΝΙΣΧ/0603/09 & ΕΝΙΣΧ/0505/44.

We thank Dr. Michalis P. Michaelides for his comments and suggestions during the manuscript preparation.

References

Ball, D. L., (1993). With an eye on the mathematical horizon: dilemmas of teaching elementary school. *The Elementary School Journal*, 93 (4): 373-397.

Blanchette, I., & Dunbar, K. (2001). Analogy use in naturalistic settings: The influence of audience, emotion, and goals. *Memory and Cognition*, 29(5): 730-735.

Bugallo Rodrıguez, A., & Jimenez-Aleixandre, M. P. (1996, August). Using Toulmin's argument pattern to analyze genetics questions. *Paper presented at the Third European Science Education Research Association (ESERA) Summer school*, Barcelona.

Clement, J. (1998). Expert novice similarities and instruction using analogies. *International Journal of Science Education*, 20, 1271– 1286.

Driver, R., Newton, P. & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, 84(3): 287-312.

Erduran, S., Simon, S. & Osborne J. (2004). TAPping into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. *Science Education*, 88(6): 915-933.

Feldman, D. H. (1994). *Beyond Universals in Cognitive Development.* Norwood, NJ: Ablex.

Gallas, K. (1995). *Talking their way into science: hearing children's questions and theories, responding with curricula.* NY: Teachers College Press.

Hammer, D. (1995). Student Inquiry in a physics class discussion. *Cognition and Instruction*, 13(3): 401-430.

Hawkins, D. (1974). *The Informed Vision: Essays on Learning and Human Nature*. New York: Agathon Press.

Karmiloff-Smith, A. (1992). *Beyond Modularity*. Cambridge, MA: MIT Press.

Koslowski, B. (1996). T*heory and evidence: the development of scientific reasoning*. Cambridge, Mass: MIT Press.

Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review*, 96(4): 674-689.

Kuhn, D. (1993). Science as argument: Implications for teaching and learning scientific thinking. *Science Education*, 77(3): 319-337

Kuhn, D. & Udell, W. (2003). The development of argument skills. *Child Development*, 74(5): 1245-1260.

Louca, L. & Hammer, D. Student Nascent Abilities for Scientific Argumentation: The Case of a 5th-6th-Grade Conversation about a Dropped Pendulum. Paper submitted to *Cognition & Instruction*.

Louca, L., Elby, A., Hammer, D., & Kagey, T. (2004). Epistemological resources: Applying a new epistemological framework to science instruction. *Educational Psychologist*, 39(1): 57-68.

May, D. B., Hammer, D., & Roy, P. (2006). Children's analogical reasoning in a 3rd-grade science discussion. *Science Education*, 90(2): 316-330.

Metz, K. (1995). Reassessment of developmental constraints on children's science instruction. *Review of Educational Research*, 65(2): 93-127.

Minstrell, J. A. & Van Zee, E. H. (2000). *Inquiring into inquiry: learning and teaching in science.* Washington, D.C., AAAS.

National Research Council [NRC]. (1996). *National Science Education Standards*. Washington DC: National Academy Press.

Osborne, J., Erduran, S. & Simon, S. (2004). Enhancing the quality of argumentation in school science. *J. of Research in Science Teaching*, 41(10): 994-1020.

Russ, R., Scherr, E., R., Hammer, D., & Mikeska, J. Recognizing mechanistic reasoning in scientific inquiry. Paper submitted in *Science Education*.

Samarapungavan, A. (1992). Children's judgments in theory choice tasks: Scientific rationality in childhood. *Cognition*, 45: 1-32.

Schauble, L. (1996). The development of scientific reasoning in knowledge-rich contexts. *Developmental Psychology*, 32 (1): 102-119.

Can Tutored Problem Solving Benefit From Faded Worked-Out Examples?

Rolf Schwonke (rolf.schwonke@psychologie.uni-freiburg.de)
Department of Psychology, University of Freiburg
Engelbergerstr. 41, D-79085 Freiburg, Germany

Jörg Wittwer (wittwer@ipn.uni-kiel.de)
Leibniz Institute for Science Education at the University of Kiel
Olshausenstr. 62, D-24098 Kiel, Germany

Vincent Aleven (aleven@cs.cmu.edu)
Human-Computer Interaction Institute, School of Computer Science, Carnegie Mellon University
5000 Forbes Ave, Pittsburgh, PA 15213 USA

Ron Salden (rons@cs.cmu.edu)
Human-Computer Interaction Institute, School of Computer Science, Carnegie Mellon University
5000 Forbes Ave, Pittsburgh, PA 15213 USA

Carmen Krieg (krieg@ psychologie.uni-freiburg.de)
Department of Psychology, University of Freiburg
Engelbergerstr. 41, D-79085 Freiburg, Germany

Alexander Renkl (renkl@psychologie.uni-freiburg.de)
Department of Psychology, University of Freiburg
Engelbergerstr. 41, D-79085 Freiburg, Germany

Abstract

Although problem solving supported by Cognitive Tutors has been shown to be successful in fostering initial acquisition of cognitive skills, this approach does not seem to be optimal with respect to focusing the learner on the domain principles to be learned. In order to foster a deep understanding of domain principles, we developed a Cognitive Tutor that contained, on the basis of the theoretical rational of example-based learning, faded worked-out examples. We conducted two experiments in which we compared the example-enriched Cognitive Tutor with a standard Cognitive Tutor. In Experiment 1, we found no significant differences in the effectiveness of the two tutor versions. However, the example-enriched Cogntive Tutor was more efficient (i.e., students needed less learning time). A problem that was observed is that students had great problems in appropriately using the example-enriched tutor. In Experiment 2, we, therefore, provided students with additional instructions on how to use the tutor. Results showed that students in fact acquired a deeper conceptual understanding when they worked with the example-enriched tutor and they needed less learning time than in the standard tutor. The results are suggestive of ways in which instructional models of problem-solving and example-based learning can be fruitfully combined.

Introduction

ognitive Tutors® (a trademark of Carnegie Learning, Inc.) an intelligent tutoring system – have been proven to be :ry effective in supporting students' learning in a variety of domains, including mathematics, computer programming, and genetics (for an overview, see Anderson, Corbett, Koedinger, & Pelletier, 1995; Koedinger & Corbett, 2006). On the basis of an online assessment of the student's learning, they provide individualized support for guided learning by doing. Specifically, the tutor selects appropriate problems, gives just-in-time feedback, and presents hints. Despite their effectiveness, a shortcoming of these tutors is that they primarily focus on students' problem solving and do not necessarily support a conceptual understanding about the domain to be learned.

Previous research has attempted to address this limitation by introducing self-explanation prompts to the students who work with the tutor. The prompts require students to provide an explanation for each of their solution steps, by making an explicit reference to the underlying principle. Empirical findings show that this instructional approach makes the cognitive tutor indeed more effective (Aleven & Koedinger, 2002). However, from a cognitive load perspective (e.g., Sweller, van Merriënboer, & Paas, 1998), it might be objected that the technique is nevertheless suboptimal because the induction of self-explanation activities in addition to problem solving places fairly high demands on students' limited cognitive capacity, particularly in the early stages of skill acquisition. Therefore, the tutor's effectiveness might be further improved by reducing cognitive load (e.g., van Merriënboer, Kirschner, & Kester, 2003), allowing students to spend more attentional capacity to engage in meaningful learning activities.

Against this background, it might be sensible to provide students with worked-out examples. The instructional model of example-based learning developed by Renkl and Atkinson (in press) suggests that learners gain a deep understanding of a skill domain when they receive worked-out examples at the beginning of cognitive skill acquisition. A worked-out example consists of a problem formulation, solution steps, and the final solution. When studying worked-out examples instead of solving problems, the learners are freed from performance demands and they can concentrate on achieving a deep understanding. Assuring that learners have a basic understanding before they start to solve problems should help them to deal with the problem-solving demands by referring to already acquired principles, which should prevent them from using only shallow strategies, such as means-end analysis or copy-and-adapt strategies (e.g., using the solution of a previously solved problem that is adapted with respect to the specific numbers). The use of principles enables learners to deepen their knowledge by applying the principles to new problems and, in addition, will cause them to notice gaps in their principle-related understanding when they reach an impasse (cf. VanLehn et al., 2005).

There is ample empirical evidence showing that learning from worked-out examples leads to superior learning outcomes as compared to the traditional method of problem solving (for an overview, see Atkinson, Derry, Renkl, & Wortham, 2000). However, it is important to note that studying worked-out examples loses its effectiveness with increasing expertise. In later stages of skill acquisition, the skillful execution of problem-solving activities plays a more important role because emphasis is put on increasing speed and accuracy of performance (Renkl & Atkinson, 2003). For example, Kalyuga, Chandler, Tuovinen, and Sweller (2001) found that learning from worked-out examples was superior in the initial phase of cognitive skill acquisition. However, when learners already had a basic understanding of the domain, solving problems proved to be more effective than studying examples (*expertise reversal effect*; Kalyuga, Ayres, Chandler, & Sweller, 2003). Therefore, Renkl and Atkinson (2003) proposed a fading procedure in which problem-solving elements are successively integrated into example study until the learners are expected to solve problems on their own. First, a complete example is presented. Second, a structurally identical incomplete example is provided in which one single step is omitted. In the subsequent isomorphic examples, the number of blanks is increased step by step until just the problem formulation is left, that is, a problem to be solved. Hence, by gradually increasing problem-solving demands, the learners should retain sufficient cognitive capacity to successfully cope with these demands and, thereby, to focus on domain principles and on gaining understanding. In a number of experiments, Renkl and colleagues provided empirical evidence for the effectiveness of a smooth transition from example study to problem solving (e.g., Atkinson, Renkl, & Merrill, 2003; Renkl, Atkinson, & Große, 2004).

Against this background, we expected that a Cognitive Tutor that not only prompts students to engage in self-explaining but also provides them with gradually faded worked-out examples should foster students' learning, particularly with respect to their conceptual understanding. In addition, the empirical results on the *worked-example effect* (positive effect of studying examples) also leads to the expectation that the learners need less study time (cf. e.g., Sweller & Cooper, 1985) when they use an example-enriched Cognitive Tutor as compared to the standard version. Accordingly, we hypothesized that a combination of example study and tutored problem solving would be more effective and more efficient than tutored problem solving alone. To test this hypothesis, we modified a Cognitive Tutor to achieve a state-of-the-art implementation of example-based learning with a gradual transition into problem solving .

In this article, we present two experiments in which we investigated the question whether an 'example-enriched' Cognitive Tutor would lead to superior learning when compared with a standard version of a Cognitive Tutor. For this purpose, we used the Cognitive Tutor Geometry. Students were asked to work on geometry problems that required them to apply different geometry principles.

Experiment 1

Method

Sample and Design

Fifty students from a German high school, 22 eighth-grade students and 28 ninth-grade students, participated in the experiment (average age: 14.3 years; 22 female, 28 male). The students were randomly assigned to one of the two experimental conditions. In the experimental condition (*example condition*; n = 25), students worked with a Cognitive Tutor that presented faded worked-out examples. In the control condition (*problem condition*; n = 25), the students worked with a standard version of the tutor in which students received no faded worked-out examples.

Learning Environment – The Cognitive Tutor

The students used two versions of the Geometry Cognitive Tutor, which differed by a single factor: whether or not worked-out examples were presented. In both versions, self-explanation prompts were employed (Aleven & Koedinger, 2002). In addition, information such as text and diagrams was presented in a single worksheet (i.e., in an integrated format). For the purpose of comparing worked-out examples with problem solving, the integrated format of the tutor was important because example-based learning might be more effective than problem solving only when a 'split source format' is avoided (i.e., the advantages of examples may not materialize when related information such as text and schematics or diagrams is presented separately, cf. Tarmizi & Sweller, 1988). Thus, this Cognitive Tutor version allowed a fair and a state-of-the-art implementation of worked-out examples. The Cognitive Tutor itself is a state-of-the-art intelligent tutoring system, in regular use in about 350 schools across the United States as part of the regular geometry curriculum.

In general, Cognitive Tutors employ two algorithms to support learners. These algorithms are called 'model tracing' and 'knowledge tracing'.

Model Tracing In order to provide appropriate just-in-time feedback and hints, the Cognitive Tutor relies on a computational model that represents the domain-specific knowledge that is necessary to solve problems. The model may also include problem-solving knowledge and skills that are typical for novices (Koedinger & Corbett, 2006). In addition, the model may include incorrect problem-solving approaches that are common for novices. The problem-solving skills (so-called *knowledge components*) are represented as production rules (i.e., if-then rules that link internal goals or external cues with new goals or actions). All user interactions with the tutor are interpreted relative to this model. Student answers that correspond to production rules are marked as correct. If an answer relates to a rule that represents an incorrect strategy, an error feedback message is presented to the student. Answers that do not correspond to any production rule are marked as incorrect. At any point in time, the student can request a hint from the tutor. The tutor will use its cognitive model to decide what a good next problem-solving step will be, and it will present hints using text templates attached to the relevant production rule(s).

Knowledge Tracing The full-scale Cognitive Tutors also implement a cognitive mastery learning criterion (Corbett & Anderson, 1995) but this capability was turned off during the experiment to keep the number and order of problems constant across participants.

Learning material In total, students were asked to work on seven problems. The first three problems required the application of only one geometry principle. In order to solve the last four more complex problems, it was necessary to apply these geometry principles in combination. In the *problem condition*, solving a problem required students (a) to enter a numerical value (such as the measure of an angle) in an entry field that was embedded in a graphical representation of the problem (in a worksheet), and (b) to justify each given numerical answer. This justification could be entered either by typing the name of a relevant principle into a text entry field (next to the numerical value entry field), or by selecting a principle from a glossary that contained a list of all principles used in the unit (i.e., explanation by reference). The combination of entering a value and providing a justification is called a learning event. For example, given the measure of an angle $m\angle ABC = 145°$, a student may be asked to figure out the measure of the supplementary angle $\angle ABD$. The correct entry would be $m\angle ABD = 35°$, because $m\angle ABD = 180° - m\angle ABC$. After entering the value (or an artihmetic expression, leaving the computation to the tutor) the student has to justify (i.e., to explain) this numerical answer in a second step. In this case, a valid explanation would be 'supplementary angles'.

In the *example condition*, students were asked to study a sequence of worked-out examples that corresponded exactly to the problems that students in the problem condition were asked to solve. A worked-out example provided the students with the numerical value (to be figured out in the problem condition) together with the necessary solution steps. The examples were gradually faded out according to the fading scheme displayed in Table 1. The table shows that the application of the principle in each of the first three problems was illustrated by a worked-out example. Also, worked-out examples were used for the fourth problem that required the application of the three principles in combination. In the subsequent problems, however, each of the principles was gradually faded out until just the problem formulation was left (problem 7).

Table 1: The sequencing of problems and fading of worked-out steps.

	Examples			Problem solving		
	Principles					
Problems	P1	P2	P3	P1	P2	P3
P1	W			S		
P2		W			S	
P3			W			S
P4	W	W	W	S	S	S
P5	W	W	S	S	S	S
P6	W	S	S	S	S	S
P7	S	S	S	S	S	S

Note. W stands for worked-out examples and *S* for problem solving.

In order to hold the self-explanation activities across the two experimental conditions constant, students in both versions of the Cognitive Tutor were asked to provide justifications for all solution steps and worked-out steps. Hence, when working on the first four problems, students in the example condition had to enter justifications for the numerical answers that were provided in the worked-out examples by the tutor. Like in the problem condition, the justifications could be typed in or selected from the glossary. From problem five to problem seven, problem-solving demands in the example condition were gradually increased. Hence, students were required not only to give justifications but also to solve the problem on their own.

Instruments

Pretest A short pretest on circles geometry containing 4 problems examined the topic-specific prior knowledge of the students. The maximum score to be obtained in the pretest was 12 points (3 points for each problem that was solved correctly).

Post-test The post-test that measured students' learning consisted of 13 questions. Two questions required the students to solve problems that were isomorphic to the problems previously presented by the Cognitive Tutor (near transfer items). In addition, 2 questions were devised to test students' ability to apply their knowledge about the geometry principles to new geometry problems (far transfer items). As both transfer scores correlated with .69 (p < .001), we aggregated them to an overall transfer score. Finally, 9 questions assessed the conceptual understanding that students acquired with the help of the tutor. Students were asked to explain the geometry principles (a maximum of 22 points could be obtained).

Procedure

The experimental sessions lasted, on average, 90 minutes and were divided into three parts: pretest and introduction, tutoring, and post-test. In the pretest and introduction part, students were asked to complete the pretest measuring their prior knowledge. Afterwards, they read an instructional text that provided them with information about the rules and principles that were later addressed by the Cognitive Tutor. In addition, they received a brief introduction on how to use the tutor. In the tutoring part, students worked either with the standard Cognitive Tutor Geometry or with the example-enriched version. In the post-test part, all students answered the transfer questions and the questions assessing their conceptual knowledge.

Results

First, we analyzed students' prior knowledge in order to assure that the experimental conditions did not differ with respect to this important learning prerequisite. There were no significant differences between the experimental groups, $t(48) = -0.75$, $p > .05$, $d = -0.21$. The low test scores obtained in the pretest (cf. Table 2) indicate that students in both experimental conditions were in fact in the initial phase of skill acquisition.

In a second step, we analyzed whether learning with a combination of example study and tutored problem solving was better than tutored problem solving alone. We found, however, no significant differences in students' learning outcomes, neither for conceptual knowledge, $t(48) = -0.11$, $p > .05$, $d < 0.01$, nor for transfer, $t(48) = 0.22$, $p > .05$ $d = 0.08$. Hence, both versions of the cognitive tutor were similarly effective.

In a last step, we examined how efficiently students worked with the tutor. For this purpose, we compared the time that students spent working on the problems or examples provided by the tutor. The analysis revealed significant time on task differences, $t(48) = -3.11$, $p < .001$ (one-tailed), $d = -0.88$. Students in the problem condition spent more time for learning than students in the example condition (cf. Table 2).

In order to quantify the differences in efficiency, we adopted the efficiency measure developed by Paas and colleagues (Paas, Tuovinen, Tabbers, & Van Gerven, 2003; Paas & Van Merriënboer, 1993). This measure relates performance in terms of learning outcomes to mental effort in terms of cognitive load as measured, for example, by questionnaires.

Table 2: Means and standard deviations of pretest and post-test scores, learning time, and learning efficiency for the experimental conditions in Experiment 1.

	Example		Problem	
Variable	*M*	*SD*	*M*	*SD*
Pretest[a]	.13	.11	.15	.11
Learning time[b]	30.0	6.56	35.4	5.72
Conceptual knowledge[a]	.54	.21	.54	.21
Transfer[a]	.12	.12	.11	.13
Conceptual knowledge acquisition efficiency[c]	0.28	1.13	-0.28	1.13
Transfer acquisition efficiency[c]	0.31	1.28	-0.31	1.11

Note. [b]Solution probability. [a]Learning time in minutes. [c]Efficiency = $(z_{\text{Post-test}}\text{-}z_{\text{Learning time}})$/SQRT(2).

More specifically, the efficiency score equals to the difference of z-scores of mean performance and effort measures (i.e., $z_{\text{performance}}\text{-}z_{\text{effort}}$) divided by the square root of two. For our purposes, we related performance in terms of the acquisition of conceptual knowledge and of transferable knowledge respectively to effort in terms of time on task (i.e., the time spent working on the problems). This relationship is depicted in the following formula:

$$\text{learning efficiency} = (z_{\text{Post-test}}\text{-}z_{\text{TimeOnTask}})/\sqrt{2}$$

Applying this efficiency formula to our data, we found significant differences between the experimental conditions for both the efficiency of conceptual knowledge acquisition, $t(48) = 1.73$, $p < .05$ (one-tailed), $d = 0.50$, and the efficiency of the acquisition of transferable knowledge, $t(48) = 1.82$, $p < .05$ (one-tailed), $d = 0.52$, which both represent medium sized effects (see Table 2).

Discussion

Both tutored problem-solving and learning with a smooth transition from worked-out examples to problem solving led to comparable levels of conceptual and procedural knowledge (in terms of near and far transfer). However, about the same learning outcomes were achieved in shorter learning times in the example-enriched Cognitive Tutor. Accordingly, the efficiency of learning was superior in this latter learning condition.

Contrary to our expectation, there was no difference in the effectiveness of the two conditions. The lack of difference might be explained by the fact that even the standard version of the Cognitive Tutor is very supportive. Thus, there might not have been much room for improvement (cf. Koedinger & Aleven, in press). Both versions of the tutor provided corrective feedback on errors and induced students to engage in self-explaining activities. Both versions provided on-demand hints (even if students did not use them very frequently). However, students in both experimental conditions achieved relatively low post-test scores making this explanation not very likely.

Informal observations and analyses of the log-file data suggested that students had difficulties in working with the Cognitive Tutor. These problems were clearly more pronounced in the example condition than in the problem condition. Although students received instructions on how to use the tutor, students in the example condition in particular had trouble in understanding the purpose of the worked-out examples. One severe and persistent misunderstanding related, for example, to the justifications that students had to give for a solution step. In the majority of cases, the students assumed that they had to enter the justification 'given' (because the numerical value had been provided by the tutor) instead of the mathematical principle relevant to the task at hand.

In order to examine whether students' problems in using the Cognitive Tutor, especially working on the worked-out geometry tasks, diminished possible differences between the two experimental conditions, we conducted another experiment. In this experiment, we gave the students more detailed and specific instructions on how to use the tutor.

Experiment 2

In the second experiment, we provided the students in both experimental conditions with more specific instructions prior to using the tutor. In addition, when students worked on the two warm-up examples provided by the tutor, they received, in case of problems in understanding, scaffolding from the experimenter.

Method

Experiment 2 was identical to Experiment 1 with respect to the experimental set up, the learning environment, and the instruments (e.g., pretest and post-test). Yet, the experiment was different with respect to the level of detail of the instruction and scaffolding provided in advance, as explained before. In addition, students in experiment 2 participated in individual sessions in the study, whereas Experiment 1 took place in a group session format.

Sample and Design

In Experiment 2, 16 ninth-grade students and 14 tenth-grade students of a German high school (average age: 15.7 years; 17 female, 13 male) took part. As in experiment 1, one half of the students were assigned to the example condition (n = 15) and the other half to the problem condition (n = 15). The procedure was similar to the procedure of Experiment 1.

Results and Discussion

In a first step, we analyzed students' prior knowledge. Again, there were no significant differences between the experimental conditions, $t(28) = 0.27$, $p > .05$, $d = 0.1$ (cf. Table 3). We then examined whether students in the example condition benefited more from the example-enriched Cognitive Tutor than students in the problem-solving condition. With regard to students' conceptual understanding, we indeed found an advantage of the example condition over the problem condition, $t(28) = 1.85$, $p < .05$ (one-tailed), d = 0.73 (medium sized effect). However, again there were no significant differences in students' transfer knowledge, $t(28) = -0.61$, $p > .05$, $d = -0.21$.

Table 3: Means and standard deviations of pretest and post-test scores, learning time, and learning efficiency for the experimental conditions in Experiment 2.

	Example		Problem	
Variable	*M*	*SD*	*M*	*SD*
Pretest[a]	.14	.16	.13.	.12
Learning time[b]	30.0	6.48	39.2	9,31
Conceptual knowledge.[a]	.61	.14	.50	.16
Transfer[a]	.19	.22	.24	.25
Conceptual knowledge acquisition efficiency[b]	0.58	0.90	-0.58	0.94
Transfer acquisition efficiency[c]	0.31	1.28	-0.31	1.11

Note. [a]Solution probability. [b]Learning time in minutes. [c]Efficiency = $(z_{Post\text{-}test} - z_{Learning\ time})/SQRT(2)$.

In a last step, we computed the efficiency of students' using the Cognitive Tutor. The differences found in experiment 1 could be replicated. This time, the differences were even more pronounced. Again, students in the problem condition spent more time working with the tutor than students in the example condition, $t(28) = -3.14$, $p < .001$ (one-tailed), d = -1.17 (large sized effect). Hence, when we related performance in terms of the acquisition of conceptual knowledge to the effort in terms of time on task, we obtained a large effect, $t(28) = 3.48$, $p < .001$ (one-tailed), d = 1.26. The efficiency of transferable knowledge acquisiton failed to reach the level of significance, $t(28) = 1.44$, $p = .08$ (one-tailed), $d = 0.52$.

General Discussion

In the two experiments, we compared a standard Cognitive Tutor with an example-enriched Cognitive Tutor. Both versions of the tutor offered corrective feedback and self-explanation prompts. The present research extends previous research in important ways. First, we found evidence, that a state-of-the-art implementation of a faded worked-out steps procedure can lead to a deeper conceptual understanding than intelligently tutored problem solving. In contrast to previous studies that usually compared example-based learning with largely unsupported problem solving (cf. Atkinson et al., 2000), in the present study, the Cognitive Tutor (as used in the control condition) provided students with a substantial amount of support by hints and corrective feedback. Therefore, it was comparatively difficult to find incremental effects on students' performance by adding worked-examples (cf. the results of McLaren, Lim, Gagnon, Yaron, & Koedinger, 2006). This probably accounts for the fact that we did not find differences in procedural knowledge (as did previous studies). Nevertheless, the results on learning time and hence on efficiency clearly show that example-based learning can be less time consuming without a loss or even a gain in conceptual knowledge. This result is also of practical relevance as it offers an alternative for the allocation of the precious resource 'learning time'. Second, and on a more general level, the present research is an example of how different instructional approaches (i.e. tutored problem solving and worked-out examples) can be productively combined to the benefits of learners.

Finally, some remarks on potentially fruitful future research directions It can be speculated that faded examples could be even more beneficial to learning if they take into account the individual prerequisites of the students. It is plausible to assume that students might differ considerably in the speed and accuracy with which they learn domain principles. Therefore, it should be sensible to adapt the speed of fading worked-out steps to the students' individual learning progress. We will conduct an experiment in which we examine the surplus value of a fading procedure that dovetails with the students' specific needs.

Acknowledgments

This research was supported by the National Science Foundation (NSF).

References

Aleven, V., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based cognitive tutor. *Cognitive Science, 26*, 147-179.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of the Learning Sciences, 4*, 167-207.

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. W. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research, 70*, 181-214.

Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning from studying examples to solving problems: Combining fading with prompting fosters learning. *Journal of Educational Psychology, 95*, 774-783.

Corbett, A. T., & Anderson, J. R. (1995). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction, 4*, 253-278.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38*, 23-31.

Kalyuga, S., Chandler, P., Tuovinen, J., & Sweller, J. (2001). When problem solving is superior to studying worked examples. *Journal of Educational Psychology, 93*, 579-588.

Koedinger, K. R. & Aleven, V. (in press). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review.*

Koedinger, K. R., & Corbett, A. T. (2006). Cognitive tutors: Technology bringing learning sciences to the classroom. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences*. New York, NY: Cambridge University Press.

McLaren, B. M., Lim, S., Gagnon, F., Yaron, D., & Koedinger, K. R. (2006, June, 26-30). *Studying the effects of personalized language and worked examples in the context of a web-based intelligent tutor.* Paper presented at the 8th International Conference on Intelligent Tutoring Systems (ITS-2006), Jhongli, Taiwan.

Renkl, A. & Atkinson, R. K. (2003). Structuring the transition from example study to problem solving in cognitive skills acquisition: A cognitive load perspective. *Educational Psychologist, 38*, 15-22.

Renkl, A., & Atkinson, R. K. (in press). Cognitive skill acquisition: Ordering instructional events in example-based learning. F. E. Ritter, J. Nerb, E. Lehtinen, T. O'Shea (Eds.), *In order to learn: How ordering effect in machine learning illuminate human learning and vice versa*. Oxford, UK: Oxford University Press.

Renkl, A., Atkinson, R. K., & Große, C. S. (2004) How fading worked solution steps works – A cognitive load perspective. *Instructional Science, 32*, 59-82.

Paas, F., & van Merriënboer, J. J. G. (1993). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human Factors, 35*, 737-743.

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*, 63-71.

Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition & Instruction, 2*, 59-89.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251-296.

Tarmizi, R. A., & Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology, 80*, 424-436.

van Merriënboer, J. J. G., Kirschner, P. A., & Kester, L. (2003). Taking the load off a learner's mind: Instructional design for complex learning. *Educational Psychologist, 38*, 5-13.

VanLehn, K., Lynch, C., Schulze, K., Shapiro, J., Shelby, R., Taylor, L. et al. (2005). The Andes physics tutoring system: Lessons learned. *International Journal of Artificial Intelligence in Education, 15*, 147-204.

Effects of Feedback on the Strategic Competence of Gifted Children

Ageliki Foustana (afousta@primedu.uoa.gr)
Department of Primary Education, University of Athens
Harilaou Trikoupi 24, 106 79 Athens, Greece

Koen Luwel (Koen.Luwel@ped.kuleuven.be)
Department of Educational Sciences, University of Leuven
Vesaliusstraat 2, B-3000 Leuven, Belgium

Lieven Verschaffel (Lieven.Verschaffel@ped.kuleuven.be)
Department of Educational Sciences, University of Leuven
Vesaliusstraat 2, B-3000 Leuven, Belgium

Abstract

The effect of two types of feedback on the strategic competence of high and average intelligent children was examined in the context of a numerosity judgment task. We used a pretest-intervention-posttest design in which children's strategic competence in the pre- and the posttest session was assessed by means of the choice/no-choice method (Siegler & Lemaire, 1997). The intervention session involved the administration of solely a choice condition to two different feedback groups: half of the participants in each intelligence group received strategy-related feedback and the other half got outcome-related feedback. Results show differential effects of feedback type and intelligence level on several aspects of strategic competence.

Theoretical Background

The last 20 years have witnessed a great progress in the research on strategy choice and strategy use in many domains of human cognition (Siegler, 1996, 2005). This has resulted in new theoretical insights and formal (computer) models regarding the (development of) choice and use of strategies, in important methodological innovations (such as the microgenetic method and the choice/no-choice paradigm), and in educational applications aimed at supporting a varied and flexible use of strategies (see Torbeyns et al., 2004, for a critical review).

A powerful theoretical framework to analyse individuals' strategic competence has been proposed by Lemaire and Siegler (1995). In this framework, a distinction is made between four dimensions of strategic competence: (a) the strategic *repertoire* (i.e., which strategies an individual uses to solve a specific task), (b) the relative *frequency* of strategy use (i.e., how often each of the different strategies are applied), (c) the *efficiency* of strategy execution (i.e., how fast and accurate each strategy is executed), and (d) the *adaptiveness* of strategy choices (i.e., the extent to which strategy choices are calibrated towards problem characteristics as well as towards the one's own strategy efficiency). According to Lemaire and Siegler, improvements in overall task performance can be the result of changes in any of these dimensions.

The different aspects of strategic competence can be assessed by means of the choice/no-choice method devised by Siegler and Lemaire (1997). This method involves testing each participant under two types of conditions: (a) a choice condition in which participants can freely choose which strategy to use, and (b) a number of no-choice conditions in which participants are required to use one specific strategy on all problems. In principle, the number of no-choice conditions equals the number of strategies occurring in the choice condition. Data from the choice condition provide information about participants' strategic repertoire as well as about their frequency of strategy use. Since all participants are required to apply a given strategy on all items of the task in each no-choice condition, confounds between strategy selection and execution are excluded. As a result, data from the no-choice conditions can offer unbiased measures of strategy efficiency. Finally, the adaptiveness of strategy choices can be assessed by comparing the performance in the choice and no-choice conditions.

Although Lemaire and Siegler's (1995) framework in combination with the choice/no-choice method was originally used to study the development of strategy performance by comparing different age groups on the four parameters of strategic competence (e.g., Siegler & Lemaire, 1997; Luwel, Lemaire, & Verschaffel, 2005), it has recently also been used to examine the effect of situational variables on strategic performance by comparing different experimental conditions on these four dimensions (e.g., Imbo & Vandierendonck, in press,), or to study differences in strategic competence of different mathematical ability groups (e.g., Torbeyns, Verschaffel, & Ghesquière, 2004). Although intelligence has proven to play a major role in many cognitive tasks (Hong, 1999), this theoretical framework has, to the best of our knowledge, never been used to assess the contribution of intelligence to the different aspects of strategic performance.

Of course, some studies have already examined the effect of intelligence on strategic competence, but all these studies focussed on only one or two parameters of strategic competence and never on all four of them. As far as the repertoire of strategies is concerned, previous studies have shown that gifted children seem to use qualitatively different strategies to solve problems than did average intelligent children (e.g., Larkin, et al, 1980; Priest, & Lindsay, 1992), whereas other researchers found that the strategy repertoire is similar for gifted children and their average peers (e.g., Bouffard-Bouchard, Parent & Larivee, 1993; Gaultney, Bjorklund, & Goldstein, 1996).

Studies that examined the efficiency of strategy execution suggested that gifted children show more efficient strategy processes compared to children of a lower intellectual ability (e.g., Geary & Brown, 1991; Saccuzzo, Johnson, & Guertin, 1994). The speed and the accuracy of the responses were greater for the gifted groups.

With respect to strategy selection, it has been reported that the source of gifted children's generally superior cognitive performance is in their more frequent and more adaptive use of particular strategies and in the subsequent generalization of these strategies to new tasks (e.g., Ippel & Beem, 1987; Scruggs & Cohn, 1983; Wong, 1982). In sum, there are individual differences in strategy choices when solving problems. The nature and quality of these strategy shifts seems to be aptitude-dependent, meaning that individuals who are particularly adept at making optimal strategy selections tend to have higher (fluid) intelligence scores.

As mentioned earlier, previous research on the contribution of intelligence to strategic performance only lead to partial conclusions due to the limited number of dimensions of strategic performance addressed in these studies. Moreover, they yielded contradictory findings with respect to some of these dimensions. With the present study we wanted to investigate the role of intelligence in strategic competence by contrasting a group of gifted children with a group of average intelligent children on all four parameters of strategic performance. Moreover, we wanted to examine the extent to which different types of feedback would lead to improvements in one or more of these parameters (Kluger & De Nisi, 1996) and to test the assumption whether gifted children would benefit less from this feedback than average intelligent children (Rohwer, 1973).

The Present Study

In the present study, participants were asked to judge different numerosities of green blocks that were presented in a 7 x 7 grid. This task allows for two strategies: (a) an addition strategy through which the given quantity of blocks is divided into a number of subgroups and the judged numerosities of the different subgroups are added, and (b) a smart subtraction strategy in which the number of empty squares is subtracted from the total number of squares in the grid.

The experiment consisted of three sessions: a pretest, an intervention and a posttest session. In the pre- and the posttest session, the same set of experimental trials was administered under three different conditions: one choice condition and two no-choice conditions. In the choice condition participants were allowed to choose freely between the two strategies (addition or subtraction) on all trials of the task. In the no-choice/addition condition participants were required to determine all the numerosities by means of the addition strategy, whereas in the no-choice/subtraction condition only the use of the subtraction strategy was allowed.

The intervention only involved a choice condition in which half of the participants in each intelligence group received outcome feedback (OFB), in which they were informed about the accuracy of their numerosity judgment in each trial, while the other half received strategy feedback (SFB), with information about the appropriateness of their strategy choice in each trial.

The predictions we formulated concern intelligence-related differences and the effect of feedback in each of the four strategic dimensions. First, more gifted than average children will use the subtraction strategy in the pretest and the frequency of the subtraction use will increase with intelligence in the pretest. Second, there will be an intelligence-related increase in the efficiency of both strategies. Third, gifted children will select their strategies more adaptively than average children. These differences are assumed to decrease in the course of the study, since it is expected that average will benefit more than gifted from the intervention. Finally, the effects of strategy feedback are expected to be more beneficial than those of the outcome feedback.

Method

Participants

Participants were 40 intellectually gifted and 40 average intelligent first-grade students from several private and public secondary schools in the county of Attica (Greece). The mean chronological age was 12.54 yrs., ranging from 11.41 to 13.67 yrs. (SD = 0.36). In each group boys and girls were almost equally represented. The mean WISC- III full scale IQ of the gifted sample was 128.67 (range: 123-145, SD = 5.98), whereas that of the average intelligent sample was 103.57 (range: 90-110, SD = 6.88).

Materials

The numerosity judgment task was presented to the participants using an Acer personal computer and a 17-inch monitor with a resolution set to 1024 X 768 pixels. Stimuli were square grids consisting of 7 X 7 little square units that were intersected by red lines. The outline of the grid was visible and colored red. Each square unit in the grid had a size of 1 X 1 cm. These squares units could either be "on" (i.e., being filled with a green colored block) or "off" (remaining empty, i.e., having the same black color as the background of the whole of the screen). In all experimental sessions (pretest, intervention, posttest) and conditions (choice and no-choice conditions), participants ran 26 trials whereby all numerosities of blocks between 20 and 45 were presented. For each participant, the sequence of the stimuli as well as the placement of the blocks in the grid was randomized by the computer. We chose the relatively small 7 X 7 grid to ensure that all participants could solve all trials relatively easily by solely using the addition or the subtraction strategy (Luwel et al., 2005).

Design

As explained above, each participant was examined in three different sessions: a pretest, an intervention and a posttest session. The pretest and the posttest sessions consisted of three conditions: one choice condition and two no-choice conditions. The presentation order of the different conditions in both test sessions was counterbalanced across participants with the important restriction that the choice condition was always presented first, so that strategy choices could not be affected by recency effects. After the choice condition half of the participants in each intelligence and feedback group were enrolled in the no-choice/addition condition followed by the no-choice/subtraction condition, whereas the other half went through both no-conditions in the opposite order. At posttest participants received both no-choice conditions in the opposite order as in the pretest.

Procedure

All participants were tested individually and were seated at about 50 cm from the computer screen. Before the start of the actual experiment they were given five example trials that were representative for the whole continuum of numerosities in the grid (i.e., 7, 15, 25, 40, 46). Participants were instructed to determine each of the numerosities as accurately and fast as possible. After each example trial, participants were asked to explain briefly how they had handled the task.

At the beginning of the choice condition from both the pretest and the posttest, participants were told that they had to determine each numerosity by using either the addition or the subtraction strategy. For each trial, participants were instructed to indicate with their finger the units (i.e., green blocks/empty squares) they were counting. This instruction allowed the experimenter to determine whether participants were applying the addition strategy (i.e., when pointing to the green blocks) or the subtraction strategy (i.e., when pointing to the empty squares).

In the two no-choice conditions from the pretest and the posttest participants were told that they had to determine all numerosities of blocks by using only one strategy, either the addition strategy (no-choice/addition) or the subtraction strategy (no-choice/subtraction). Participants were again asked to indicate the blocks/empty squares that they were counting in order to guarantee that participants always used the required strategy.

The stimulus remained on the screen until the participants had made their numerosity judgment. They were asked to verbally state their answer as soon as they knew it. The experimenter then immediately pressed a key that stopped the computer timer and at the same time emptied the grid. After the response was typed in by the experimenter, a new stimulus appeared on the screen. After each trial, the computer recorded participants' response (entered by the experimenter) and response time (with an exactitude of 0.1 s). A brief pause was given between the different choice and no-choice conditions.

In the intervention session, participants were presented all numerosities between 20 and 45 in a choice condition only. The procedure and instructions were exactly the same as for the choice condition from the pretest and posttest session, except for the feedback that was given after each trial. Half of the participants in each intelligence group received outcome feedback (OFB) whereby the students were informed about the accuracy of their numerosity judgment in each trial (i.e., the number of blocks that their answer deviated from the actual numerosity)., whereas the other half received strategy feedback (SFB), which informed students about the adaptiveness of their strategy choice on each trial as indicated by the no-choice data from the pre-test session (see further).[1]

Results

Strategy Repertoire

Figure 1 presents the number of children in each intelligence and feedback group that uses both the addition and subtraction strategy in the choice condition of the different sessions.

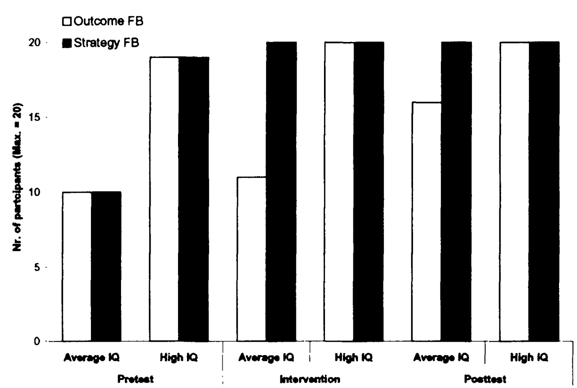


Figure 1: Number of participants in both intelligence and feedback groups that used the addition and subtraction strategy in the choice condition of the pretest, intervention, and posttest session.

As can be derived from the figure, there were more gifted than average intelligent children who applied both strategies in the pretest session and, as expected, there was no difference between both feedback groups in the pretest session. In the intervention session, there was a large increase in the number of average intelligent children with

[1] For each subject, the trials appeared in a predetermined random order so that the experimenter knew in advance which feedback to provide.

both strategies in their repertoire in the SFB group but only a minor increase in the OFB group. In the posttest session, the number of average intelligent children in the OFB group that used both strategies further increased, however, without reaching the maximum. In the high intelligence group, this maximum was already reached for both feedback groups in the intervention session.

Frequency of Strategy Use

A 2 (intelligence: average vs. high) X 2 (feedback type: OFB vs. SFB) X 3 (session: pretest, intervention, and posttest) ANOVA with repeated measures on the last factor was conducted on the percentage use of subtraction in each of the three choice conditions.

The analysis showed a significant main effect of intelligence, $F(1, 76) = 24.62, p < .0001$, indicating that the gifted children used the subtraction strategy more often ($M = 61\%$) than the average intelligent children ($M = 41\%$). We also observed a significant main effect of feedback, $F(1, 76) = 8.77, p = .004$, demonstrating that children in the SFB condition ($M = 57\%$) used the subtraction strategy more frequently than children in the OFB condition ($M = 45\%$) There was also a significant main effect of session, $F(2, 152) = 32.00, p < .0001$, showing an increase in the frequency of the subtraction strategy during the course of the study (*Ms*: 40%, 54%, and 58% for pretest, intervention and posttest, respectively). We also observed a significant intelligence X session interaction, $F(2, 152) = 10.96, p < .0001$,which showed that the significant increase in the use of the subtraction strategy from the pretest to the intervention session was restricted to the average intelligence group. Finally, we observed a significant feedback X session interaction, $F(2, 152) = 3.58, p = .03$, revealing that the more frequent use of the subtraction strategy in the SFB group than compared to the OFB group occurred in the intervention and posttest session but not in the pretest session.

Strategy Efficiency

Strategy efficiency was examined in terms of solution times (i.e., RTs) and of error rates (i.e., absolute deviations between the given response and the actual numerosity). Since only the no-choice conditions provide unbiased measures of strategy performance (Siegler & Lemaire, 1997), we will only discuss the strategy efficiency as measured under no-choice conditions.

Solution times Solution times were only analysed for items that were solved correctly. A 2 (intelligence: average vs. high) X 2 (type of feedback: OFB vs. SFB) X 2 (session: pretest vs. posttest) X 2 (strategy: addition vs. subtraction) ANOVA with repeated measures on the last two factors revealed a significant main effect of intelligence, $F(1, 76) = 11.58, p = .001$, indicating that gifted children ($M = 11.68$ s) were faster in their strategy execution than average intelligent children ($M = 13.53$ s). We also found a significant main effect of session, $F(1, 76) = 21.27, p < .0001$, showing that children were faster in the posttest ($M = 12.10$ s) than in the pretest session ($M = 13.98$ s). Moreover, there was a significant main effect of strategy, $F(1, 76) = 335.76, p < .0001$, indicating that the subtraction strategy ($M = 10,85$ s) was executed faster than the addition strategy ($M = 14.36$ s)[2]. Furthermore, we observed a significant intelligence X strategy interaction, $F(1, 76) = 22.74, p < .0001$ and a significant session X strategy interaction, $F(1, 76) = 35.81, p < .0001$. Both interactions were involved in a significant intelligence X strategy X session interaction, $F(1. 76) = 20.26, p < .0001$. This interaction showed that, notwithstanding that the average intelligent children showed a dramatic increase in subtraction strategy speed from the pre- to the posttest session, this group remained significantly slower in the execution of that strategy compared to the gifted children. There were no significant differences in the speed of the addition strategy, neither between intelligence groups nor between sessions.

Error Rates A similar analysis as for the solution times was conducted on the error rates. This analysis revealed a significant main effect of intelligence, $F(1, 76) = 10.24, p = .002$, indicating that gifted children ($M = 0.22$) were more accurate in their responses than the average intelligent children ($M = 0.43$). A significant main effect of strategy was also observed, $F(1, 76) = 10.27, p = .002$, showing that the addition strategy ($M = 0.24$) was more accurate than the subtraction strategy ($M = 0.42$). We also found a significant main effect of session, $F(1, 76) = 12.56, p = .0007$, revealing an increase in overall accuracy between the pre- ($M = 0.41$) and the posttest session ($M = 0.24$). We observed an intelligence X strategy interaction, $F(1, 76) = 6.04, p = .02$, showing that the subtraction strategy was less accurate than the addition strategy in the average intelligence group but not in the gifted group. Furthermore, it showed that the gifted children applied the subtraction strategy more accurately than the average intelligent children, but not the addition strategy. Finally, there was a session X strategy interaction, $F(1, 76) = 4.11, p < .05$, showing that the initial difference between the addition and subtraction strategy in the pretest session disappeared in the posttest session due to a significant increase in the accuracy of the subtraction strategy from pre- to posttest.

Adaptiveness of Strategy Choices

Adaptiveness of strategy choices was analyzed by means of an analytical technique devised by Luwel et al. (2003). With this technique, one can compare the location of the actual change point (i.e., the trial on which participants switched from the addition towards the subtraction strategy in the

[2] Given that previous findings (e.g., Luwel, et al., 2005) have shown that the addition strategy is faster than the subtraction strategy, this finding seems to be counter-intuitive. However, it is important to note that 76% of the items included a grid that was more than half filled with green blocks. It is especially on these large-numerosity items that the subtraction strategy becomes faster than the addition strategy.

choice condition) with the location of the optimal change point as estimated by the no-choice data.

Since the actual change point cannot always be determined unambiguously, we applied the same criterion as Luwel, et al (2005): the first numerosity on which participants started to use the subtraction strategy and did so for at least three consecutive numerosities. The optimal change point was determined by fitting a linear regression on the individual response time patterns of the correctly solved trials of both no-choice/addition and no-choice/subtraction conditions (see Figure 2). The numerosity on which both regression lines intersect each other is considered as the optimal change point, since from this trial on, the subtraction strategy becomes faster than the addition strategy without a loss of accuracy. Since the projected change point indicates for each individual the trial on which it would be most efficient to switch from the addition strategy towards the subtraction strategy, the absolute difference in location between the actual and the optimal change point can be conceived as a measure of adaptiveness: the smaller this difference, the better an individual's strategy choices are calibrated to his/her unbiased estimates of strategy performance.

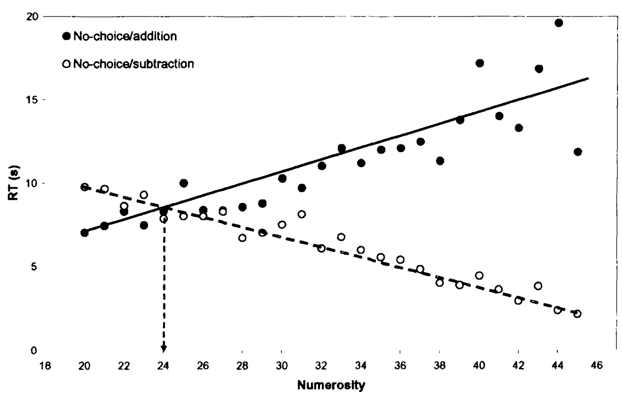


Figure 2: Example of two individual response-time patterns from respectively a no choice/addition and no-choice/subtraction session with their corresponding linear regression lines.

A 2 (intelligence: average vs. high) x 2 (session: pretest vs. posttest) x 2 (type of feedback: outcome feedback vs. strategy feedback) ANOVA was run on the difference scores between the actual and the optimal change point. First, we found a main effect of intelligence, $F(1, 76) = 18.14$, $p < .0001$, indicating that the gifted children ($M = 4.34$) were more adaptive than the average intelligent children ($M = 7.74$). A significant main effect of feedback type was also observed, $F(1, 76) = 6.44$, $p = .01$, showing that children in the SFB group ($M = 5.03$) were more adaptive than children in the OFB group ($M = 7.05$). We also found a significant main effect of session, $F(1, 76) = 36.16$, $p < .0001$, indicating that the adaptiveness increased drastically from the pre- ($M = 7.71$) to the posttest session ($M = 4.36$). Furthermore, we observed a feedback X session interaction, $F(1, 76) = 6.77$, $p = .01$ and an intelligence X session interaction, $F(1, 76) = 9.04$, $p = .004$. Both interactions were involved in an intelligence X feedback X session interaction, $F(1, 76) = 4.44$, $p = .04$. Additional testing showed that, in the pretest session, there was no difference between both feedback groups and that the gifted children were more adaptive than the average intelligent children. In the posttest session, we found that the average intelligent students in the SFB group had made a significantly greater improvement in adaptiveness than the average intelligent children from the OFB group. Actually, the increase in adaptiveness in the SFB group was so large that the initial difference with the gifted children disappeared. The gifted children only showed a slight (non-significant) increase in adaptiveness from the pre- to posttest and within the gifted children there was also no difference in adaptiveness between both feedback groups.

Discussion

The present study demonstrated the effect of intelligence on each of the four parameters of strategic competence. Gifted children used the smart subtraction strategy (next to the addition strategy) in each session and with a greater frequency compared to the average intelligent children. Almost all gifted students' repertoire included both strategies right from the beginning of the task showing an inclination to invent and use more advanced strategies than average intelligent children. The measures of the strategy efficiency and the adaptiveness of strategy choices showed a superiority of the gifted children, as well. Of interest is the finding that gifted children didn't improve in each of the four parameters of the strategic competence as a result of the intervention. The type of feedback played a crucial role: on almost all strategic parameters, except for strategy efficiency, it was found that strategy feedback lead to greater improvements than outcome feedback. In general, one can conclude that the gifted children were already performing at an almost optimal strategic level and, therefore, there was little or no room for further strategic improvement in this group.

The present research showed qualitative differences of the gifted children compared to the average ones in their learning, since they revealed a different cognitive profile in their strategy use and execution. This finding has very important implications in relation to the Education of gifted students since strategic competence is apparent in almost all school lessons. The present study suggests that educating gifted children in learning and applying strategies that they already know has little effect on them. A more appropriate approach would be to teach them by their own learning pace and/or cognitive style. There have already been developed a number of instructional systems and techniques that take such a differentiated approach into account, such as providing differentiated curriculums (VanTassel-Baska, 1997; Ward, 1961), the Problem-Based Learning System (Gallagher, 1997) and the Self-Directed Learning System

(Treffinger, 1986) and teaching methods like the problem solving and the independent study method (Coleman & Cross, 2001).

To conclude, the present study demonstrated the value of the theoretical framework of Lemaire and Siegler (1995) in combination with the choice/no-choice method to unravel the contribution of intelligence on different aspects of strategic performance. These findings, in conjunction with the findings from the Educational research, could help orientating Gifted Education.

Acknowledgments

Koen Luwel is Postdoctoral Fellow of the Fund for Scientific Research-Flanders (Belgium). This study is co-financed within Op. Education by the ESF (European Social Fund), National Greek Resources, and the GOA grant 2006/1 from the Research Fund K.U. Leuven, Belgium to the CIP&T.

References

Bouffard-Bouchard, T., Parent, S., & Larivee, S. (1993).Self-regulation on a concept–formation task among average and *gifted* students. *Journal of Experimental Child Psychology, 56,* 115-134.

Coleman, L., & Cross, T. (2001).*Being gifted in school.* Prufrock Press, Inc.

Gallagher, S. (1997).Problem-based learning: Where did it come from, what does it do, and where is it going? *Journal for the Education of the Gifted, 20,* 332-362.

Gaultney, J.F., Bjorklund, D.F., Goldstein, D. (1996).To be young, gifted and strategic: Advantages for memory performance. *Journal of Experimental Child Psychology, 61,* 43-66.

Geary, D.C., & Brown, S.C. (1991). Cognitive addition: Strategy choice and speed-of-processing differences in gifted, normal, and mathematically disabled children. *Developmental Psychology, 27,* 398-406.

Hong, E. (1999). Studying the Mind of the Gifted. *Roeper Review, 21 (4),* 244-52.

Imbo, I., & Vandierendonck, A. (in press). The role of phonological and executive working-memory resources in simple arithmetic strategies. *European Journal of Cognitive Psychology.*

Ippel, M.J., & Beem, A.L. (1987). A theory of antagonistic strategies. In E.D. Corte, H.Lodewijks, R.Parmentier, & P. Span (Eds.), *Learning and instruction* (Vol. 1). New York: Pergamon.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance. *Psychological Bulletin, 119,* 254-84.

Larkin, J. H., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Models of competence in solving physics problems. *Cognitive* Science, *4*, 317-345.

Lemaire, P., & Siegler, R. S. (1995). Four aspects of strategic change: Contributions to children's learning of multiplication. *Journal of Experimental Psychology: General, 124,* 83-97.

Luwel, K., Lemaire, P., & Verschaffel, L. (2005). Children's strategies in numerosity judgement. *Cognitive Development, 20,* 448-471.

Luwel, K., Verschaffel, L., Onghena, P., & De Corte, E. (2003). Analysing the adaptiveness of strategy choices using the choice/no-choice method: The case of numerosity judgement. *European Journal of Cognitive Psychology, 15(4),* 511-537.

Priest, A. G., & Lindsay, R. O. (1992). New light on novice-expert differences in physics problem solving. *British-Journal-of-Psychology, 83,* 389-405.

Rohwer, W. D., Jr. (1973). Elaboration and learning in childhood and adolescence. In H. W. Reese (Ed.), *Advances in child development and behaviour.* New York: Academic Press.

Saccuzzo, D. P., Johnson, N. E., & Guertin, T. L. (1994). Information processing in gifted versus nongifted African American, Latino, Filipino, and white children: Speeded versus nonspeeded paradigms. *Intelligence, 19,* 219-243.

Scruggs, T. E., & Cohn, S. J. (1983). Learning characteristics of verbally gifted students. *Gifted Child Quarterly, 27,* 169-172.

Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking.* New York: Oxford University Press.

Siegler, R. S. (2005). Children's learning. *American Psychologist, 60,* 769-778.

Siegler, R.S., & Lemaire, P. (1997).Older and younger adult's strategy choices in multiplication: Testing predictions of ASCM using the choice/no choice method. *Journal of Experimental Psychology: General, 126,* 71-92.

Torbeyns, J., Arnaud, L., Lemaire, P., & Verschaffel, L. (2004). Cognitive change as strategy change. In A. Demetriou & A. Raftopoulos (Eds.), *Cognitive Developmental Change: Theories, models and measurement.* Cambridge: Cambridge University Press.

Torbeyns, J., Verschaffel, L. & Ghesquière, P.(2004). Strategy development in children with mathematical disabilities: insights from the choice/no-choice method and the chronological-age/ability-level-match design. *Journal of Learning Disabilities, 37,* 119-131.

Treffinger, D. (1986).*Blending gifted education with the total school program.* Buffalo, NY: D.O.K.

VanTassel-Baska, J. (1997).What matters in curriculum for gifted learners: Reflections on theory, research, and practice. In N. Colangelo & G. Davis (Eds.), *Handbook of gifted education* (2nd ed., pp. 113-125). Boston: Allyn and Bacon.

Ward, V. S. (1961).*Educating the gifted: An axiomatic approach.* Columbus, OH: Charles E. Merrill.

Wong, B. (1982). Strategic behaviors in selecting retrieval cues in gifted, normal achieving and learning-disabled children. *Journal of Learning Disabilities, 15,* 33-37.

"Knowledge and Information Awareness" for Enhancing Computer-Supported Collaborative Problem Solving by Spatially Distributed Group Members

Tanja Engelmann (née Keller) **(t.engelmann@iwm-kmrc.de)**
Knowledge Media Research Center, Konrad-Adenauer-Str. 40
72072 Tuebingen, Germany

Sigmar-Olaf Tergan (s.tergan@iwm-kmrc.de)
Knowledge Media Research Center, Konrad-Adenauer-Str. 40
72072 Tuebingen, Germany

Abstract

Computer-supported collaboration via the Internet becomes increasingly important in many educational and workplace settings. However, there are still problems regarding computer-supported collaboration, especially interaction problems within groups. In this paper, we suggest enhancing "knowledge and information awareness" (KIA) to solve these problems. KIA is defined as awareness of a group member with regard to task-relevant domain knowledge and information underlying this knowledge of her/his collaborators. In this paper, an empirical study is presented, which investigates whether KIA is an efficient means to foster computer-supported collaborative problem solving. In this study, an experimental condition, in which the group members of a triad are provided with an environment for enhancing "knowledge and information awareness", is compared to a control condition, in which the group members are not provided with this environment. Results showed that groups with a KIA environment performed better in their problem-solving tasks than groups without one.

Introduction

Today's information society involves significant changes in the world of learning and working. Given the complexity of modern problems and the ill-structuredness of subject matter, combined with the impossibility of everyone meeting at the same location, computer-supported collaboration between individuals becomes necessary. In order to collaborate effectively, there is a need to be aware of the subject-matter knowledge of the collaborators and the information their knowledge is based upon. However, fostering awareness of the individual group members' task-relevant knowledge and information is still a major problem in virtual collaboration settings.

In this paper, the potential and the problems associated with computer-supported collaboration are first outlined. Afterwards, an innovative solution for problems in computer-supported collaboration is described. This solution is built upon an approach for making individual group members aware of the knowledge and information resources of other members, which are necessary for coping effectively with a task. In order to confirm the efficiency of the suggested approach, an empirical study is presented and its results are discussed. The paper ends with conclusions.

Potential and Problems Associated with Computer-Supported Collaboration

Computer-supported collaboration becomes increasingly important when learners have to construct a shared knowledge and information basis in order to cooperatively solve problems by using the Internet as a communication medium. According to Koschmann (2002), computer-supported learning (CSCL) could be characterized as "practices of meaning-making in the context of joint activity, and the ways in which these practices are mediated through designed artifacts" (p. 18). Following this often cited definition, there are two important features that characterize CSCL: First, the collaboration aspect implies that a group, not only an individual, is involved. Stahl, Koschmann, and Suthers (2006) explain that this group learning is not merely accomplished interactionally, but is actually constituted of the interactions between participants. This statement points out that, in such situations, the interaction between the group members is essential for group efficiency. Second, Koschmann's definition highlights the aspect of mediation through designed artifacts. This aspect refers to the computer support of the group interaction, i.e., the technology should be designed to mediate and encourage social acts that lead to efficient group work. It is important to mention that the research area of CSCL does not only include learning settings, but also settings that are learning relevant, such as computer-supported collaborative decision making or problem solving (e.g., Fjermestad, 2004). In this paper, computer-supported collaborative problem solving is the focus.

Results of empirical research suggest that learners in computer-supported collaboration may provide more complete reports, may make decisions with higher quality, and may be better in idea generation (Fjermestad, 2004). However, research results also show that efficient computer-supported collaboration is not easy to achieve (Dewiyanti, Brand-Gruwel, & Jochems, 2005). According to Janssen, Erkens, Jaspers, and Broeken (2005), groups who are collaborating with computer support often have communication and interaction problems. They may perceive their discussion as confused (Thompson & Coovert, 2003), they may need more time to arrive at a consensus and for making decisions (Fjermestad, 2004), and they may need more time for solving tasks (Baltes, Dickson, Sherman, Bauer, & LaGanke, 2002). Following the conclusions of Carroll, Neale, Isenhour, Rosson, and McCrickard (2003), in CSCL settings, the group task is often not perceived as a group task; i.e., the group members work individually instead of collaboratively, and coordination is missing. In addition, the individual group members often do not trust in the fact that the others are doing their part of the work.

In the CSCL research community, there are different strands of research addressing such problems of CSCL. On the one hand, there are approaches that foster computer-supported collaboration by explicit methods like scripting (e.g., Kollar, Fischer, & Slotta, 2005), i.e., the learners are instructed how they should behave to be efficient. On the other hand, there are approaches that seek to support computer-supported collaboration by using implicit methods focused on enhancing different kinds of *group awareness* (e.g., Gross, Stary, & Totter, 2005). These implicit approaches provide no instructions, but inform learners about relevant information. They assume that the group members have the ability to collaborate efficiently if they are informed regarding relevant information, i.e., if they are aware regarding this information.

"Knowledge and Information Awareness" as an Innovative Solution for Enhancing Computer-Supported Collaboration

Group awareness according to Gross et al. (2005) is defined as "consciousness and information of various aspects of the group and its members" (p. 327). However, in the literature, there is no consensus about the definition of the term group awareness. Some authors try to differentiate it according to several dimensions. Carroll et al. (2003), for example, differentiate three different types of group awareness on the working processes level: While social awareness is defined as awareness regarding who is currently available for collaboration, action awareness additionally provides information regarding who is doing what at the moment, as well as who did what recently. This last type of awareness refers to feedback on single occurrences. However, Carroll et al. (2003) point out the importance of activity awareness for computer-supported collaborative scenarios. They defined activity awareness as awareness regarding not only who is currently available and is doing what, but also awareness regarding the relevance of an activity with regard to the group goal.

In most papers, the meaning of awareness refers to both social awareness and action awareness. However, in specific situations, social and action awareness may not be enough to support effective collaboration, but rather knowledge is needed about the mental representations regarding the task domain of each of the group members, the concepts and information resources they use and share, as well as the knowledge gaps that are responsible for misunderstandings, ineffective shared knowledge construction, and deficient problem solving. In such situations, "*knowledge and information awareness*" (*KIA*) is needed. KIA is defined as awareness of a group member regarding both the knowledge and the information underlying this knowledge of her/his collaborators (Keller, Tergan, & Coffey, 2006).

A situation in which KIA is necessary arises, for example, when spatially distributed group members with different domain expertise have to solve a task together that requires not only the expertise of the group members, but also knowledge about a large amount of task-relevant information resources that is distributed among the experts. The need for KIA results from the explosive increase in information and information resources. This information flood requires a changed handling of information, namely a self-regulated, resource-based activity (Rakes, 1996). The handling of complex contents or a large amount of information or information resources can lead to cognitive overload, which may hinder the efficiency of an individual or a group (Chandler & Sweller, 1991).

Visualizations are suggested to reduce cognitive load while interacting with large and complex amounts of information, because they could be used as cognitive tools for overcoming limits of cognitive capacity (Ware, 2005). Concept maps, developed by J. D. Novak (e.g., Novak & Gowin, 1984), are a type of knowledge visualization for representing the knowledge of an individual by means of nodes displaying concepts and labeled links between the nodes representing the relations between the concepts. While traditional concept maps were created using paper and pencil, computer-based concept mapping tools allow for the creation of digital concept maps. An example is CmapTools developed by the Florida Institute of Human Machine and Cognition in Pensacola, (USA). Traditional concept maps have been criticized for some shortcomings in representing knowledge. For example, they only visualize abstract concept knowledge, leaving the information underlying the concepts (e.g., examples and images of a concept) unconsidered (e.g., Tergan, Keller, & Burkhard, 2006). By contrast, advanced digital concept mapping tools allow the representation of information underlying the conceptual knowledge. These types of visualizations combine the advantages of both traditional knowledge visualizations and information visualizations (Tergan et al., 2006). Information visualizations, with their origin in computer science, are interactive, spatial-visual representations of abstract data (Card, Mackinlay, & Shneiderman, 1999). In Tergan and Keller (2005), as well as in Tergan et al. (2006), the potential of synergistic approaches between information visualizations and knowledge visualizations is presented and discussed. KIA could be enhanced by means of such knowledge and information visualizations by using CmapTools.

When using an environment based on knowledge and information visualizations, users are not only able to check visually which concept is based on an information resource, but can also access information relevant for an explanation of a concept and its relation to other concepts. It is suggested that being aware of one's own knowledge and the knowledge of others, as well as the information resources linked to the concepts, may help cooperative problem solvers in shared knowledge-construction and problem-solving tasks. This assumption is based on the theory of transactive memory (Wegner, 1986). According to this theory, a transactive memory system is a set of individual memory systems combined with communication between the group members. This enhances the expertise of each group member, because everyone has access to the knowledge and information of the others.

It is assumed that KIA is helpful in a computer-supported, collaborative problem-solving scenario, because it can be expected that KIA will have a positive impact on interaction, especially on the processes and the effectiveness of communication, coordination, and collaborative problem solving. On the one hand, according to Clark and Brennan (1993), shared understanding in communication is crucial for individuals working in a group. Making visual representations of the knowledge structures and the underlying information of each group member available to the group should facilitate shared understanding and knowledge construction. On the other hand, the exchange of unshared information is very important (e.g., Stasser, Vaughan, & Stewart, 2000). It has been shown that information that is shared by all group members is often mentioned in group discussion, while unshared information that is known, e.g., by only one group member, mostly remains unmentioned. Such unshared information could be important for problem solving. Therefore, it is important to recognize unshared information. By comparing the external representations of the knowledge structures of the collaborators and the information resources linked to the knowledge elements, group members can easily recognize which knowledge and information is shared and which is not. This should have a positive effect on group coordination. In addition, it is assumed that the capability to view the knowledge and underlying information of the others in the group provides a kind of affordance to make use of these representations (Suthers, 2005).

Experimental Study

This experiment investigated whether an environment for fostering awareness regarding the knowledge and underlying information of the collaborators leads to more efficient collaboration (in the sense of coordination and communication) of a group and, as a result, to more efficient problem solving compared to a condition with groups in which the group members do not have a KIA environment.

Method

Participants Participants were 90 students (58 female, 32 male) of the University of Tuebingen, Germany. Average age was 24.47 (*SD* = 3.83). The students were randomly assigned to the experimental condition or to the control condition. Each group consisted of three participants, resulting in 15 control groups and 15 experimental groups.

Materials and Procedures The participants worked in groups of three students in a room that was divided by partition walls into three separate sections. Each of the sections was equipped with a desk and a computer. The participants could not see each other, but could speak with each other. The experimental environment used in this study provided information elements that are necessary to care for a fictitious kind of spruce forest. These information elements consisted of 13 concepts, 30 relations between these concepts, and 13 background resources, i.e., information underlying a concept, and were evenly distributed among the three group members. Each participant had access to several concepts, relations, and background resources that were unshared, shared with one collaborator, or shared with both collaborators. The experimental environment consisted of two software components. The first was an information space that contained the different information units the group members needed for solving the problems. This information space was based on a Zope3-based groupware that was developed by the Knowledge Media Research Center in Tuebingen, Germany. The other was CmapTools (described above).

Procedure: (1) At the start of the study, the participants took a pencil-paper diagnostic test aimed at assessing the *control variables*, i.e., their experience with computers, mapping techniques, and group work. (2) Afterwards, they received an introduction and *practice using CmapTools* (without time limit). (3) After ensuring that all participants could use CmapTools without problems, they started with *individual phase 1* of the experiment: At the outset of this phase, participants were told that they are experts who have to protect a spruce forest and that they first have to refresh their domain expertise before they start to collaborate and find a common solution for the problems. During this phase, which lasted 23 minutes, the group members worked separately, accessing the information elements in their own information window located at the left side of their screen and structuring their information and knowledge in their own working window located at the right side of their screen. (4) In the *individual phase 2*, each participant of the control group had 5 minutes to examine her/his own map (see Figure 1, left side). Each participant of the experimental group, however, had 5 minutes to view her/his own map, as well as the maps of her/his collaborators (see Figure 1, right side). The 23-minutes and 5-minutes time slots were based on experience from a pilot study. (5) After this activity, all participants had to fill out a 15-items questionnaire used as a *manipulation check* to measure the amount of knowledge the participants acquired from the maps. (6) Subsequently, the three group members had to *collaborate to solve two problems*, i.e., which pesticide and which fertilizer they would use to protect the spruce forest. To solve these problems, the participants needed to compile the knowledge and information they had structured and visualized in the individual phase 1 in the form of a digital concept map.

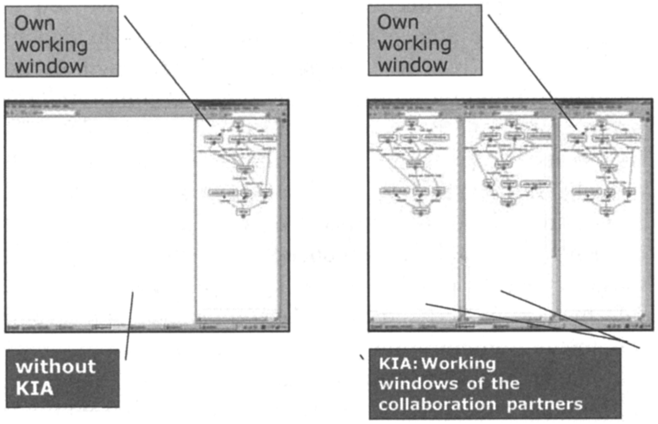


Figure 1: Individual phase 2
(left: control group; right: experimental group).

To do this, they used a shared working space to create a common digital concept map containing all the knowledge and information they acquired in the individual phase. Based on experience from a pilot study, they had 40 minutes for collaboration. During this phase, they could speak with one another. They were told that they were using a kind of hands-free speaking system. In the control condition, the participants could only see their own working window and the shared working window (see Figure 2, above). In the experimental condition, the participants also saw the individual maps of their collaborators, i.e., they were also aware of the knowledge and information their collaborators had (see Figure 2, below). The individuals' interactions were recorded as log files and audio data.

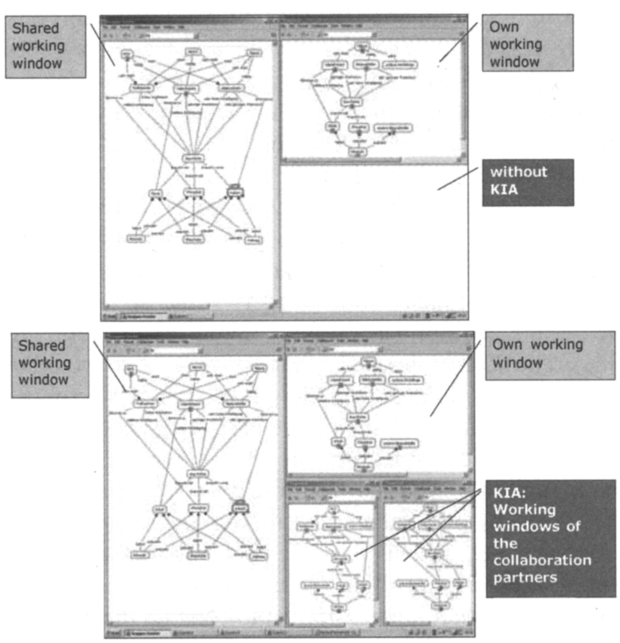


Figure 2: Collaborative phase
(above: control group; below: experimental group).

(7) After this collaborative phase, the participants were given another *test* containing 30 items *to measure the knowledge* they had acquired regarding taking care of the spruce forest. In this test phase, the experimental environment was no longer available. There were no time limits on this test. (8) At the end of the study, participants had to fill out a *questionnaire* asking about difficulties regarding communication and collaboration, the use of CmapTools, and the helpfulness of the KIA environment.

Design and Dependent Measures The analysis was based on a comparison of the control and the experimental condition. In the experimental condition, the participants were provided with a KIA environment, i.e., they could see the individual concept maps of their collaboration partners and, therefore, could become aware of the knowledge and information their collaborators had. In the control condition, the participants were not provided with a KIA environment.

With regard to the dependent measures, the distinction was made between product-related measures and process-related measures. The *product-related measures* could be divided into three categories:

First, the *domain knowledge* measured with 30 multiple-choice test items: Several sub-variables could be differentiated, for example, knowledge regarding relations and contents underlying a concept, as well as knowledge pertaining to whether it is unshared, shared with one other member, or shared with both members.

Second, the *quality of the common concept map* that the group created in the collaboration phase: Several sub-variables were used, for example, the number of correct nodes and relations.

Third, the *quality of the group answers to the two problem-solving tasks,* measured by means of the number of correct solutions and correct reasons.

Regarding the *process-related measures*, the communication and collaboration aspects were of interest: In the collaboration phase, the development of the group map was recorded in a *log file* and the verbal communications were recorded in an *audio file* for later analysis. In addition, *subjective items* were captured through a questionnaire.

Results and Discussion

Several ANCOVAs were performed. In all analyses of variance reported in this paper, the control measure item "experience in creating computer-based graphics" was used as a covariate. The reason is that with regard to this item, a significant difference existed between the control condition and the experimental condition, with a higher value in the sense of more experience in the control condition. In addition, this item was strongly associated with dependent measures. With regard to other control items, there were no significant differences between the control and the experimental condition. All analyses presented here are based on group level, that is, the group values are calculated as means of the values of the individuals of a group.

Analysis on the group level was necessary, due to the fact that the group members were not independent of each other.

The first analysis determined whether the KIA environment was used by the participants in the experimental condition. For this purpose, the questionnaire items were analyzed through the use of a five-point rating scale, with the number one for "no agreement", the number three for "partial agreement", and the number five for "complete agreement". The experimental groups agreed on average that it was helpful to have an overview of the maps of the collaborators ($M_E = 4.27$; $SD_E = 0.75$) and that seeing the maps of the others was useful ($M_E = 3.58$; $SD_E = 0.58$).

The second analysis explored whether the use of the KIA environment had an effect on the dependent measures: The questionnaire at the end of the study showed that the study was more stressful for participants in the control condition ($M_C = 3.2$; $M_E = 2.7$; $F(1,27) = 4.66$; $MSE = 0.28$; $p < .05$), although the experimental condition had more problems regarding the use of the different windows on the desktop ($M_C = 1.8$; $M_E = 2.2$; $F(1,27) = 6.25$; $MSE = 0.25$; $p < .05$) compared to the control groups. This last result was not unexpected, due to the fact that, in the experimental condition, the participants had to work with two more windows than in the control condition. The previous result showed that the cognitive load in the control condition was higher than in the experimental condition. In addition, in the experimental condition, the participants stated that the collaboration with each other led to a better overview regarding the relations of the domain compared to the control groups ($M_C = 4.0$; $M_E = 4.3$; $F(1,27) = 5.89$; $MSE = 0.22$; $p < .05$). This could be confirmed by the analysis of the domain knowledge measures: The analysis revealed marginally better performance for the experimental groups regarding the knowledge on domain relations compared to the control groups ($M_C = 3.4$; $M_E = 3.7$; $F(1,27) = 3.43$; $MSE = 0.21$; $p = .075$). Regarding the domain knowledge performance, the experimental condition gained a higher performance on domain relations that were shared by a participant collaborator dyad as compared to the control groups ($M_C = 2.1$; $M_E = 2.4$; $F(1,27) = 4.2$; $MSE = 0.14$; $p < .05$). This result constitutes evidence for the helpfulness of the KIA environment, because the participants were aware of which other collaborator had the same relation knowledge that they had. In addition, the analyses revealed higher performance by the experimental groups with regard to knowledge about information that is linked to concepts: In this context, the experimental groups gained higher values in knowledge regarding information that is only shared by the other collaborators; that is, the participant himself did not have this information ($M_C = 2.6$; $M_E = 2.9$; $F(1,27) = 4.17$; $MSE = 0.41$; $p = .05$). This result *also* provides evidence of the efficiency of the KIA environment: Considering information underlying a concept, participants in the experimental condition did remember more often items that both other collaborators had. In respect of the quality of the group maps there were no significant differences between the conditions with regard to the included correct relations ($M_C = 23.3$; $M_E = 21.5$; $F(1,27) = 1.81$; $MSE = 22.21$; $p = .19$) or correct nodes ($M_C = 12.9$; $M_E = 12.6$; $F(1,27) = 1.71$; $MSE = 0.46$; $p = .20$). With regard to the problem-solving tasks, the experimental groups tended to be more confident that they had solved the two tasks correctly as compared to the control group (w.r.t. the pesticide problem: $M_C = 3.8$; $M_E = 4.2$; $F(1,27) = 3.38$; $MSE = 0.47$; $p = .077$; w.r.t. the fertilizer problem: $M_C = 3.8$; $M_E = 4.2$; $F(1,27) = 3.17$; $MSE = 0.57$; $p = .086$). This subjective estimation is partly mirrored in objective results, namely in the group answers given: Regarding the number of correct answers to the pesticide problem, the data did not show a significant difference between the conditions ($Pearson\text{-}\chi 2$ (2) = 3.20; $p = .20$). However, with regard to the reasons given as to why they chose the correct pesticide, the experimental condition was marginally superior to the control condition ($M_C = 0.2$; $M_E = 0.8$; $F(1,27) = 3.36$; $MSE = 0.7$; $p < .1$). By contrast, regarding the number of correct answers to the fertilizer problem, the experimental condition achieved a marginally higher performance compared to the control condition ($Pearson\text{-}\chi^2$ (2) = 4.9; $p < .1$). But with regard to the reasons given as to why they chose the correct fertilizer, there was no significant difference between the groups ($M_C = 0.7$; $M_E = 1.3$; $F(1,27) = 0.79$; $MSE = 1.01$; $p = .38$).

Conclusions

The presented study demonstrated that computer-supported collaborative problem solving can be supported by enhancing KIA, i.e., awareness of a group member with regard to the knowledge and the underlying information of the other collaborators. In this study, an experimental condition using an environment for enhancing KIA was compared to a control condition that worked without it. Results of the analysis showed that the participants of the experimental condition evaluated the use of the KIA environment as helpful. Comparing the two conditions, it could be shown that the study was more stressful under the control condition, although the experimental condition had more difficulties in using the windows. Therefore, the benefit of using a KIA environment seems to be great enough to compensate for the higher cognitive load caused by the need to use more windows on the screen. The analyses also showed that the experimental groups achieved higher performance in both knowledge regarding content information that was only shared by the other collaborators and in knowledge regarding relation information that both an individual and another collaborator had. In addition, the study demonstrated that using a KIA environment was helpful for problem-solving performances. The results support hypotheses concerning the support of computer-supported collaborative problem solving by enhancing KIA. Further research activities will investigate in greater detail the factors that are causative for the efficiency of the KIA environment.

Acknowledgements

This project was supported by the KMRC (Tuebingen, Germany). We thank Dr. J. Coffey of the IHMC (Pensacola, USA) who adjusted CmapTools for us and the technical group of the KMRC for the technical assistance.

References

Baltes, B. B., Dickson, M. W., Sherman, M. P., Bauer, C. C., & LaGanke, J. (2002). Computer-mediated communication and group decision making: A meta-analysis. *Organizational Behavior and Human Decision Processes, 87*(1), 156-179.

Card, S. K., Mackinlay, J. D. & Shneiderman, B. (1999). Information visualization. In S. K. Card, J. D. Mackinlay & B. Shneiderman (Eds.), *Information visualization. Using vision to think* (pp. 1-34). San Francisco: Morgan Kaufmann.

Carroll, J. M., Neale, D. C., Isenhour, P. L., Rosson, M. B., & McCrickard, D. S. (2003). Notification and awareness: Synchronizing task-oriented collaborative activity. *International Journal of Human-Computer Studies, 58*(5), 605-632.

Chandler P., & Sweller J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*, 293-332.

Clark, H. H., & Brennan S. E. (1993). Grounding in communication. In R. E. Baecker (Ed.), *Readings in groupware and computer-supported cooperative work assisting human collaboration* (pp. 222-233). San Mateo, CA.: Morgan Kaufman

Dewiyanti, S., Brand-Gruwel, S., & Jochems, W. (2005). Learning together in an asynchronous computer-supported collaborative learning environment: The effect of reflection on group processes in distance education. *Paper presented at Earli*, 2005, Nicosia, Cyprus.

Fjermestad, J. (2004). An analysis of communication mode in group support systems research. *Decision Support Systems, 37*(2), 239-263.

Gross, T., Stary, C., & Totter, A. (2005). User-Centered Awareness in Computer-Supported Cooperative Work-Systems: Structured Embedding of Findings from Social Sciences. *International Journal of Human-Computer Interaction, 18*, 323-360.

Janssen, J., Erkens, G., Jaspers, J., & Broeken, M. (2005). Effects of visualizing participation in computer-supported collaborative learning. *Paper presented at Earli*, 2005, Nicosia, Cyprus.

Keller, T., Tergan, S.-O., & Coffey, J. (2006). Concept maps used as a "knowledge and information awareness" tool for supporting collaborative problem solving in distributed groups. In A. J. Cañas, & J. D. Novak (Eds.), *Concept Maps: Theories, Methodology, Technology* (pp. 128-135). San José: Sección de Impresión del SIEDIN.

Kollar, I., Fischer, F., & Slotta, J. D. (2005). Internal and external collaboration scripts in webbased science learning at schools. In T. Koschmann, D. Suthers, & T. W. Chan (Eds.), *Computer Supported Collaborative Learning 2005: The Next 10 Years* (pp. 331-340). Mahwah, NJ: Lawrence Erlbaum.

Koschmann, T. (2002). Dewey's contribution to the foundations of CSCL research. In G. Stahl (Ed.), *Computer support for collaborative learning: Foundations for a CSCL community* (pp. 17-22). Boulder, CO: Lawrence Erlbaum.

Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn.* New York: Cambridge University Press.

Rakes, G. C. (1996). Using the internet as a tool in a resource-based learning environment. *Educational Technology, 36*(5), 52-56.

Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-supported collaborative learning. In R. K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (pp. 409-426). Cambridge: University Press.

Stasser, G., Vaughan, S. I., & Stewart, D. D. (2000). Pooling unshared information: The benefits of knowing how access to information is distributed among members. *Organizational Behavior and Human Decision Processes, 82*, 102-116.

Suthers, D. D. (2005). Technology affordances for intersubjective learning: A thematic agenda for CSCL. In T. Koschmann, D. D. Suthers & T. W. Chan (Eds.), *Computer Supported Collaborative Learning 2005: The Next 10 Years!* (pp. 662-672). Mahwah, NJ: Lawrence Erlbaum.

Tergan, S.-O., & Keller, T. (Eds.) (2005). *Knowledge and information visualization: Searching for synergies.* LNCS 3426. Berlin:. Springer.

Tergan, S.-O., Keller, T., & Burkhard, R. (2006). Integrating Knowledge and Information: Digital Concept Maps as a Bridging Technology. *Information Visualization, 5* (3), 167-174.

Thompson, L. F., & Coovert, M. D. (2003). Teamwork online: The effects of computer conferencing on perceived confusion, satisfaction and post-discussion accuracy. *Group Dynamics, 7*(2), 135-151.

Ware, C. (2005). Visual queries: The foundations of visual thinking. In S.-O. Tergan & T. Keller (Eds). *Knowledge and information visualization: Searching for synergies* (pp. 27-35). LNCS 3426. Berlin: Springer.

Wegner, D. M. (1986). Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G. R. Goethals (Eds.), *Theories of group behaviour* (pp. 185-208). New York: Springer.

Computational Evidence for Two-Stage Categorization as a Process of Adjective Metaphor Comprehension

Akira Utsumi (utsumi@se.uec.ac.jp)
Department of Systems Engineering
The University of Electro-Communications
1-5-1, Chofugaoka, Chofushi, Tokyo 182-8585, Japan

Maki Sakamoto (sakamoto@hc.uec.ac.jp)
Department of Human Communication
The University of Electro-Communications
1-5-1, Chofugaoka, Chofushi, Tokyo 182-8585, Japan

Abstract

Most of existing metaphor studies address comprehension of nominal metaphors like "My job is a jail" and predicative metaphors like "He shot down all of my arguments". However, little attention has been given to how people comprehend adjective metaphors such as "red voice". In this paper, we address adjective metaphors and argue that adjective metaphors are comprehended via a two-stage categorization process. In a two-stage categorization process, the adjective of an adjective metaphor evokes an intermediate category, which in turn evokes an abstract category of property to be mapped onto the target noun, rather than directly creating a category of property as predicted by the categorization theory. We then test our argument by means of computer simulation in which the meanings of adjective metaphors are computed from the representations of the adjective and the noun in a multidimensional semantic space constructed by latent semantic analysis. In the simulation, three algorithms for adjective metaphor comprehension, i.e., two-stage categorization, categorization and comparison, were compared in terms of how well they mimic human interpretation of adjective metaphors. The simulation result was that the two-stage categorization algorithm best mimicked human interpretation of adjective metaphors, thus suggesting that the two-stage categorization theory is a more plausible theory of adjective metaphor comprehension than the categorization theory and the comparison theory.

Keywords: Metaphor comprehension; Computational modeling; Latent semantic analysis (LSA); Adjective metaphor; Two-stage categorization

Introduction

Many studies in the domain of cognitive science have been made on the mechanism of metaphor comprehension. Although they have paid much attention to nominal metaphors such as "*My job is a jail*" (e.g., Bowdle & Gentner, 2005; Gentner, Bowdle, Wolff, & Boronat, 2001; Glucksberg, 2001; Jones & Estes, 2006; Utsumi & Kuwabara, 2005) and predicative metaphors such as "He *shot down* all of my arguments" (e.g., Lakoff & Johnson, 1980; Martin, 1992), little attention has been given to adjective metaphors such as "argumentative melody" and how they are comprehended. Some studies (e.g., Shen & Cohen, 1998; Werning, Fleischhauer, & Beşeoğlu, 2006; Yu, 2003) have focused on a synesthetic metaphor, a kind of adjective metaphor in which an adjective denoting the perception of one sense modality modifies a noun denoting a different modality. However these studies only examine how the acceptability of synesthetic metaphors can be explained by the pairing of adjective's and noun's modalities, rather than exploring the mechanism of adjective metaphor comprehension.

In this paper, we address the problem of how adjective metaphors are comprehended and argue that adjective metaphors are comprehended via a *two-stage categorization* process, which is an extended view of Glucksberg's categorization theory (Glucksberg, 2001; Glucksberg & Keysar, 1990). We then test our argument by means of computer simulation of adjective metaphor comprehension. For this purpose, we use a semantic space constructed by latent semantic analysis (LSA) (Landauer & Dumais, 1997) and provide a computational model of the two-stage categorization process, together with computational models of other possible processes for adjective metaphor comprehension such as categorization and comparison. In the computer simulation, we examine how well a computational model embodying each metaphor theory mimics human comprehension by comparing the interpretations of metaphors obtained by the computer simulation with human interpretations of the same metaphors obtained in a psychological experiment (Sakamoto & Sano, 2004). The metaphor theory that achieves the best simulation performance can be seen as the most plausible theory of adjective metaphor.

Adjective Metaphor Comprehension

Metaphor comprehension can be viewed as the process of finding relevant features (or predicates) that constitute the metaphorical meaning from the interaction between a source concept and a target concept, i.e., the process of generating the modified target concept in which some features or properties are highlighted and some other features are downplayed. In the case of adjective metaphors, the target concept is expressed by the head noun and modified by the source concept expressed by or associated with the adjective. The problem is how people determine which features of the target concept are highlighted or downplayed by the source concept.

One probable theory that can explain the mechanism of adjective metaphor comprehension would be the categorization theory of metaphor proposed by Glucksberg and his colleagues (Glucksberg, 2001; Glucksberg & Keysar, 1990). The categorization theory addresses mainly nominal metaphors and argues that people understand nominal metaphors by seeing the target concept as belonging to the superordinate metaphorical category exemplified by the source concept. Glucksberg (2001) has also argued that predicative metaphors function very much as do nominal metaphors; just as nominal metaphors use vehicles that epitomize certain categories of objects or situations, predicative metaphors use verbs that epitomize certain categories of actions. Some empirical evidence in favor of this view of predicative metaphors was also provided by Torreano, Cacciari, and Glucksberg (2005). Therefore, although they do not explicitly mention adjective metaphors in their works, it is likely that the same argument can be applied to adjective metaphors, that is, adjective metaphors use adjectives that epitomize certain categories of properties. According to this view, an adjective metaphor "red voice", for example, is comprehended so that the source concept *red* evokes an ad hoc category of property like "scary, screaming and dangerous" and such metaphorical property is mapped onto the target concept.

Against the categorization theory of adjective metaphors, we propose a two-stage categorization theory. The intuitive idea behind two-stage categorization is that correspondences between the properties literally expressed by the adjective and the properties to be mapped onto the target concept would be indirect, mediated by an intermediate category, rather than direct as predicted by the categorization theory. In the case of "red voice" metaphor, for example, the adjective *red* first evokes an intermediate category "red things", to which blood, fire, passion, apple, danger typically belong. Then exemplars relevant to the target concept *voice* such as blood, passion and danger are selected and they evoke a final abstract category of property like "scary, screaming and dangerous". [1]

An alternative, but probably less likely, explanation of adjective metaphor comprehension is given by the comparison theory of metaphor (Gentner, 1983; Gentner et al., 2001). This theory argues that metaphors are processed via a comparison process consisting of an initial alignment process between the source and the target concepts followed by a process of projection of aligned features into the target concept. According to the comparison theory, the "red voice" metaphor is comprehended in such a way that two concepts *red (or redness)* and *voice* are aligned, some features such as ones about scariness, scream or danger are found, and they are mapped onto the target noun.

In the rest of this paper, we examine which of these three theories best explains the mechanism of adjective metaphor comprehension by comparing them in terms of how accurately computational models embodying these theories simulate human behavior.

Computational Model

Vector Space Model

A vector space model is the most commonly used geometric model for the meanings of words. The basic idea of a vector space model is that words x are represented by high-dimensional vectors $\boldsymbol{v}(x)$, i.e., *word vectors*, and the degree of semantic similarity $sim(x, y)$ between any two words x and y can be easily computed as the cosine $\cos(\boldsymbol{v}(x), \boldsymbol{v}(y))$ of the angle formed by their vectors.

Word vectors are constructed from the statistical analysis of a huge corpus of written texts in the following way. First, all content words in a corpus are represented as m-dimensional feature vectors, and a matrix A is constructed using n feature vectors as rows. Then the dimension of M's rows is reduced from m to l. A number of methods have been proposed for computing feature vectors and for reducing dimensions (Utsumi & Suzuki, 2006). In this paper, we used an LSA technique (Landauer & Dumais, 1997) for constructing word vectors. LSA uses the term frequency in a paragraph as an element of feature vectors, and singular value decomposition as a method for dimensionality reduction. LSA was originally proposed as a document indexing technique for information retrieval, but several studies (e.g., Kintsch, 2001; Landauer & Dumais, 1997) have shown that LSA successfully mimics many human behaviors associated with semantic processing.

For example, using a semantic space derived from a corpus of Japanese newspaper used in this paper, similarity between *computer* ("*konpyuta*" in Japanese) and *Windows* ("*uindouzu*" in Japanese; Microsoft's OS) is computed as .63, while similarity between *computer* and *window* ("*mado*" in Japanese; glass in the wall) is computed as –.02.

Metaphor Comprehension Algorithms

In the vector space model, a vector representation $\boldsymbol{v}(s)$ of a piece of text s (e.g., phrase, clause, sentence, paragraph) consisting of constituent words $w_1, \cdots, w_n$ can be defined as a function $f(\boldsymbol{v}(w_1), \cdots, \boldsymbol{v}(w_n))$. Therefore, adjective metaphor comprehension is modeled as computation of a vector $\boldsymbol{v}(M) = f(\boldsymbol{v}(w_T), \boldsymbol{v}(w_S))$ which represents the meaning of an adjective metaphor M with the noun w_T (target) and the adjective w_S (source). In the rest of this paper, I use the phrase "n neighbors of a word (or a category) x" to refer to words with n highest cosine similarity to x, and denote a set of n neighbors of x by $N_n(x)$.

Categorization The algorithm of computing a metaphor vector $\boldsymbol{v}(M)$ by the process of categorization is as follows.

1. Compute $N_{m_1}(w_S)$, i.e., m_1 neighbors of the source w_S.
2. Selects k words with the highest similarity to the target noun w_T from $N_{m_1}(w_S)$.
3. Compute a vector $\boldsymbol{v}(M)$ as the centroid of $\boldsymbol{v}(w_T)$, $\boldsymbol{v}(w_S)$ and k vectors of the words selected at Step 2.

This algorithm is identical to Kintsch's (2000) predication algorithm and it is also used as a computational model of the categorization process in Utsumi's (2006) simulation experiment. As Kintsch suggests, this algorithm embodies the categorization view in that a set of k words characterizes an abstract superordinate category exemplified by the vehicle.

Two-stage categorization We propose the algorithm of two-stage categorization as follows.

1. Compute $N_{m_1}(w_S)$, i.e., m_1 neighbors of the source w_S.
2. Selects k words with the highest similarity to the target noun w_T from $N_{m_1}(w_S)$.
3. Compute a vector $\boldsymbol{v}(C)$ of an intermediate category C as the centroid of $\boldsymbol{v}(w_T)$, $\boldsymbol{v}(w_S)$ and the vectors of k words selected at Step 2.
4. Compute $N_{m_2}(C)$, i.e., m_2 neighbors of the intermediate category C.
5. Compute a metaphor vector $\boldsymbol{v}(M)$ as the centroid of $\boldsymbol{v}(w_T)$, $\boldsymbol{v}(w_S)$ and m_2 vectors selected at Step 4.

[1] Our preliminary experiment demonstrated that figurative meanings of adjectival metaphors with color adjectives were not directly associated with adjectives, but could be explained more appropriately by considering intermediate concepts associated with both adjectives and target nouns. This finding may lend support to our view based on two-stage categorization.

The first three steps, which are identical to the original categorization algorithm, correspond to the process of generating an intermediate category. Steps 4 and 5 correspond to the second categorization process.

Comparison The algorithm of computing a metaphor vector $\boldsymbol{v}(M)$ by the process of comparison is as follows.

1. Compute a set of k words (i.e., alignments between the target w_T and the source w_S) by finding the smallest i that satisfies $|N_i(w_T) \cap N_i(w_S)| = k$.
2. Compute a metaphor vector $\boldsymbol{v}(M)$ as the centroid of $\boldsymbol{v}(w_T)$ and k vectors computed at Step 1.

This algorithm is proposed by Utsumi (2006). Step 1 corresponds to the initial alignment process, while Step 2 corresponds to the later projection process.

Besides these three models, for comparison purposes, we also consider a simple combination algorithm by which a metaphor vector $\boldsymbol{v}(M)$ is computed as the centroid of the target vector $\boldsymbol{v}(w_T)$ and the source vector $\boldsymbol{v}(w_S)$.

Simulation Experiment

Method

Human experiment For human interpretation of adjective metaphors, we used the result of the psychological experiment reported in Sakamoto and Sano (2004). The materials used in the experiment were 50 Japanese adjective metaphors. They were created from all possible adjective-noun combinations of five adjectives (red ["*akai*"], blue ["*aoi*"], yellow ["*kiioi*"], white ["*shiroi*"], black ["*kuroi*"]) with 10 nouns (voice ["*koe*"], sound ["*oto*"], mind ["*kokoro*"], feeling ["*kimochi*"], words ["*kotoba*"], atmosphere ["*funiki*"], character ["*seikaku*"], past ["*kako*"], future ["*mirai*"], taste ["*aji*"]).

Thirty-eight undergraduate students of the University of Electro-Communications, who were all native speakers of Japanese, were assigned to all the 50 metaphors. They were asked to choose among 24 perceptual adjectives (i.e., features) appropriate ones for the meaning of each adjective metaphor. For each chosen feature w_i of an adjective metaphor M, the degree of salience $sal(w_i, M)$ is then assessed as the number of participants who chose that adjective. These features were used as landmarks with respect to which model's interpretation and human interpretation were compared for evaluation. Note that any adjective chosen by only one participant was not included in the analysis. For example, as shown in the bar graph of Figure 1, seven adjectives were chosen for the metaphor "black future", and the adjective *dark* had the highest salience, i.e., the number of participants (26 participants) who listed it was largest.

Computer simulation The semantic space used in the simulation experiment was constructed from a Japanese corpus of 251,287 paragraphs containing 53,512 different words, which came from a CD-ROM of Mainichi newspaper articles (4 months) published in 1999. The dimension l of the semantic space was set to 300, and thus all words were represented as 300-dimensional vectors.

In the computer simulation, for each of the 50 adjective metaphors, four kinds of metaphor vectors were computed using the four comprehension algorithms presented in the preceding section, i.e., categorization, two-stage categorization, comparison and simple combination. In computing the metaphor vectors, we varied the parameter m_1 in steps of 50 between 50 and 500, and the parameters k and m_2 from 1 to 10. After that, for all the features $w_i, \cdots, w_n$ chosen for a metaphor M in the human experiment, similarity to the metaphorical meaning $sim(w_i, M)$ was computed separately using the four metaphor vectors. Features with higher similarity to the metaphorical meaning can be seen as more relevant to the interpretation of the metaphor. In Figure 1, for example, the word *dark* has the highest similarity to both the metaphor vectors computed by the categorization algorithm and by the two-stage categorization algorithm, but a least salient word *calm* is also highly similar to the metaphor vectors.

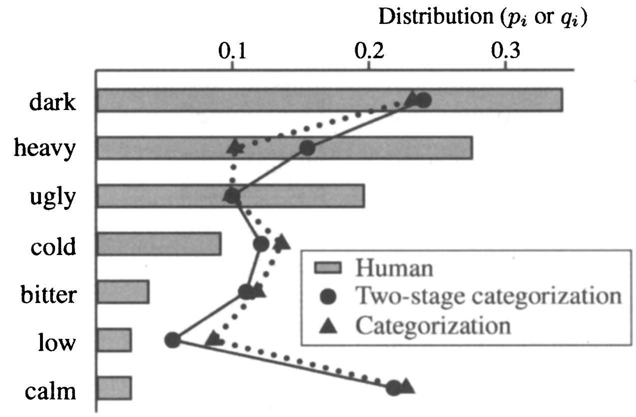


Figure 1: "Black future" metaphor

Evaluation measures To evaluate the ability of the model to mimic human interpretations, we use the following measures, which were also used in Utsumi's (2006) simulation experiment for nominal metaphors.

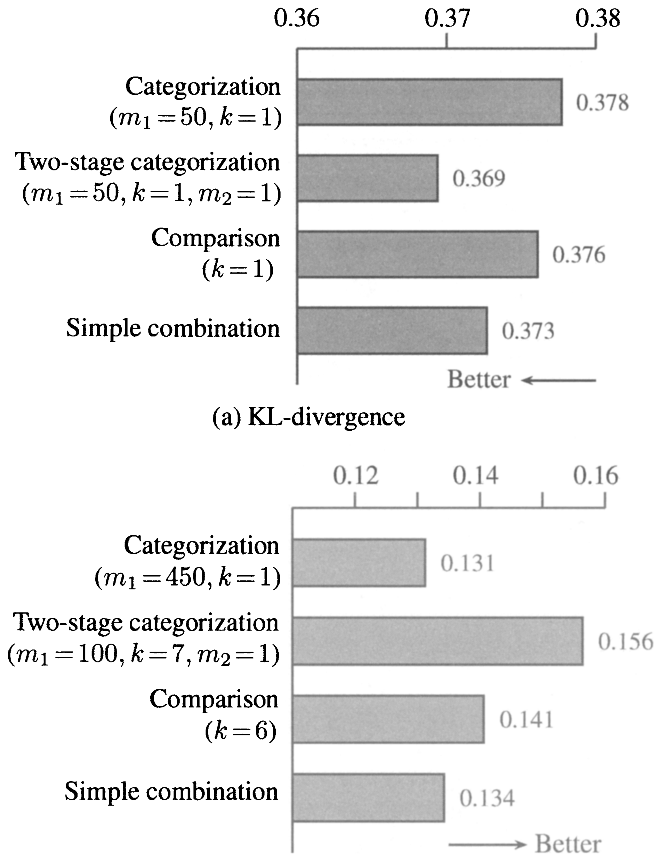
- *Kullback-Leibler divergence (KL-divergence)*:

$$D = \sum_{i=1}^{n} p_i \log \frac{p_i}{q_i} \quad (1)$$

$$p_i = \frac{sal(w_i, M)}{\sum_{j=1}^{n} sal(w_j, M)} \quad (2)$$

$$q_i = \frac{sim(w_i, M) - \min_x sim(x, M)}{\sum_{j=1}^{n} \{sim(w_j, M) - \min_x sim(x, M)\}} \quad (3)$$

It measures how well a model simulates the salience distribution of features relevant to human interpretation, or in other words, the degree of dissimilarity between human interpretation p_i and computer's interpretation q_i. Hence, *lower divergence means that the model achieves better performance.* In Figure 1, for example, KL-divergence between the salience distribution of human interpretation and the similarity distribution of computer interpretation is 0.546 for the categorization model ($m_1 = 50$, $k = 1$) and 0.396 for the two-stage categorization model ($m_1 = 50$, $k = 1$, $m_2 = 1$). This result suggests that, in this case, the two-stage categorization model better mimics human interpretation than the original categorization model.



(a) KL-divergence

(b) Rank correlation

Figure 2: Simulation results: Comparison among the four comprehension models for adjective metaphors

- *Spearman's rank correlation*:

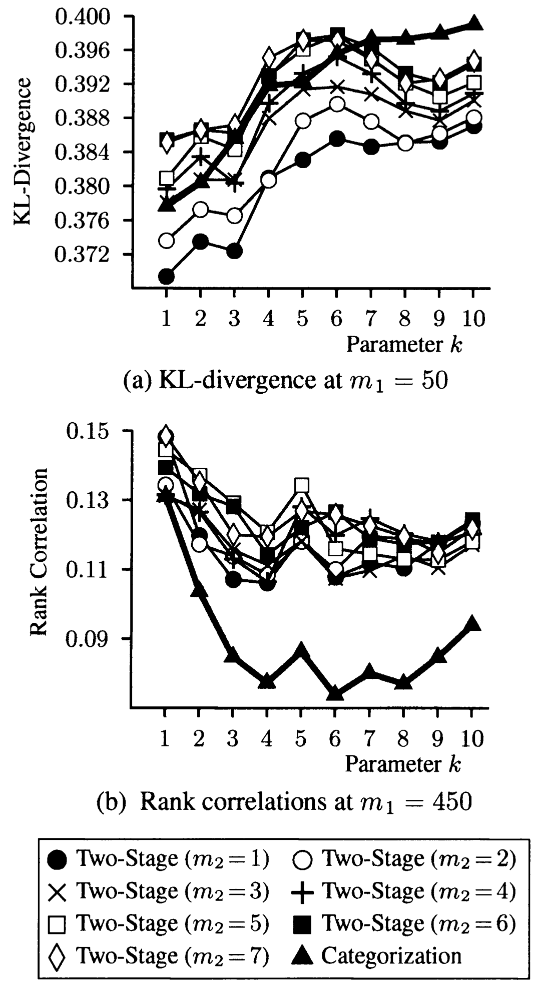
$$r = 1 - \frac{6 \sum_{i=1}^{n} d_i^2}{n^3 - n} \tag{4}$$

$$d_i = rank(sim(w_i, M)) - rank(sal(w_i, M)) \tag{5}$$

It measures how strongly the computed similarity of relevant features is correlated with the degree of salience of those features. *A higher correlation means that the model yields better performance.* In Figure 1 the two-stage categorization model yields a higher correlation ($r = .46$) than the categorization model ($r = .28$), which again indicates that the two-stage categorization model is superior to the categorization model.

Result

For each of the 50 metaphors, KL-divergences and rank correlations were computed using the four metaphor vectors. These values were then averaged across metaphors. Concerning KL-divergence, the categorization algorithm achieved the best performance when $m_1 = 50$ and $k = 1$, the two-stage categorization model did the best performance when $m_1 = 50$, $k = 1$ and $m_2 = 1$, and the comparison model did the best performance when $k = 1$. Concerning rank correlation, the combination of $m_1 = 450$ and $k = 1$ was optimal for the



(a) KL-divergence at $m_1 = 50$

(b) Rank correlations at $m_1 = 450$

Figure 3: Simulation results of the two-stage categorization model and the categorization model obtained with various values of parameters k and m_2

categorization model, while the combination of $m_1 = 100$, $k = 7$ and $m_2 = 1$ was optimal for the two-stage categorization model. For the comparison model, $k = 6$ was optimal.

Figure 2 shows mean divergences and correlations calculated using these optimal parameters. The two-stage categorization model outperformed the other three models on both measures. It suggests that *the two-stage categorization theory is the most plausible theory of adjective metaphor comprehension.* Furthermore, in order to demonstrate that this simulation result in favor of the two-stage categorization theory is general, not specific to the particular value of the parameters, we show the simulation results obtained with various values of parameters in Figure 3. Figure 3(a) shows that, when they were compared at the same value of k, the two-stage categorization algorithm had lower divergence (i.e., better performance) than the categorization algorithm at almost all the values of m_2, although it had worse performance at some higher values of m_2 and lower values of k. Similarly, as shown in Figure 3(b), the two-stage categorization algorithm achieved a higher correlation (i.e., better performance) regardless of values of m_2. These results clearly indicate the plausibility of the two-stage categorization model as a cognitive theory of adjective metaphor comprehension.

Discussion

Related Work

Until now there have been some computational studies on metaphor comprehension. For nominal metaphors, Thomas and Mareschal (2001) proposed a connectionist implementation of comprehending nominal metaphors on the basis of the categorization theory, but they did not test the validity of their models in a systematic way, nor did they make a new contribution to the psychological or cognitive theory of metaphor. Kintsch (2000) proposes an LSA-based computational model of metaphor comprehension. His predication algorithm is also used in this study as a model of categorization, but he did not test its psychological validity as a model of metaphor comprehension. In addition, his study does not allow for the fact that some metaphors are comprehended as comparisons. Lemaire and Bianco (2003) also employ LSA to develop a computational model of referential metaphor comprehension. However, they do not address how well it mimics human interpretations; they only showed that it mimics processing time difference between when supporting context is provided and when it is not provided. Moreover, their model is theoretically less well motivated. For adjectival metaphors, Weber (1991) proposed a connectionist model of adjectival metaphors, which can be seen as one computational implementation of the categorization theory. This model uses two methods (direct value transference and scalar correspondence) for establishing semantic correspondences between the properties literally expressed by the adjective and the properties to be mapped onto the target concept. However, her model was not tested in a systematic way, either.

In contrast, our LSA-based computational methodology used in this study tests the validity of competing metaphor theories and predicts which is most plausible. Utsumi (2006) has applied this methodology to nominal metaphors and demonstrated that the interpretive diversity view of metaphor (Utsumi, 2007; Utsumi & Kuwabara, 2005) best explains the mechanism of nominal metaphor comprehension.

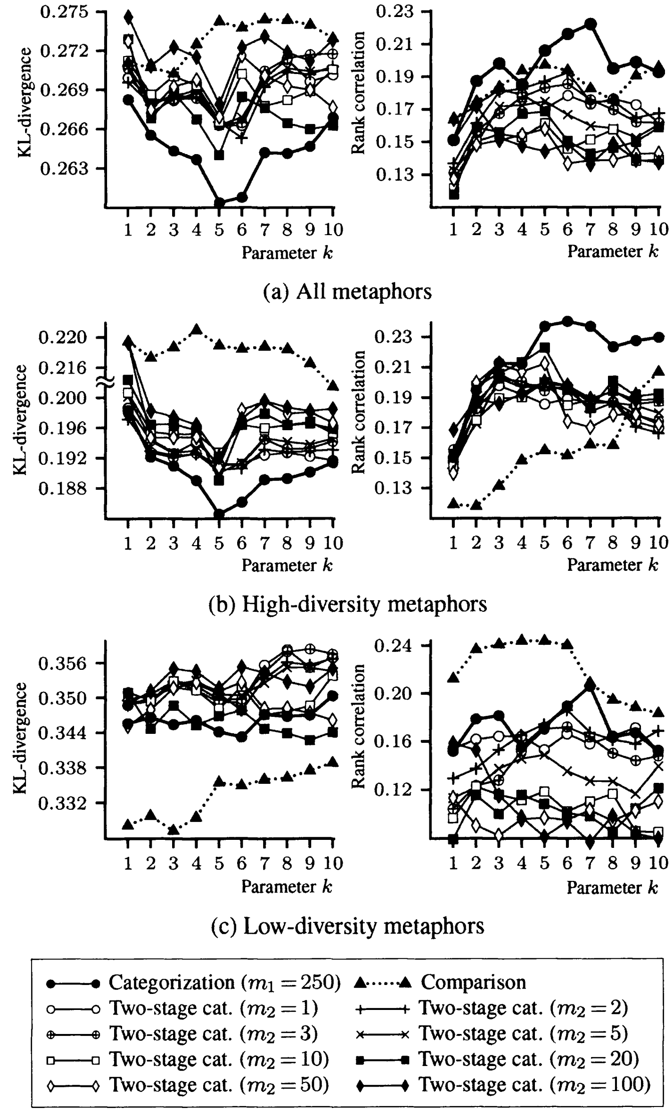


Figure 4: Simulation results of nominal metaphor comprehension ($m_1 = 250$)

Does Two-Stage Categorization Better Explain Nominal Metaphor Comprehension?

In this paper, we have shown that adjective metaphors are comprehended via a two-stage categorization process, rather than via a categorization process or a comparison process. This raises a new interesting question whether or not people also comprehend other types of metaphors, especially nominal metaphors, via a two-stage categorization process.

Recent studies have claimed that people comprehend nominal metaphors as categorizations or comparisons depending on a metaphor property such as vehicle conventionality (Bowdle & Gentner, 2005), metaphor aptness (Jones & Estes, 2006) or interpretive diversity (Utsumi, 2007; Utsumi & Kuwabara, 2005). Especially Utsumi (2007) has demonstrated through a psychological experiment that interpretively diverse metaphors are processed as categorizations but less diverse metaphors are processed as comparisons. Utsumi (2006) also confirmed this finding by means of computer simulation. Therefore, the question mentioned above can be refined as follows: Does the two-stage categorization process better explain comprehension of high-diversity metaphors than the categorization process, and comprehension of low-diversity metaphors than the comparison process?

In order to tackle this question, we conducted an additional simulation experiment in which the metaphorical meanings of 40 nominal metaphors such as "Life is a game" were computed by the two-stage categorization algorithm, and the results were compared with the results of the categorization algorithm and the comparison algorithm obtained in our preceding study (Utsumi, 2006). The simulation method and evaluation measures used in this additional experiment were identical to those used in the main simulation experiment of this study. For human interpretation of the nominal metaphors, the result obtained in a psychological experiment (Utsumi, 2005) was used. (For further details of the simulation experiment of nominal metaphors, see Utsumi, 2006).

The overall result was that the two-stage categorization al-