Talking Nets:
exploring a multi-agent, connectionist model of distributed communication and cognition

Research Proposal for a "Geconcerteerde Onderzoeksactie",

submitted by

Francis Heylighen and Frank Van Overwalle,

Vrije Universiteit Brussel

Summary

This research project is proposed by a multidisciplinary team led by F. **Heylighen** and F. **Van Overwalle**. Its members have expertise in cognitive science, psychology, AI, philosophy, economics, political science, and linguistics, and advanced research experience in connectionist simulation, multi-agent systems, computer-mediated communication and group experiments.

The project aims to further test and develop our new "Talking Nets" simulation model of distributed cognition published in Psychological Review [**Van Overwalle** & **Heylighen**, 2006]. Distributed cognition concerns the information processing and learning that occurs on the social level, by the propagation of information from agent to agent across a social network. Our simulation has accurately replicated the results of half a dozen key experiments in social psychology, illustrating such effects like persuasion, polarization and the reinforcement of stereotypes during group communication. The simulation models individuals as recurrent connectionist networks who are themselves connected by a social network of trust-based links. Both individual and social networks learn from experience, by strenghtening links whose expectations are confirmed.

While the model is relatively simple and apparently powerful in predicting known effects, its assumptions and implications are broad, abstract and complex, requiring extensive further testing and exploration. We propose to investigate the remaining issues using four complementary methodologies: 1) multi-agent simulations, 2) individual psychological experiments, 3) group experiments, 4) development and monitoring of Internet-based social software to support distributed cognition. These methods should clarify at least the following issues: What is the precise role of trust in communication? Which social structures emerge from trust-based links? Which information is preferentially propagated across such network? Under which circumstances is the eventual consensual belief better ("collective intelligence") or worse ("groupthink") than the individual beliefs?

Table of Contents

Preface 3
1. Presentation of the research team 4
1.1. Collaboration between the groups 4
1.2. Members of the team 4
1.3. Short biographies of the team members 5
2. Activities and achievements of the research team 9
2.1. Previous research 9
2.2. Teaching 10
2.3. Phd's delivered 10
2.4. Organization of conferences 10
2.5. Contacts and collaborations 11
3. The research subject: background and state-of-the-art 12
3.1. Investigating distributed cognition 12
3.2. Cognition at the social level 13
3.3. Psychological studies of group cognition 13
3.4. Multi-agent simulations 14
4. A novel theoretical perspective: the "talking nets" model 15
4.1. A connectionist perspective on individual cognition 15
4.2. Communication between individual cognitive nets 17
4.3 Memetics: which information is propagated? 19
4.4 Research questions and hypotheses. 20
4.4.1 Further development of theory and simulations 20
4.4.2 Individual aspects of distributed cognition and trust 21
4.4.3 Social structures and group communication 22
4.4.4 Internet-based experiments and applications 22
5. Methodology and investigations 23
5.1 Development of the theoretical model and simulation 23
5.1.1 Replicating existing and novel empirical findings 23
5.1.2 Large scale network structure 24
5.2 Investigations on trust and information intake in individuals 25
5.2.1 Antecedents of trust 25
5.2.2 Manipulation of trust and some of its consequences 26
5.2.3 The automaticity of trust and novelty 26
5.3 investigations on group communication 27
5.3.1 Consequences of trust in groups 28
5.3.2 Novel concepts emerge 28
5.3.3 Nformation spreading in groups and stereotype change 29
5.4 Internet-based experiments and applications 30
5.4.1 Measuring the collective intelligence of propagation networks 31
5.4.2 Monitoring the automatic propagation of messages 32
6. Deliverables and expected results 34
6.1 Publications 34
6.2 Simulation environments 34
6.3 Empirical data 34
6.4 Workshops, conferences and lectures 35
6.5 Social relevance and potential applications 35
7. Project planning 36
Requested funding 38
Relevant publications of the research team 39
Bibliography (publications by others) 44

Preface

The following is a research proposal, submitted to the GOA ("Geconcerteerde Onderzoeksacties") funding program of the Vrije Universiteit Brussel (VUB), in the category social sciences and humanities ("humane wetenschappen"). A GOA project is intended to further support and stimulate the activities of an experienced VUB research group, lifting it to the level of a "center of excellence". As such, the funding can be seen as complementary to the funding that the team already has collected for individual collaborators and more specific projects, allowing it to fill in the gaps and develop a more long-term, integrated research strategy.

In practice, if the proposal is accepted, the available GOA funding would allow us to employ 4 additional researchers over a 5-year period (2007-2011). These would join a team of some 10 people who are already funded, e.g. by the VUB as professors, teaching assistants or research assistants, and by outside sources such as the Flemish government (FWO), European Commission, or US government. As such, it would help us to grow to an "excellence" level by providing the necessary stability, security and continuity, making us less dependent on smaller-scale and shorter-term contracts, and giving us the resources to develop a more ambitious and in-depth research strategy.

Francis **Heylighen** & Frank **Van Overwalle** Brussels, April 2006

1. Presentation of the research team

In the following, the names of the members of the team are printed in **bold** for easy reference.

The present proposal is an initiative of Francis **Heylighen** and Frank **Van Overwalle**, supported by their PhD students, PostDocs and research assistants. As such, the research team is a collaboration between two groups, the interdisciplinary *Evolution, Complexity and Cognition* group (ECCO), led by **Heylighen**, which is affiliated with the Center Leo Apostel (CLEA), and the *Social Cognition Lab* (SCL), led by **Van Overwalle**, which is part of the *Personality and Social Psychology* department (PESP). The two groups have built up complementary expertise concerning distributed cognition: ECCO mostly in transdisciplinary, theoretical frameworks, multi-agent simulations and web applications; SCL in experimental psychology and connectionist simulations.

For an overview of their activities, check their respective websites:

- http://ecco.vub.ac.be/
- http://www.vub.ac.be/PESP/VanOverwalle.html.

1.1. Collaboration between the groups

ECCO and SCL have been collaborating since 1990 on the dynamics of cognition, with special focus on causal attribution, connectionist learning, and the development of shared or collective knowledge. This resulted in several co-authored publications, including [**Van Overwalle** & **Heylighen**, 1991, 1995; **Van Overwalle**, **Heylighen** et al., 1992; **Bollen**, **Heylighen** & **Van Rooy**, 1998]. In the last few years, this work has led to the formulation of a "connectionist" framework for the understanding of distributed cognition [**Heylighen**, Heath & **Van Overwalle**, 2004]. Part of this framework was implemented as a simulation environment that allowed us to reconstruct several key experiments in social psychology [**Van Overwalle** & **Heylighen**, 2006]. The present proposal wishes to further explore and test different implications of this theoretical model.

Heylighen and **Van Overwalle** were moreover co-promotors of a PhD dissertation [**Bollen**, 2001], and have jointly submitted several research proposals, including the following :

- Evolutionary Construction of Knowledge Systems (main promotor F. **Heylighen**, funded FWO 1994-1999)
- The Social Construction of Shared Concepts: empirical study and computer simulation of a distributed cognitive process (main promotor F. **Heylighen**, funded FWO 2004-2007)

1.2. Members of the team

The following is a list of the members of the two groups whose contributions and expertise we will draw on to realize this project, approximately in the order of when they first joined the team. Individual members are likely to take part in those aspect of the research project that best matches their interest and experience (as summarized in the next section with biographies), while the overall approach and strategy will be formulated and supervised by **Heylighen** and **Van Overwalle**.

Presently funded:

- Prof. Dr. Francis **Heylighen** (promotor, ECCO)
- Prof. Dr. Frank **Van Overwalle** (co-promotor, SCL)
- Dr. Bertin **Martens** (researcher European Commission, ECCO)
- Carlos **Gershenson** (researcher FWO, ECCO)
- Bert **Timmermans** (researcher FWO, SCL)
- Marijke **Van Duynslaeger** (researcher FWO, SCL)
- Klaas **Chielens** (researcher OZR, ECCO)
- Mixel **Kiemen** (researcher DISC, ECCO)
- Marko **Rodriguez** (researcher GAANN, ECCO)
- Tanguy **Coenen** (teaching assistant VUB, MOSI/ECCO)

Looking for funding:

- Erden **Göktepe** (PhD student, ECCO)
- Dirk **Bollen** (PhD student, ECCO)
- Clément **Vidal** (PhD student, ECCO)

Former members:

These are group members who have left our local center at the VUB, but with whom we remain in contact and who are likely to still contribute to certain parts of the project.

- Dr. Johan **Bollen** (assistant professor, Computer Science Dept., Old Dominion University / Researcher, Digital Library Research group, Los Alamos National Laboratory (USA), ECCO)
- Dr. Dirk **Van Rooy** (lecturer, Social Cognitive Neuroscience, Birmingham University (UK), SCL)
- Dr. Tim **Vanhoomissen** (lecturer, Psychology Dept., Lessius Hogeschool (Belgium), SCL)

1.3. Short biographies of the team members

Francis **Heylighen** is a research professor affiliated with the *Department of Philosophy* and the interdisciplinary *Center Leo Apostel,* and director of the *Evolution, Complexity and Cognition* group at the *Vrije Universiteit Brussel*. He has worked during most of his career for the *Fund for Scientific Research-Flanders* (FWO), first as research assistant ("aspirant"), then PostDoc, and finally tenured Senior Research Associate ("onderzoeksleider"). He received his MSc in mathematical physics in 1982, and defended his PhD in 1987, on the cognitive processes and structures underlying physical theories [**Heylighen**, 1990b]. He then shifted his research to the self-organization and evolution of complex, cognitive systems, which he approaches from a cybernetic perspective.

Francis **Heylighen** has authored over 90 scientific publications in a variety of disciplines, including a monograph and four edited books. Since 1990 he is an editor of the *Principia Cybernetica Project*, an international organization devoted to the computer-supported, collaborative development of an interdisciplinary knowledge network. He created (and still administers) the project's website [**Heylighen**, Joslyn & Turchin, 2004] in 1993, as one of the first complex, interactive webs in the world. Since 1996 he chairs the *Global Brain Group*, an international discussion forum reflecting on the emerging information society. He is the present editor-in-chief of the *Journal of Memetics: Evolutionary models of information transmission,* which he co-founded in 1996, and member of the editorial boards of the *Journal of Happiness Studies*, and the journals *Informatica* and *Entropy*.

His work has received a wide and growing international recognition from peers, students and the general public. This is shown by such indicators as the number of references to his work in the *Science Citation Index* (more than 220), in the *Google Scholar* paper database (over a 1000), and on the web (about 70,000 according to www.google.com), in the national and international media (articles about his work have appeared among others in *New Scientist*, *Frankfurter Algemeine Zeitung*, *Die Zeit*, *Le Monde*, the *Washington Post*, and *Knack*). This recognition is confirmed by the invitations he regularly gets to lecture in different countries or to write review articles for leading reference works [e.g. **Heylighen**, 2002; **Heylighen** & Joslyn, 1995, 2001]. He is a Fellow of the *World Academy of Art and Science*, and his biography has been listed in *Who's Who in the World* and other international directories.

Frank **Van Overwalle** is a full professor affiliated with the *Department of Psychology* at the *Vrije Universiteit Brussel*. He has worked first as research assistant in the VUB department for new media and computer technology in education, then as PostDoc at the *University of California at Los Angeles* (1988-1989), and finally as PostDoc and tenured professor at the VUB psychology department.

He got his MSc in psychology in 1980, and defended his PhD in 1987 on "Causes of success and failure of freshmen at university: An attributional approach", for which he received the *Tobie Jonckheere Award* of the *Belgian Royal Academy of Sciences, Letters and Arts*. He continued to work on attribution and social cognition, and then applied his and others' research to the development of artificial neural network models of social cognition. He has received several grants from his university and the *Fund for Scientific Research-Flanders* in order to test some unique predictions derived from these theoretical proposals. This enabled him to employ several PhD students in his social cognition lab, who generate scientific output either as a PhD or in empirically oriented articles.

Frank **Van Overwalle** has authored some 45 peer-refereed scientific publications, in the domain of social cognition. His recent research focuses on artificial neural network models of various phenomena in the domain of social cognition at large, to demonstrate the common cognitive processes underlying many social findings. The aim is to abolish ad-hoc hypothesis building which is currently very flourishing in social psychology, and to attempt to develop a general cognitive theory encompassing the whole of social psychology, in line with general theories of psychological information processing. This has resulted in a number of publications in top-ranking journals such as *Psychological Review* and *Personality and Social Psychology Review* with an impact score (SSCI) between 3 and 7.

His work is receiving a broad and growing international recognition from peers, as evidenced by some 220 references to his work in the *Web of Science Citation Index*. He is a member of the *Royal Flemish Academy of Art and Science's* committee of Psychology*,* the *American Psychological Association*, and the executive board of the *Belgian Federation of Psychologists* (BFP). He is a past secretary-general and president of the *Belgian Society of Psychology* (BVP), and is in the editorial board of the *European Journal of Social Psychology* and *Psychologica Belgica*.

Bertin **Martens** is an economist with a MSc (1979) from the *Katholieke Universiteit Leuven*. Since 1989 he works at the *European Commission* in Brussels on project design and evaluation, macro-economic modelling and implementation of structural reform programmes. He has combined his professional career with academic research by working part-time and taking sabbaticals to visit research institutes around the world. As such, he held Visiting Fellow positions at the *University of New South Wales*, the *Max Planck Institute for Research into Economic Systems*, *George Mason University,* and *Stanford University*—where he worked for six months with the Nobel Prize winner Douglas North. He focuses on cognitive science approaches to economic development and social change. In 2004, he defended his PhD thesis on the role of distributed knowledge in social and economic evolution, with F. **Heylighen** and M. Despontin as promotors, which was published as a book by Routledge [**Martens**, 2005].

Carlos **Gershenson** is a computer scientist with a BEng (2001) from the *Fundación A. Rosenblueth* in México, and a MSc (2002) from the *School of Cognitive and Computer Sciences* at the *University of Sussex*. He is making a PhD on the design and control of self-organizing systems under the supervision of **Heylighen**. His research interests include distributed cognition, philosophy of mind, complex systems, artificial societies and computer simulation. At the age of 26, he already had published over 25 scientific papers in international proceedings and journals. He is a contributing editor to *Complexity Digest* and Book Review Editor of the journal *Artificial Life*. His research has been covered in the national and international media, including *Nature News*, *Trends*, and *Technology Research News*.

Bert **Timmermans** is a psychologist with a MSc (1998) from the *Vrije Universiteit Brussel* and an additional MSc in Cognitive Sciences (1999) from the *Université Libre de Bruxelles*. He is finishing his PhD under the supervision of **Van Overwalle** on the way summary information is represented and processed in social judgments, and how this can be modelled by a connectionist network. Other fields of interest are implicit learning, neural networks, consciousness, self-consciousness and personality, and artificial intelligence.

Marijke **Van Duynslaeger** studied Clinical Psychology at the *Vrije Universiteit Brussel*. She obtained her MSc in 2002 and an additional MSc in Cognitive Science from the *Université Libre de Bruxelles* in 2003. She is making a PhD under the supervision of **Van Overwalle**, on whether and in what contexts observers spontaneously infer the overt or hidden motives of a person when given information about that person's actions. This research project is funded by the FWO. Her other research interests include attitude formation and persuasive communication.

Marko **Rodriguez** is computer scientist with a BSc in Cognitive Science from the *University of California at San Diego* (2001), and a MSc in Computer Science from the *University of California at Santa Cruz* (2004). He was awarded a GAANN fellowship by the US Department of Education, which allows him to work as a researcher at ECCO. Together with D. Steinbock (Stanford University), he has developed the "particle-flow network" as a general methodology and software environment to model self-organization and distributed cognition. He has applied this to support collective intelligence in decision-making and scientific collaboration, and has written several papers on these topics. Marko is on track to receive his Ph.D in Computer Science from the *University of California at Santa Cruz.* Since summer 2005, he has been working as a visiting researcher on distributed knowledge systems and digital libraries at the *Los Alamos National Laboratory* with J. **Bollen** and C. Joslyn.

Klaas **Chielens** is a linguist with a MA (2003) in Germanic philology from the Vrije Universiteit Brussel. His Master's thesis [**Chielens**, 2002] made an empirical investigation of selection criteria for the spread of memes. He is working towards a PhD under the supervision of **Heylighen** on the same subject, funded by the Vrije Universiteit Brussel. He is the new managing editor of the *Journal of Memetics*, assisting the editor, F. **Heylighen**, with the publishing and refereeing process.

Mixel **Kiemen** is a computer scientist with a MSc in Theoretical Informatics (2003) from the *Vrije Universiteit Brussel*. For his Master's thesis, he built a software agent simulation to investigate the creative process of tool-making. In 2005, he was responsible for the Cartography of Research Actors project of *DISC*, the Brussels center for the knowledge society. He now works for the Knowledge Sharing over Social Software project, focusing on context-aware information technology to support virtual communities.

Tanguy **Coenen** is a commercial engineer from the Vrije Universiteit Brussel (1999) who works as a teaching assistant at the MOSI department of the VUB. He has been doing research on knowledge management systems in organizations, creating a prototype system for expertise location management based on social network analysis. His present research focus is on how knowledge sharing takes place in social networks over computer mediated communication channels, and how this may enhance creativity.

Erden **Göktepe** studied Political Science in *Ankara University* (1996) and the *Université Robert Schuman* (1999), and has an M.A. in International Relations from *Galatasaray University* (2004). He worked as research and teaching assistant at the *International Relations Department, Galatasaray University* in Istanbul, Turkey before joining ECCO in 2005. In his Master's thesis he approached international relations from the point of view of complexity theory and self-organising systems. He is preparing a PhD thesis on the emergence of cognitive actors as a part of complex social evolution in international politics, with F. **Heylighen** and G. Geeraerts as supervisors.

Dirk **Bollen** has a Master's degree in psychology (2003) from the *University of Maastricht*, with specialization in artificial intelligence and cognitive science. He worked as a teaching assistant at the faculty of psychology and guided some robotic workshops for students at the computer science department, University of Maastricht. He is interested in dynamical systems models of situated and embodied cognition, and their applications to the self-organization of multi-agent systems.

2. Activities and achievements of the research team

2.1. Previous research

The Evolution, Complexity and Cognition group has been focusing on the self-organization [**Heylighen**, 2002; **Heylighen** & **Gershenson**, 2003] and evolution [**Heylighen**, **Bollen** & Riegler, 1999] of complex systems that exhibit some form of intelligence or cognition. This includes individuals, groups, societies, and computer networks. The emphasis is on how higher-order, collective organizations emerge from the interactions between individual agents. Much of their research is theoretical, aimed at formulating fundamental principles [**Heylighen**, 1992e] and integrating conceptual frameworks [**Heylighen**, 2000].

However, this work has also led to concrete applications in the design of a self-organizing, "learning" web, that assimilates the implicit knowledge of its users [**Bollen**, 2001; **Bollen** & **Heylighen**, 1996, 1998; **Heylighen** & **Bollen**, 2002], and the representation of knowledge through "bootstrapping" semantic and associative networks [**Heylighen**, 2001a, 2001b]. A related strand of work, on the selection criteria that determine which knowledge is transmitted in a large group [**Heylighen**, 1993; 1997a; 1998] has received partial empirical confirmation from the statistical analysis of linguistic data [**Heylighen** & Dewaele, 2002; **Chielens** & **Heylighen**, 2005].

More recently, models of (collective) cognition and learning were also investigated by means of multi-agent computer simulations [**Gershenson**, 2002a,b, 2003, 2004; **Gershenson** & **Heylighen**, 2004b] and "particle-flow" networks [**Rodriguez**, 2004; **Rodriguez** & Steinbock, 2006, 2004].

The Social Cognition Lab has worked mainly on the following topics:

- causal attribution [**Van Overwalle** & **Heylighen**, 1995; **Van Overwalle**, **Heylighen**, Casaer, & Daniëls, 1992; **Van Overwalle** & **Timmermans**, 2005; **Van Overwalle**, 1989; 1997a, b; 1998],
- implicit and spontaneous learning and inferences [**Timmermans** & Cleeremans, 2000; **Van Overwalle**, 2004; **Van Overwalle**, Drenth & Marsman,1999],
- connectionist modeling of attribution phenomena [**Van Overwalle**,1998, 2003, under revision; **Van Overwalle** & **Van Rooy**, 1998; 2001a, b; **Van Overwalle** & **Timmermans**, 2001, 2005] and
- connectionist modeling of social psychology at large.

The latter topics have led to a series of publications on connectionist models, including two publications in *Psychological Review*: one on group impression formation and biases [**Van Rooy**, **Van Overwalle**, **Vanhoomissen**, Labiouse & French, 2003] and another on multi-agent models of distributed cognition and communication [**Van Overwalle** & **Heylighen**, 2006]. In addition, there were three publications on person impression formation, cognitive dissonance and on attitude formation in *Personality and Social Psychology Review* [**Van Overwalle** & Jordens, 2002; **Van Overwalle** & Labiouse, 2004; **Van Overwalle** & Siebler, 2005]. There is recent empirical work supporting some unique predictions of the connectionist approach on group processes and biases [**Vanhoomissen** & **Van Overwalle**, submitted; **Vanhoomissen**, **Van Overwalle** & De Haan, submitted] and on attitude formation [Jordens & **Van Overwalle**, 2004; 2006].

2.2. Teaching

Frank **Van Overwalle** teaches three introductory and advanced courses on Social Psychology (with emphasis on social cognition), and one on Group Dynamics. These courses are followed by hundreds of students from different social sciences and humanities. He also teaches a course on Social Connectionism as a guest lecturer at the ULB, as part of the Master's program in Cognitive Sciences.

Francis **Heylighen**, as a research professor, only teaches a single course on Complexity and Evolution for some 20 students in philosophy. Next year, he will teach a second course on Cognitive Systems in the new Master's program in philosophy.

Both have been active in the formation of PhD students, from their own and other departments, by organizing and chairing series of seminars and discussions: the “Foundations Lectures” (1996-2001, **Heylighen** with other CLEA members), “CLEA/ECCO seminars” (2002-present, **Heylighen**) and “Boterhammen in de faculteit” (2002-present, **Van Overwalle**). These (mostly interdisciplinary) seminars have always been open to the whole academic community, and have attracted researchers from all the main departments of the university.

2.3. PhD’s delivered

Several researchers have prepared and defended their doctorate within the research team, under the (individual or joint) supervision of **Van Overwalle** and **Heylighen**:

- Dirk **Van Rooy** [2000]
- Johan **Bollen** [2001]
- Bertin **Martens** [2004]
- Tim **Vanhoomissen** [2006]

The other members of the team are expected to defend their PhD within the next few years.

2.4. Organization of conferences

Both **Heylighen** and **Van Overwalle,** with their collaborators, have organized and chaired several international conferences and workshops on topics related to distributed cognition:

- International Symposium and Workshop on “Self-steering and Cognition in Complex Systems” (VUB, May 20-23, 1987). Proceedings: [**Heylighen**, Rosseel & Demeyere, 1990]
- Summer School on “Self-organization of Cognitive Systems” (Rijksuniversiteit Groningen, Netherlands, August 1988)

- 1st Workshop of the Principia Cybernetica Project: computer-supported cooperative development of an evolutionary-systemic philosophy (VUB, Belgium, July 2-5, 1991)
- Symposium “the Principia Cybernetica Project”, as part of the 13th Intern. Congress on Cybernetics (Namur, Belgium, August 1992)
- Symposium “Cybernetic Principles of Knowledge Development”, as part of the 12th European Meeting on Cybernetics and Systems Research, (Vienna, Austria, April 1994)
- Symposium “The Evolution of Complexity,” as part of the international congress “Einstein meets Magritte” (VUB, Belgium, June 1995). Proceedings: [**Heylighen**, **Bollen** & Riegler, 1999]
- 1st Symposium on “Memetics”, as part of the 15th Intern. Congress on Cybernetics (Namur, Belgium, August 1998)
- International Workshop “Classic and Connectionist Approaches to Causal Inference and Social Judgment” (Aix-en-Provence, France, 1999)
- International Workshop “From Intelligent Networks to the Global Brain” (VUB, Belgium, July 3-5, 2001)
- One-day International Workshop on “Trends in Distributed Cognition: towards a formulation of a research agenda” (VUB, July 6, 2002)
- Workshops on “Social psychology in Belgium” (2002 and 2003).
- International Meeting on “Social Connectionism”, (16-19 June 2004, Genval, Belgium)
- Session on “Philosophy and Complexity” at the Complexity, Science & Society Conference (Liverpool, UK, 11-14 Sep., 2005)
- Symposium on “Social Connectionism” at the EAESP conference (Wurzburg, Germany, 26-25 July 2005)

2.5. Contacts and collaborations

Francis **Heylighen** and his students actively take part in several international networks related to collective knowledge development and information transmission: The *Principia Cybernetica Project* develops and manages a knowledge web (administered by **Heylighen**) that contains over 2000 documents, including many papers and complete electronic books, which are consulted some 80 000 times a day by people around the world. The *Global Brain Group*, co-founded and chaired by **Heylighen**, groups most of the important researchers in its domain (the emergence of computer-supported, collective intelligence at a world scale), including V. Turchin, B. Goertzel, J. de Rosnay, G. Stock and C. Joslyn. The group organized the first conference on the domain. **Heylighen** administers its electronic mailing list which is used by over 100 selected contributors to discuss advanced issues. **Heylighen** is also involved as editor-in-chief and founding editorial board member in the *Journal of Memetics: Evolutionary Models of Information Transmission*, where most researchers in the domain publish,.

His group has been closely collaborating for many years with the *Distributed Knowledge Systems and Modelling* team, led by C. Joslyn at *Los Alamos National*

Laboratory, producing several joint publications [e.g. **Heylighen** & Joslyn, 1993, 1995, 2001; Rocha & **Bollen**, 2000]. They also have kept in contact for many years with B. Edmonds in the *Center for Policy Modelling* (Manchester Metropolitan University) and S. Umpleby, director of the *Center for Social and Organizational Learning*, *George Washington University,* and many other researchers in the domain of systems theory and cybernetics. Recently, they have presided over two international consortia of research groups to submit projects for the European NEST programme, one including Nagarjuna G. (Tata Institute, Mumbay), S. Solomon (Hebrew University of Jerusalem) and C. Joslyn, and one involving several key members of the Journal of Memetics board of editors (R. Aunger, P. Marsden, M. de Jong, N. Krasnogor).

At the international level, Frank **Van Overwalle** collaborates with renowned researchers in the area of connectionist modeling of social phenomena, including Eliot Smith (*Purdue University*, USA), Stephen Read (*USC, Los Angeles)*, Yoshi Kashima (*University of Melbourne*, Australia), and Fred Vallée-Tourangeau (*University of Hatfield*, U.K.). He is also a member of a research community of the FWO on "Acquisition and representation of evaluative judgments and emotion". There is also intense collaboration and joint publications with well-known connectionist researchers in other domains of psychology in Belgium, such as at the *Université Libre de Bruxelles* (Axel Cleeremans) and *Université de Liège* (Robert French; Christophe Labiouse).

Locally, within the *Vrije Universiteit Brussel*, our research team maintains and plans to further develop a variety of interdisciplinary contacts, including N. Gontier and J-P. Van Bendegem at the Center for Logic and Philosophy of Science (CLWF) on the evolution of language and the extended mind, E. Verstraeten and E. Soetens of the Cognitive and Physiological Psychology group (COPS) on brain physiology and implicit learning, G. Geeraerts and M. Tezcan of the Political Science Department (POLI) on complex systems models of social interaction, B. Manderick and A. Nowé of the Computational Modelling Lab (COMO) on multi-agent systems, and G. Vancronenburgh, and N. Deschacht of the Economics Department (MOSI) on evolutionary and systems dynamics models of social and economic interaction.

At our sister university, the *Université Libre de Bruxelles*, we plan to stay in touch with T. Lenaerts of the AI-lab (IRIDIA) on the evolution of cooperation, A. Cleeremans of the Cognitive Science Research Unit on connectionist models of cognition, the group around J-L. Deneubourg at the Unit of Social Ecology on animal models of collective intelligence, and O. Klein at the Social Psychology Department on communication and maintenance of stereotypes in groups.

3. The Research Subject: background and state-of-the-art

3.1. Investigating Distributed Cognition

Cognition is not limited to the mind of an individual, but involves interactions with other minds. Sociologists and developmental psychologists have long noted that most of our knowledge of reality is the result of communication and social construction rather than of individual observation. Moreover, complex real-world problems are typically solved by a committee or association of individuals who pool their individual expertise while forming a collective decision that is supposedly better than what they might have achieved individually. This phenomenon may be labeled *distributed* *cognition* [Hutchins, 1995; **Heylighen**, Heath & **Van Overwalle**, 2004]: the processes of knowledge development and problem-solving are distributed across a group of communicating individuals.

To understand such collective information processing, we must investigate the distributed system constituted by different individuals with different forms of knowledge and experience together with the social network that connects them and that determines how information is propagated. In addition to qualitative observations and conceptualizations, the phenomenon of distributed cognition has been studied in a quantitative, operational manner, albeit from two very different traditions: *social psychology* and *multi-agent simulations*.

The present section introduces the different relevant approaches. The next section introduces our recently developed Talking Nets model, which is an attempt to integrate the main findings, together with the as yet unsolved issues that the present proposal wishes to tackle.

3.2. Cognition at the Social Level

The study of cognition—*cognitive science*—is in essence multidisciplinary, integrating insights from approaches such as psychology, philosophy, artificial intelligence (AI), linguistics, anthropology, and neurophysiology. To this list of sciences of the *mind*, we now also must add the disciplines that study *society*. Indeed, an increasing number of approaches are proposing that cognition is not limited to the mind of an individual agent, but involves interactions with other minds.

Sociologists [e.g. Berger & Luckman, 1967] have long noted that most of our knowledge of reality is the result of a social construction rather than of individual observation. Philosophers [e.g. Searle, 1995] and cyberneticians [e.g. **Heylighen** & Joslyn, 2001] have pointed out the profound implications of this observation for epistemology and theory of mind. Economists too have studied the role of knowledge in innovation, diffusion of new products and technologies, and the organization of the market. More recently, they have begun to understand its essential role in social, organizational and economic development [e.g. **Martens**, 1998, 2005; Senge, 1990].

The nascent science of memetics [Aunger, 2001; **Heylighen**, 1998], inspired by evolutionary theory and culture studies, investigates the spread of knowledge from the point of view of the idea or *meme* being communicated between individuals rather than the individual that is doing the communication.

Biologists and computer scientists have built models that demonstrate how collectives of simple agents, such as ant colonies, bee hives, or flocks of birds, can process complex information more effectively than single agents facing the same tasks [Bonabeau et al., 1999]. Building on the tradition of distributed artificial intelligence which studies the interactions and collaborations within a system of intelligent software agents [Weiss, 1999; **Gershenson**, 2001], the subject of collective cognition is now even being investigated mathematically [Crutchfield et al. 2002].

3.3. Psychological studies of Group Cognition

Psychologists have studied cognition at the group level by using laboratory experiments to investigate various biases and shortcomings [e.g. Brauer et al., 2001; Klein et al., 2003; **Van Rooy, Van Overwalle** et al., 2004]. Social psychology has a long record of research on the cognitive processes responsible for the formation of stereotypic impressions about other groups and for the lack of efficiency in group problem-solving.

Research has revealed that we often fall prey to biases and simplistic impressions about other groups, and that many of these distorted representations are emergent properties of our cognitive dynamics. Some of these biased processes are *illusory correlation* or the creation of an unwarranted association between a group and undesirable characteristics, *accentuation* of differences between groups, *subtyping* of deviant members [for a review see **Van Rooy, Van Overwalle, Vanhoomissen** et al., 2003] and the *communication* of stereotypes [e.g., Lyons & Kashima, 2003, for a review see **Van Overwalle** & **Heylighen,** 2005]. Many of these processes have been modeled by connectionist networks, to be discussed shortly [**Van Rooy** et al., 2003; **Van Overwalle** & **Heylighen,** 2005].

With respect to processes within a group, different types of social dynamics lead to a less than optimal performance. These include *conformity* and *polarization* which move a group as a whole towards more extreme opinions [Ebbesen & Bowers, 1974; Mackie & Cooper, 1984; Isenberg, 1986], *groupthink* that leads to unrealistic group decisions [Janis, 1972], the lack of *sharing of unique information* so that intellectual resources of a group are underused [Larson et al., 1996, 1998; Stasser, 1999; Wittenbaum & Bowman, 2003] and the suboptimal use of relevant *information channels* in social networks [Leavitt, 1951; Mackenzie, 1976; Shaw, 1964].

Although such experimental research produces empirical evidence about real people and thus allows us to convincingly confirm or refute particular hypotheses, practical constraints imply that it is usually limited to small groups with minimal structures performing tightly controlled tasks of limited duration. As such, its applications to more complex social systems, where there is no limit on the number of agents or the purpose, timing or intensity of their interactions, remain limited.

3.4. Multi-Agent Simulations

This limitation has been partially overcome by the method of Multi-Agent Systems or Multi-Agent Simulations (MAS), which has been used to model human societies and other large-scale social processes [e.g. Epstein & Axtell, 1996; Nowak, Szamrej & Latané, 1990; **Gershenson**, 2001, 2002a]

In these systems, agents interact in order to reach their individual or group objectives, which may be conflicting. However, the combination of their local, individual actions produces emergent behavior at the collective level. In contrast to experimental psychology, this approach has been used to investigate intrinsically *complex* situations, such as the emergence of cooperation and culture, and the self-organization of the different norms and institutions that govern the interactions between individuals in an economy. It manages to tackle such problems by means of computer simulations in which an unlimited number of software agents interact according to whatever simple or complex protocols the researcher has programmed them to obey.

Although the complexity of situations that can be studied is much greater than in an experimental paradigm, there is no guarantee that the results produced by the simulation say anything meaningful about real human interaction. Moreover, most of the earlier simulations of agent interaction are too rigid and simplistic to be psychologically plausible. The behavior of a typical software agent is strictly rule-based: if a particular condition appears in the agent's environment, then the agent will respond with a particular, preprogrammed action.

More sophisticated cognitive architectures may allow agents to learn, that is, adapt their responses to previous experience, but the learning will typically remain at a symbolic, rule-based level. Perhaps the most crucial limitation of many models is that the individual agents lack their own psychological interpretation and representation of the environment. As Sun (2001, p. 6.) deplored, MAS need "…better understanding and better models of individual cognition".

Another shortcoming of many MAS is their rigid representation of relations between agents. In the most common models based on *cellular automata*, each agent ("automaton") occupies a cell within a geometrical array of cells, typically in the form of a checkerboard. Agents will then interact with all the agents in their geometric neighborhood. This is clearly a very unrealistic representation of the social world, in which individuals interact differentially with other individuals depending on their previous experiences with those others. Other types of simulations, such as *social network* models, are more realistic with respect to human relationships, but still tend to be rigid in the sense that the network topology is typically fixed by the programmer, so that the agents cannot change the strength of their relationship with other agents.

4. A Novel Perspective: the "Talking Nets" model

Over the last 2-3 years, our group has developed a generic model, dubbed "Talking Nets", that addresses many of the shortcomings of earlier multi-agent simulations. A detailed exposition of the simulation model and several of its applications is to appear in the July issue of the top-ranking journal *Psychological Review* [**Van Overwalle** & **Heylighen**, 2006]. We will here summarize the main contributions, and go into more detail about the assumptions, their implications and extensions in the next subsections.

The core idea is to extend our connectionist philosophy of cognition from the individual to the social level. This means that:

1) We model individual agents as recurrent connectionist networks, that process incoming information on the basis of their knowledge resulting from earlier learning. This ensures a realistic model of individual cognition (see section 4.1).

2) We model the social links between agents as a higher-order connectionist network that also learns from experience. This ensures a very flexible, and—as we are about to argue—more realistic representation of social interaction (see section 4.2).

In addition to these theoretical innovations, our model offers an excellent fit with some of the available empirical evidence. Our "Talking Nets" simulation which implements the above principles was able to accurately reproduce the quantitative results of half a dozen key experiments in social psychology that illustrate well-known effects in group problem-solving, communication and persuasion, such as polarization and the reinforcement of stereotypes [**Van Overwalle** & **Heylighen**, 2006]. Thus, our model appears to combine the advantages of multi-agent simulations and psychological experiments, while offering a simple, integrated model of much of distributed cognition and communication.

Still, the model is very new and relatively untested, and most of its underlying assumptions, implications and applications still need to be systematically investigated. The present proposal sets out to do just that.

4.1. A Connectionist Perspective on Individual Cognition

Cognition can be defined as the processing of information by an agent in order to support understanding, decision-making and problem-solving. The cognitive agent uses its *knowledge* to interpret incoming data or stimuli, derive inferences from it, and select actions appropriate to the thus perceived situation and to its internal preferences. This knowledge is in general the result of previous *learning*, i.e. adapting the internal structure responsible for processing the information so as to maximize the quality of the inferred predictions and selected actions, while taking into account the feedback from the environment.

From this "cybernetic" perspective [**Heylighen** & Joslyn, 2001; **Van Overwalle** & **Van Rooy**, 1998; **Van Overwalle** 1998; **Van Overwalle** & Labiouse, 2003], knowledge is not a discrete collection of beliefs, propositions or procedures, but a continuously evolving network of connections between perceptions, interpretations and actions, which allows the agent to anticipate and adapt to changes in its environment [**Heylighen**, 1990]. Thus, the understanding of a situation that knowledge enables can be seen as a form of preparedness, expectation, or anticipation for what may come [Neisser, 1976; **Heylighen**, 2006b]. The anticipatory connections between perceptual, abstract, and motor categories have the general form A → B [**Heylighen**, 2001a]. Examples are *banana → yellow* or *lack of preparation → exam failure*, where the weights of the links reflect a certain probability or intensity of occurrence. Such basic connections underlie not only prediction, but causal attribution or explanation of *B*, given *A* [**Van Overwalle** & **Heylighen**, 1991, 1995; **Van Overwalle**, 2003].

The collection of weighted connections defines an associative or connectionist network [**Heylighen**, 2001a]. Such a network can perform complex inferences by activating the initial or perceived categories, letting the activation propagate recurrently through connected categories, and noting which "inferred" or output categories gather most activation. The connections and their continuously varying weights can be learned through the closely related *Hebbian* [e.g. **Heylighen** & **Bollen**, 2002] algorithm, which reinforces the connection between concepts each time they are co-activated, and *Delta* algorithm [**Van Overwalle**, 1998, 2003; **Van Overwalle** & **Van Rooy**, 1998, 2001a,b], which reinforces connections that match outside activation sequences.

Connectionist networks have proven to provide very flexible and powerful models of cognitive systems [e.g. McLeod et al., 1998; **Van Overwalle** & Labiouse, 2004; **Timmermans** & Cleeremans, 2000]. Their processing is intrinsically parallel and distributed [Rumelhart & McClelland, 1986]. Because of the inherent redundancy, they are much more robust than sequential architectures, surviving destruction of part of their nodes and connections with merely a "graceful" degradation of their performance. These systems are wholly decentralized and self-organizing, eliminating the need for a central executive that deliberatively processes information. Moreover, since activation spreads automatically from the nodes that received the initial stimuli to associated nodes, connectionist networks can complete and generalize patterns. Thus, they can fill in lacking data and infer plausible conclusions on the basis of limited information.

Most importantly, connectionist networks inherently support learning, by means of the continuous adaptation of the connection strengths to the ways in which they are used. Thus, successfully used connections become stronger, making it easier for information to be propagated along them, while connections that are rarely used or whose use led to erroneous results weaken, and eventually disappear.

In the Talking Nets model, the individual processing of information is modeled by a standard recurrent connectionist network that has served us well in the past to model the development of individual impressions, opinions and beliefs [e.g. **Van Overwalle** & Labiouse, 2004; **Van Rooy**, **Van Overwalle, Vanhoomissen**, Labiouse & French, 2003; **Van Overwalle** & Siebler, 2005]. It has the following characteristics:

- In a recurrent architecture, all nodes within an agent are interconnected with all of the other nodes of the same agent. Thus, all nodes send out and receive activation.
- Received information is represented by *external activation*, which is automatically spread among all interconnected nodes within an agent in proportion to the weights of their interconnections. The activation coming from the other nodes within an agent is called the *internal activation*.
- The short-term activations are stored in long-term *weight changes* of the connections; these are driven by the difference between the internal activation received from other nodes in the network and the external activation received from outside sources (the *Delta* algorithm).

4.2. Communication between Individual Cognitive Nets

Our multi-agent approach to social structure involves a "connectionist" perspective, where individual recurrent nets that define each agent in the interaction, are connected in a larger structure of social connections [**Van Overwalle** & **Heylighen**, 2006]. This idea was inspired on the classic notion of a *social network* [Mitchell, 1969; **Rodriguez** & Steinbock, 2004], in which individuals are linked by ties of friendship or acquaintance. But rather than using the vague notion of "acquaintance", we interpret a social link as expressing a measure of (cognitive) *trust*, in the sense of the confidence that one has in the ideas, opinions or expertise of the other. Such trust can vary continuously, taking on values between 1 (complete trust) and 0 (no trust at all) and therefore the network is weighted. Furthermore, trust can be asymmetric (e.g. A trusts B more than B trusts A) and therefore the network is directed. We moreover assume that trust is learned by reinforcement, using a modification of the *Delta* algorithm: the more often A has heard B express opinions that appear truthful to A, the more A will trust B, i.e. expect that the next idea uttered by B will also be credible.

Thus, in the "Talking Nets" approach, the trust in the social network *between* individuals will develop and learn in a manner similar to the associative networks *within* each individual. Communication is implemented by transferring the activation of nodes from "talking" agents to "listening" agents. This is accomplished by activation spreading between agents in much the same way as activation spreading within the mind of a single agent, with the restriction that activation spreading between agents is (a) limited to nodes representing identical attributes and (b) in proportion to the trust connection weights linking the attributes between agents. Because agents can play the role of speaker or listener, the trust connections in the model go in two directions for each agent: *Sending* connections for a speaking agent and *receiving* connections for a listening agent.

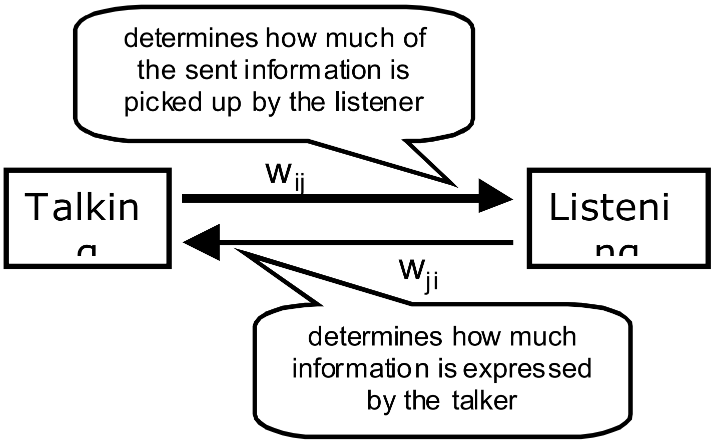

Figure: The role of trust weights in communication between agents.

An important aspect of social communication, canonized by the philosopher Grice [1975] as the maxim of *quality*, is that communication is more effective if the information is (believed to be) *trustworthy*. This is implemented in the *sending* trust connection w_{ij} from an agent i expressing its ideas to the receiving agent j (see Fig.). When trust is maximal (+1), the information expressed by the talking agent is accepted as such by the listening agent. When trust is lower, information processing by the listener is attenuated in proportion to the trust weight. When trust is minimal (0), no information is processed by the listening agent. Thus, the listener sums all information received from talking agents in proportion to the respective trust weights, and then processes this information internally.

In addition, Grice [1975] put forward another maxim of *quantity* which suggests that communicators transmit only information that adds to the audience's knowledge. In addition, research on group minority suggests that communicators tend to increase their interaction with an audience that does not agree with their position. This is implemented in the model by the *receiving* trust weights w_{ji} from the listening agent j to the talking agent i (see Fig.). These weights reflect the *novelty* of the issue and are the result of earlier communications in which the listening agent expressed judgments on an issue that were congruent with the talking agent's knowledge. When these trust weights are high, consensual knowledge on an issue is assumed and the talking agent will refrain from expressing these ideas further. In contrast, when these weights are low, the talking agent tends to express its ideas on this issue more strongly, as if to compensate for the expected attenuation of the message by a skeptical receiver.

Grice's [1975] maxims of quality and quantity, implemented in the Talking Nets model as principles of trustworthiness and novelty, have several important implications. First, information coming from a trusted individual will be assimilated more quickly or easily than information expressed by a non-trusted individual [**Van Overwalle** & **Heylighen**, 2006]. Second, an individual trying to convey a message to an individual who distrusts the sender will need to do additional effort to be persuasive.

At first sight, a Talking Nets approach may not seem a cognitive system because it appears to merely transmit information between individuals, while the intelligent interpretation and manipulation happens within the individuals. But information is not transmitted unchanged, since normally each individual reformulates, selects from and adds to the information s/he receives before passing it on to the others. While these transformations still seem to be limited to the individual networks, the overall effect, in the sense of the collective "decision" or outcome of the group process, cannot be reduced to any individual contributions: it is truly emergent. For example, phenomena such as groupthink or polarization show that groups can often make much more radical decisions than any member would have dared to take individually.

This provides social systems with a form of collective intelligence that is different from the sum of the individual intelligences of its members. Thus, "Talking Nets" can be seen as a medium for distributed cognition [**Heylighen**, Heath & **Van Overwalle**, 2004] or collective decision-making [**Rodriguez** & Steinbock, 2004, 2006] where different beliefs are propagated selectively across the high-trust links in the network, undergoing various shifts and recombinations, possibly coalescing into an overall, "collective" decision.

4.3 Memetics: which information is propagated?

A third way to consider the problem of distributed cognition, after the individual and social perspective, is the cultural or "memetic" perspective, which focuses on the information that is communicated rather than on the communicating agents or their relationships. What information is transmitted not only depends on the architecture of the network or the psychological make-up of its members, but on the content of the information itself.

Before we created the Talking Nets model, we had already developed a list of selection criteria that specify which elements of information are preferentially passed on [**Heylighen**, 1993, 1997a, 1998; **Chielens** & **Heylighen**, 2005], and thus become successful elements of the agent's shared culture. In addition to being intuitively plausible and theoretically justifiable, several of these criteria have been empirically confirmed through psychological experiments [Lyons & Kashima, 2003] and analysis of linguistic data [**Heylighen** & Dewaele, 2002; **Chielens** & **Heylighen**, 2005; Heath, Bell & Sternberg, 2001]. However, this list of criteria has up to now remained more an open-ended, *ad-hoc* collection of features and the critical question now is to what extent the Talking Nets model is compatible with these features. These include the criteria of:

- *utility* (the information appears useful or valuable to the agents)
- *novelty* (the information is not already known or expected)
- *coherence* or *consistency* (the information is consistent with the knowledge that the agents already have)
- *simplicity* (the information is easy to process)
- *authority* or *trust* (the sender is considered to be trustworthy)
- *conformity* or *consensus* (the majority of agents agree on the information)
- *formality* or *explicitness* (the less context or background communicating agents share, the more important it is to express the information explicitly)
- (invariant) *evidence* (there exists a stable, unambiguous source of objective data that confirms the information)
- *publicity* or *repetition* (the same message is encountered again and again)

Our new Talking Nets model explicitly or implicitly explains the origin of these criteria. At the individual network level, the *novelty* criterion is implicit in the connectionist learning algorithm, since unexpected information will have a larger effect on an individual's cognitive structure than already known information. Moreover, spread of activation along existing connections will automatically attenuate messages that are *complex* or *inconsistent* with the knowledge that is already there [**Van Overwalle** & Jordens, 2002]. On the other hand, it will reinforce messages that are encountered more than once (*conformity, repetition, evidence*) or that activate in-built rewards or punishments (*utility*).

At the social level, the principle of *trust* is evident in the role of the social (trust) weights in spreading information to the listening agents, and the principle of *novelty* determines how much the talking agents will express some ideas (see 4.2). In other words, the reinforcement of social links through the increase of trust builds *authority* for the sending agents, while telling them which information the receiving agents are likely to already know, making it less important for them to transmit this information. This allowed our simulation [**Van Overwalle** & **Heylighen**, 2006] to accurately reproduce the results of the [Lyons & Kashima, 2003] experiments that confirmed the importance of consistency, simplicity, conformity and explicitness (or novelty).

However, we haven't as yet systematically investigated the roles of and interdependencies between all these criteria. This requires at least an extension of our connectionist simulation with an in-built *utility* function that may reinforce connections that help to produce rewards or avoid dangers. The inclusion of invariant *evidence* is easier, since this can be modeled by introducing a "passive" agent (e.g. representing a book, or other invariant information source) that repeats always the same message when consulted.

4.4 Research Questions and Hypotheses.

Distributed cognition, the way we have introduced it (and the way it is described in the literature [e.g. Hutchins, 1995; Susi & Ziemke, 2001]), remains a complex phenomenon that touches on a variety of disciplines and approaches, including psychology, sociology, communication, computer simulation and culture. Therefore, to make the problem easier to grasp, we will subdivide our research agenda it into four major aspects: 1) further development of theory and simulation; 2) individual cognition; 3) group communication and the mergence of collective intelligence; 4) propagation of information or "memes" on the internet. We will first formulate the main issues concerning each aspect. In the next section, we outline concrete methods, designs and investigations to tackle these issues, thus defining four subprojects.

4.4.1 Further Development of Theory and Simulations

What kind of social structures are likely to determine or be determined by the trust-based information exchanges that we sketched? Our Talking Nets simulation has already showed that depending on the intensity of the communication, two subgroups characterized by different beliefs may either merge their opinions, or drift apart further [**Van Overwalle** & **Heylighen**, 2006]. Yet, social structure is more complex. There has been relatively little research on the underlying mechanisms or dynamics that determine these structures: why do social networks exhibits these particular properties (e.g. being small-world, yet highly clustered)?

Our Talking Nets simulation and the theory upon which it is based may throw a new light on these questions. In particular, we plan on extending the work by **Van Overwalle** and **Heylighen** [2006] to simulate a larger population of virtual agents reflecting a social community at large. The hypotheses that arise from these extended simulations are then to be tested against (existing or novel) empirical data. In particular, we study the following questions:

- What is the large scale social structure emerging from interactions between a large set of agents? The Talking Net model predicts a sort of self-reinforcing dynamics: agents who share a number of similar beliefs will develop strong trust links, and thus have a stronger consensual influence on each other's other beliefs. The more similar their overall beliefs, the more these agents will interact, and therefore the more they will develop their own ingroup and "culture". On the other hand, they will reduce their interaction with agents whose views are inconsistent with their own, and therefore be unlikely to achieve any consensus with them.
- To what extent can this local "culture" phenomenon form the core of a generalized model of how different subcultures arise within a globalized world in which anybody can in principle communicate with anybody else, but in practice tends to remain within a restricted, community of like-minded people?
- Yet, if we leave open the possibility that an agent may have communications with someone from outside their community, it may happen that their first exchanges largely concern beliefs on which they agree so that they develop a "friendship" link outside their community. Such "cross-community" links may be what turns the network into a small world, rather than just a collection of non-overlapping clusters.

4.4.2 Individual aspects of Distributed Cognition and Trust

According to the Talking Nets approach, to understand social interaction we need to pay special attention to the crucial property of trust, which determines in how far an individual will be inclined to communicate with another individual. From a cognitive point of view, trust is the basic relation of *expectation* that the other agent will react in a certain way with a certain probability. Thus, trust can be seen as the direct extension from the internal, cognitive connections between the concepts within an agent's memory to the external, social connections between agents that allow them to form a distributed cognitive system.

Although there is increasing awareness that trust is an important "core motive" in human interaction [Fiske, 2005] and while its neurological underpinnings are gradually unveiled [King-Casas et al., 2005], there has been little empirical study of trust in social cognition, let alone on its specific role in human communication and information exchange. A trust link in the Talking Nets model is more dynamical than just the assumption that "like attracts like", as it developed over a history of communicative interactions, which may confirm or disconfirm both partner's beliefs. Hence, it is important to set out the particular conditions under which trust increases or decreases and to test these in experiments with humans. We need in particular to investigate:

- How humans perceive trust
- Which factors determine the degree of trust

- To what extent trust is working at the implicit (i.e. automatic) or explicit level.
- How trust influences information exchange (information uptake and expression)

The Talking Nets approach makes specific predictions about the trust dynamics between individuals and how that determines the changes in their beliefs and opinions.

4.4.3 Social Structures and Group Communication

We now move from the individual to the sociological dimension: what kind of social structures are likely to determine or be determined by the trust-based information exchanges that we sketched? This is the issue where we have as yet performed least research. Particularly relevant is how groups coordinate their beliefs and actions when they make collective decisions. We need in particular to investigate:

- During communication, how do individual's beliefs undergo changes due to the pressure of (biased) incoming information from other individuals and prior beliefs by the individual him- or herself?
- How is information propagated in a group? Which information is selected to survive in a group, and which information dies out? The Talking Nets approach makes specific predictions about the dynamics of information spreading in groups.
- What type of information exchange in groups (e.g., parallel, serial and so on) is most efficient, or buffers most against spreading of stereotypical ingroup information, leaving aside and ignoring less desirable countersteretypical information?
- Emergence of collective intelligence: This is perhaps the most novel and difficult aspect of our research. In how far is a distributed cognitive system more than the sum of its components? In how far does it produce a level of understanding, information-processing or intelligence that cannot be achieved at the individual level? The Talking Nets model suggests that trust may increase collective intelligence: as individuals trust each other more, they will more easily assimilate information received from each other, and be less inclined to suppress data that go against their existing beliefs.

4.4.4 Internet-based experiments and applications

Some of the foregoing questions on group communication can be studied with internet-based experiments and applications. There are several reasons for this.

- First, in the internet and similar environments information has been recorded and stored, providing a wealth of data that can be analyzed.
- Second, the internet allows studying more complex social systems, without limit on the number of agents or the purpose, timing or intensity of their interactions.
- Finally, the internet provides specific tools to control and manipulate experimentally the flow of information and communication between groups of individuals

5. Methodology and Investigations

After reviewing the research domain and sketching some of the most important unresolved issues, we will now propose four subprojects or "workpackages" that apply complementary methodologies to tackle these outstanding issues. We have charted four domains or subprojects on which there are a large number of pressing questions that require an answer, and have provided for each of them a relatively detailed program of the general methodology to tackle our research questions, together with a number of more detailed examples of experiments or simulations that appear particularly interesting and serve as basis of further experimenting. However, since this is a long term project, we have on purpose formulated the different simulations and experiments in an open-ended way: at this stage it is difficult to say with any certainty which experiment will turn out to be most fruitful to perform 5 years from now.

5.1 Development of the Theoretical Model and Simulation

Our aim is to extend and run our simulation environment so as to explore a great variety of situations. The general dynamics of our computer modeling is similar to Van Overwalle and Heylighen [2006] and can be briefly described as follows.

A distributed cognitive process is initiated when one or more agents receive one or more pieces of external information. These agents propagate their interpretation to "listening" agents, according to the trust connections. A listening agent will aggregate and process the information it receives from one or more talking agents. It will then pass on its own interpretation to others. These will again transmit their own interpretation of all the information received, to the other agents, and so on. At each transmission stage, the pattern of spreading activation undergoes a transformation determined by the connection pattern within and between agents, during which some information is irreversibly lost, until the connectionist network settles. The state of the network can be seen as the collective interpretation of the externally received information at that stage.

5.1.1 Replicating Existing and Novel Empirical Findings

Using this model, we will try to replicate a number of additional empirical findings from the psychological literature on group persuasion and communication, as we already did in the past [**Van Overwalle** & **Heylighen**, 2006]. Successful replication of the general trends will provide further confirmation of our model; failure to replicate well-known facts will be a strong indication that the model lacks essential features. Some of the specific group phenomena that we want to simulate are:

- Referencing of concrete and abstract objects during communication, and how that affects what we say and how much we talk to describe them [Krauss & Weinheimer, 1964; Schober & Clark, 1989; Steels, 1998]
- Learning of relevant information channels and trustworthy sources in social networks so that available but dormant information is exploited more efficiently [e.g., Leavitt, 1951; Mackenzie, 1976; Shaw, 1964].

In addition to such replication of small-scale, controlled experiments, we will use the simulation environment to investigate the differential effects of large numbers of parameter values, including: number of agents, number of nodes per agent, amount of consistent or inconsistent information that is provided as input to the system, distribution of this information over the agents (e.g. information can be given to one or a few agents at a time, or to the whole group) and over time, and topology of the initial communication network (e.g. random, regular, sequential, parallel, ...).

The specific configurations we will explore will of course depend on the results of existing empirical data or on the empirical data collected in the subsequent subprojects (5.2-5.4): e.g., if one of our experimental results would suggest that collective intelligence is enhanced by a particular communication topology, we will try to replicate that result in the simulation. Vice-versa, if such a result comes out of the simulation, we will configure our experiment(s) to try to reproduce this result. There will be a continuous feedback between simulation runs and real-world experiments, thus combining the advantages of MAS and social psychology approaches.

One central issue we need to explore is the nature of the distributed information processing: which types or aspects of the initially given information tend to get lost in the eventual equilibrium configuration, and how does this depend on the different parameters? Does the selective propagation obey our list of intuitive selection criteria (see 4.3)? And in how far is the quality of the collective "interpretation" reduced by such unavoidable loss of information?

5.1.2 Large Scale Network Structure

Another important issue that is more difficult to study via small-scale experiments is the structure of the emergent social network at the scale of a larger society. This could be investigated by simulation in the following way.

We start with a large population of agents with a variety of partly similar, partly different beliefs. We let the agents interact randomly at first. On the basis of this interaction they will develop trust in those agents whose interpretations and reactions appear most similar to their own. We further assume a preference for trusted agents, so that interactions with distrusted agents gradually diminish, while those with trusted ones become more frequent. This seems to lead to a self-reinforcing dynamics: if two agents happen to already have a number of similar beliefs they will develop strong trust links, and thus interact more. On the other hand, they will reduce their interaction with agents whose views are inconsistent with their own, and therefore be unlikely to achieve any consensus with them.

This is likely to lead to a clustering of agent beliefs around a specific local "culture". This may be a model of how different subcultures arise within a globalized world in which anybody can in principle communicate with anybody else, but in practice tends to remain within a restricted, "virtual community" of like-minded people. Yet, if we leave open the possibility that an agent starts a random communication with someone from outside their community, it may happen that their first exchanges largely concern beliefs on which they agree so that they develop a "friendship" link outside their community. Such "cross-community" links may be what turns the network into a small world, rather than just a collection of non-overlapping clusters.

Other interesting phenomena to study are earlier MAS simulations that were conducted with the more simple models described earlier (see 3.4). It is very likely that some of these simulation outcomes will be replicated by the Talking Nets model. However, we also expect some divergences, that may point to the greater power of our approach. In addition, we want to go beyond the issues that were addressed in the past, to deal with more in-depth problems that these earlier simulations models were unable to address. In any case, these novel simulations need to be tested to see how much they are dependent on the choice of the parameter space and compared against social-psychological and sociological data on a larger scale in order to determine how realistic the results are.

5.2 Investigations on Trust and Information Intake in Individuals

Information exchange between humans is selective, being shaped by the beliefs, needs and goals of individuals and the group. Based on our Talking Nets model [**Van Overwalle** & **Heylighen**, 2006], we hypothesize that information propagation is selective because of inherent features of the act of communication itself. Of the many factors that impact on individual's beliefs, we consider the *trustworthiness* of (the information provided by) the communicators as theoretically most crucial in this interpersonal context.

Several research strategies and designs are possible to explore how humans themselves perceive trust, how trust influences information exchange (information uptake and expression) and what the antecedents are that determine whether trust is high or low. In addition, one can explore to what extent trust is working at the implicit (i.e. automatic) or explicit level. We will discuss each of these approaches in turn.

For each of these approaches we present a number of topics and research designs with multiple experiments/studies that can be conducted, without specifying exactly how many to allow some flexibility in our planning so that further exploration of interesting results is possible. Each of the studies detailed below can be conducted with simple paper-and-pencil experiments or by computer-controlled experiments on approximately 60 participants for each study.

5.2.1 Antecedents of Trust

Which factors increase or decrease trust in information provided by other actors? This is the most crucial test of the Talking Nets connectionist model. Indeed, the model proposes that if no a priori expectations about the sender exist, people will trust information so long as it *fits with their own beliefs*. Although some degree of divergence is tolerated, if the discrepancy is too high, the information will not be trusted and hence not influence people's own belief system. Thus, rather than some internal inconsistency or ambiguity in the story told, it is the *inconsistency* with one's own beliefs that sets off the listener and make him or her distrust the information.

These opposing predictions can be easily tested empirically, by providing information that varies (a) in the degree of internal inconsistency or ambiguity, and (b) in the degree of inconsistency with prior beliefs. This latter aspect can either be manipulated experimentally by providing the participants with background information on the topic, or by measuring their existing beliefs and making up information that is inconsistent with it.

These studies can be conducted with simple paper-and-pencil experiments and questionnaires, on large samples of participants.

5.2.2 Manipulation of Trust and some of its Consequences

Generally, we predict that distrusted information will have much less impact on the participants' beliefs and actions as compared to trusted information. Under more controlled laboratory conditions, we predict that it will take more time to read and understand information from an untrustworthy source, and that it will lead to more temporal activation of distrusted pieces of information, in an attempt to reconcile the information or refute it [e.g., Schul, Mayo & Burnstein, 2004].

There are several ways to increase one's trust or distrust in information provided. One way to manipulate trust is rather *blatantly.* For instance (a) by providing verbal information on the trustworthiness of the information given by some communicators in a group discussion or task, (b) by varying prior *expectations* about the communicators, or (c) by varying whether the communicators are *member of some rival groups* or not, and so on.

More implicit vehicles of trust manipulation might be (d) nonverbal *facial expressions* of disbelief, or other nonverbal expressions of unease and untrustworthiness. The literature on emotions provides as yet little evidence on pancultural facial expressions of distrust, but it seems to us that a facial expression of disbelief might strongly affect people's impression of the trustworthiness of information. Given the crucial role of trust in communication and collaboration, we also expect that even (e) *subliminal* presentation (below awareness) of facial expressions of trust and distrust will determine the impact of the information provided. For instance, we expect that the subliminal presentation of faces expressing disbelief (while an auditory message is presented) might result in less trust of that information, even when participants are unaware of the influence of this contextual factor. In the Social Cognition Lab, we have extensive experience with subliminal presentations outside the focus of vision that are too short to be consciously seen; for instance, by the mere subliminal presentation of happy and sad faces we were able to influence one's self-esteem [**Timmermans**, **Van Duynslaeger**, Vandewaerde & **Van Overwalle**, 2006).

5.2.3 The Automaticity of Trust and Novelty

Our Talking Nets simulations lead us to expect that trust between individuals is developed and applied automatically, outside of consciousness, rather than being a deliberate, controlled process. In contrast, although automatic to some degree, we expect that other criteria such as novelty and (attenuation of) talking about known information (which, according to the Talking Nets model, depends on how much we trust the listener) can be more easily overruled by controlled processes, such as task instructions and goals, since the act of speaking itself is largely within the control of the individual.

To test that the use of trust is automatic, we can make use of an experimental paradigm on spontaneous inferences that the Social Cognition Lab has used before [**Van Overwalle**, Drenth & Marsman, 1999, **Timmermans, Van Duynslaeger**, & **Van Overwalle**, 2006]. In short, in these experiments we compare statements by trusted and distrusted sources, and see to what extent their information is spontaneously integrated in the inferences about the target. For instance, we can provide information implying some trait about the actor (e.g., the sentence "Jane solved the mystery halfway the book" implies that Jane is intelligent), and see to what extent this trait is also spontaneously believed by the receiving individual. We expect that this will be more the case for trusted sources than for distrusted sources, demonstrating that trust is automatically applied.

Similarly, the model assumes that to some degree, speakers will spontaneously refrain from telling information that the listener already knows (the novelty criterion). We can test this by using a paradigm in which spontaneous thoughts on novel versus old story elements are measured in the same manner as above [**Van Overwalle**, Drenth & Marsman, 1999]. For instance, we ask our participants to communicate a specific story to someone else who either does or does not possess the same background information. Immediately after the communication instruction or after telling the story, we can measure how spontaneously participants think about novel information rather than known or consistent information.

5.3 Investigations on Group Communication

Although well-controlled individual experiments are appropriate to study how trust determines information uptake and expression, for observing the propagation of information in the social environment it is essential that we study the processes *between* participants through *group studies* where members of a group work towards a common goal, or observe the same stimuli, after which the individual (private) or group (public) reaction is measured. Particularly relevant to our project is how groups coordinate the concepts and rules they use when collectively tackling a problem. In more mundane terms, this involves the spreading in the group of cherished or stereotypical ideas as opposed to unusual, contradictory and counterstereotypical notions.

Based on the Talking Nets model, an idea that is recurrently communicated will undergo a shift in meaning each time it is assimilated by a new individual, who adds his or her own, unique interpretation and experience to it. Like in a game of Chinese whispers [cf. Lyons & Kashima, 2003], by the time the idea comes back to the individual who initiated it, it may have changed beyond recognition. After several rounds of such passing back and forth between a diverse group of individuals, the dynamical system formed by these propagations is likely to have reached an attractor, that is, an invariant, emergent configuration. In particular, we expect that in general group impressions will become increasingly stereotypical while they are expressed by more communicators as the elements of the message that appear irrelevant, inconsistent, non-conformist, difficult to express, or too complex are filtered out. However, under some circumstances, e.g. when the non-stereotypical information is very novel or valuable, or when it is supported by external evidence (and hence undeniable), we may expect the opposite to occur.

The connectionist Talking Nets model predicts that the development and deployment of trust weights will strongly affect how opinions and beliefs are propagated in the collective. To test hypotheses concerning group interaction, we can focus on minute details of the interaction between participants, extending previous research on the individual's reaction to private stimuli into the realm of group input. Alternatively, we can explore how information may be shared or distorted at the group level, or falls apart in different sub-cultures during the communication or discussion in a group.

In these group studies, the dissemination of information can be tightly controlled by providing information prior to a group discussion or group task, so that its change and communication can be more easily traced and compared later in the process with the information provided initially. Alternatively, one can study how much existing stereotypes, measured individually before the studies, are changed afterwards. And lastly, in line with he Talking Nets model, we can measure how often group members talk and listen to each other, and what they say [see Lyons & Kashima, 2003]. For each of these approaches, we present a number of topics and designs with multiple experiments/studies that can be conducted. Each of the group studies has approximately 40 individuals.

5.3.1 Consequences of Trust in Groups

In extending experiments from the individual level to real groups, the most important issue involves the consequences of trust after information dissemination and discussion in groups (or dyads). This allows us to explore how much of the information is taken into account for generating the individuals' beliefs. As noted earlier, the Talking Nets model predicts that more trust will result in stronger change of beliefs and stronger adoption of collective views and solutions.

There are several ways to explore the extent to which (dis)trusted information influences one's beliefs. This can be measured directly, by *asking* participants at the end of a discussion how much they thought the information provided was useful, how much they privately believe the ideas expressed, how much a consensus was reached (when the task involves a collective decision) or how much they agree with the group solution (when the task is to find a solution to a problem).

More indirectly, in group discussions, one can measure the *time* it took to reach a consensus or a solution in the group, under conditions of trust or distrust. Other related measures are how often there was expression of *disagreement* in the group, opposition or denigration of others, how *sincere* the members of the other groups are seen to be and how much their behavior and messages lead to *attitude change* [e.g., Nadler & Liviatan, 2006]. In addition, another specific prediction made by the model is that if the members of the group trust each other more, there will be more *sharing of information* uniquely held by each member of a group [cf. Stasser, 1999].

Finally we can explore to what degree the quality of the decision (e.g., consensus or shared understanding reached) can be predicted on the basis of *concurrent indices* of trust or other process variables. For this question to be answered, we can videotape on-going discussions. Since groups obviously will differ in the amount of shared understanding or consensus on an issue, we can explore whether the variability in the outcomes of these groups can be predicted by some process variables related to trust (or perhaps totally unrelated to trust) observed in the videotapes.

5.3.2 Novel concepts emerge

To elaborate a key hypothesis of the Talking Nets model—that distributed cognitive systems are able to produce qualitatively new knowledge structures—we use an open-ended observation of real group processes, so that it can give us a better idea of what kind of novelty can actually appear, and which factors stimulate or inhibit this form of social construction or collective creativity.

To allow a quantitative analysis of our observations, we operationalize a concept as a process of categorization, whereby different phenomena are classified as instances of this concept to a greater or lesser degree. A concept can thus be represented as a vector, e.g. (1, 0.7, 0.3, 0), the components of which correspond to the categorization strengths [see **Heylighen**, 2001b; Foltz, 1996]. This allows us to measure individuals' concepts, by asking them in how far they consider a list of possible cases (e.g. *nut,* *apple, berry, tomato, pumpkin*, ...) to be good examples of the concept (e.g. *fruit*). By comparing the results before and after the group discussion, we can numerically estimate the cognitive changes that occurred in the group.

Starting from our general Talking Nets assumptions we expect the following to hold true. As the exchange of information between individuals strengthens consensus, the spread of individual ideas and interpretations among the participants will diminish (i.e. individual concepts will become more similar in the sense of a reduced distance between the vectors):

- We expect that in general elements (vector components) about which there was a relative agreement will be strengthened, while components important to only one or a few individuals are suppressed, or disappear altogether; if some of the members of the group have a higher authority (trustworthiness) than others, their views will carry a proportionately higher weight in the eventual consensus.
- However, if during the discussion a divergent or minority interpretation is produced that scores significantly better on one or more of the other selection criteria (simpler, more useful or novel, supported by more evidence....) this may push the dynamics into a different attractor, strengthening elements (vector components) that didn't have strong values in any of the individual concepts.
- We also expect that the trust in specific people will lead to more directed search and dyadic interaction with these individuals, so that in the end freely generated information channels will outperform *a priori* structured channels or networks (e.g., serial chain, wheels or circles).

A content analysis of the different interventions can provide us with a more qualitative picture of the arguments and factors that have influenced the outcome. The discussion is recorded on videotape, and analyzed for specific factors that appear to have influenced the outcome. The possible reasons why a particular participant has or has not searched for some information and has changed positions are explored by focused interviews.

5.3.3 Information Spreading in Groups and Stereotype Change

More specific predictions at the group level are also possible. We can explore how controlled information may be shared or become distorted (stereotypical) during the communication in a group, or how much stereotyped beliefs are changed after a group discussion and under what conditions. The Talking Nets model makes very specific predictions for several experimental procedures.

In one type of experiment, we study how information is spread in a group and how it affects its communication and the resulting beliefs of the group members. Social psychologists have used two basic paradigms that reflect alternative ways in which information is propagated between people: *parallel* and *serial* communication [e.g. Lyons & Kashima, 2003]. In a parallel communication design, the information is spread from all communicators directly to each participant, like in a group discussion. Thus, the participant has direct access to the observations and impressions of all the people who received the stimuli. In a serial reproduction design, the communication of information is passed sequentially from person to person, like in rumors and gossip. The first communicator in the chain receives the information, memorizes it, and then communicates this information to the second person in the chain, and so on. Here we can investigate how the information changes as it progresses through the chain, depending on factors such as the background knowledge that the participants have.

The Talking Nets model predicts that story elements are shaped in a process of selection, variation and retainment, and are passed along and accepted insofar as they are considered trustworthy and tap beliefs that are common to the individuals or provide novel useful information [see also Heath, Bell & Sternberg, 2001; Stasser, 1999]. This hypothesis predicts that people will be more likely to exchange information that is shared within a group or community. Consequently, communication will become increasingly stereotypical. However, to the point that the participants become aware that these stereotypical pieces of information have already been exchanged, more unique and counterstereotypical information will be expressed in the group (principle of novelty). Because this point of awareness is most quickly reached in parallel communication, unique information is likely to be exchanged earlier in this type of communication than in serial communication where such immediate feedback does not exist.

A complementary way of studying the issue of social structure and stereotyping is to impose different structures (e.g. a two-dimensional array, hub-and-spokes, circle, or everyone communicating with everyone) and then examine what effect this has on the collective cognition: does the general quality of the resulting distributed knowledge worsen (e.g. more polarized or stereotypical) or improve (e.g. more diverse, less stereotyped, more sharing of unique information)?

Simulations with the Talking Nets model point also to other variables that may affect whether and how beliefs are disseminated in parallel discussion groups. For instance, when there is less communication or less initial trust between some members of a group, it is more likely that the group will split in two sub-groups, without ever reaching a consensus. Likewise, when the initial positions diverge too much, or when tolerance of dissenting ideas is reduced, no group consensus or common belief is possible. On the other hand, when information spreading is made more difficult through the use of complex stories or restricting media (e.g., typing instead of writing down), the trust model would predict more distortion in the direction of one's stereotypes due to the noisy input. All these predictions can be tested in such a group experiment.

5.4 Internet-based Experiments and Applications

Over the last few years, a new paradigm has started to emerge for studying social systems: Internet-based social networks or on-line communities (e.g. Friendster, Orkut, LinkedIn). These are groups of people who communicate and get to know each other via the intermediary of a complex software system, which is typically reachable through a web-based interface. Members of such system typically invite their friends and acquaintances to join the same system, and then make links from their personal page in the system to the pages of their friends. This turns the system into a computerized version of a social system, where social links are now formalized as hyperlinks between personal pages. The system moreover provides various facilities (e.g., chat, email, blogs, and discussion forums) for quick and easy communication between the members of the same group, network or community.

Several members of the ECCO team (M. **Rodriguez**, M. **Kiemen**, T. **Coenen**, C. **Gershenson**...) have started to develop and research such systems, in particular via the KNOSOS (Knowledge Sharing over Social Software) project (http://www.knosos.be), and the Smartocracy project (http://smartocracy.net [**Rodriguez**, ..., **Gershenson**, **Bollen** et al., 2006]). This builds further on our experience in developing Principia Cybernetica Web, one of the first complex and interactive websites in the world

[**Heylighen** et al, 1993-2006; **Bollen** & **Heylighen**, 1996; **Bollen**, 2001]. Such systems to some degree combine the advantages of:

- *computer simulations*, since all data about who is linked with whom or who communicated with whom are automatically stored for easy analysis and the rules for communicating and linking can to some degree be manipulated by reprogramming the software,
- *group experiments*, since the "agents" are real people who willingly participate in the experiment.

This means that the same questions that we plan to investigate via laboratory experiments could in principle also be investigated using such systems. The main difference is that interaction is not direct like in face-to-face conversation, but happens over a (usually written) medium, that leaves a trace of the information that is exchanged. According to our working hypothesis, this should increase the possibilities for collective intelligence, since there is less danger of unique information being lost. Moreover, computer-mediated communication facilitates parallel communication, since there is no requirement for a single person to "speak" at a given time: different contributors can enter their own ideas and opinions into the system at the time they choose. On the other hand, the lack of face-to-face communication may make it more difficult for participants to build up a relationship of trust with someone they never actually met.

Therefore, it is worth trying to replicate our major group experiments over an on-line system, to see in how what respect the results would differ. This is less obvious for the information spreading experiments (see 5.3.3), though, since email allows the literal transmission of information, without any change, so that we cannot study a gradual transformation of a story like in the [Lyons & Kashima, 2003] experiment.

On the other hand, the emergence of novel group knowledge and solutions is easy to observe by using a "discussion forum" type of environment, in which the different participants write their arguments and counter-arguments in a common space, reacting directly or indirectly to previously made points. In addition to this shared information space, we may also allow "private" exchanges between participants via email, perhaps criticizing "outgroup" members among the ingroup, or forming coalitions. If our hypothesis is correct, the presence of a shared record should improve the quality of the outcome. On the other hand, it is not obvious whether the additional possibility of private exchanges will enhance collective intelligence (by allowing more strongly trust-based channels) or produce polarization (by allowing people with similar ideas to reinforce each other's beliefs independently of the others).

While discussion-forum type systems are already quite well-known, we would also like to investigate two types of social network applications that are as yet in a purely exploratory stage: one for collective decision-making via propagation of votes [**Rodriguez** & Steinbock, 2004, 2006] and one for the propagation of information items between "trusted" individuals [**Rodriguez**, Watkins, Steinbock & **Gershenson**, 2006]. These internet applications grew out of the same theoretical principles on distributed cognition that underly the Talking Nets model.

5.4.1 Measuring the Collective Intelligence of Propagation Networks

The simplest way to measure (individual) intelligence is by means of an IQ test, where the subject needs to select the correct solution out of given potential answers to a series of problems. A similar procedure could be applied to measure the quality of collective decision-making, thus proposing a possible test for a *collective IQ*. The idea is to collect a series of multiple choice questions with unambiguous, known answers, but that are difficult and diverse enough so that no single individual can be expected to answer all of them correctly. Such questions could for example be extracted from high level IQ tests as well as "general culture" quizzes that require advanced expertise. The score of a group of individuals answering the questions would then constitute a measure of the collective intelligence of that group. To speak meaningfully about "collective intelligence", the group should score significantly better than the average individual, and preferably better than the best individual [cf. Johnson, 1998; Kaplan, 2001].

[**Rodriguez** & Steinbock, 2004, 2006] have proposed a generic social-network based framework closely related to our Talking Nets model, in which "votes" can propagate recursively along the links between individuals, proportional to the degree of trust, until they reach an individual who is willing to express an opinion on the problem at hand. The vote of that individual is then weighted by the total number of directly or indirectly accumulated "votes of confidence" that have propagated through the network. This avoids the problem of individuals having to vote on an issue for which they do not have the expertise: for each issue, only those individuals vote who feel motivated or self-confident enough to make a decision, yet their vote carries the weight of all others they directly or indirectly represent. The underlying "particle flow" propagation algorithm, which is a discrete version of spreading activation, has been shown to perform well for simple simulations [**Rodriguez** & Steinbock, 2004], but hasn't yet been tested in real-world decision-making. The algorithm is also flexible enough to encompass other collective decision-making methods such as direct democracry, representative democracy and "smartocracy" [**Rodriguez**, Steinbock, Watkins,**Gershenson**, **Bollen** et al., 2006] as special cases.

We plan to test the potential power of collective decision-making by designing a web-based "Collective IQ" test, gather a community of people around it that know each other well enough to form a social network (as it already exists in embryonic form in the smartocracy.net website), and then let them vote on each other's perceived expertise and on different proposed answers to the questions. We can then test out different network representations, voting procedures and propagation algorithms to see which one produces the highest overall score. Moreover, we can measure in how far the collective IQ score depends on factors such as the total number of participants, the diversity in their backgrounds or opinions, or the strength of the trust links connecting them. This, together with the rest of our empirical and theoretical research, should allow us to formulate guidelines on how to optimize collective decision-making.

5.4.2 Monitoring the Automatic Propagation of Messages

The "Snurf" project formulated by [**Rodriguez**, Watkins, Steinbock & **Gershenson**, 2006] proposes to build a similar social network-based software system to propagate information between individuals. The idea is that participants link to the other participants that they trust, using weighted and labeled links (e.g. John may trust Paul's opinion about politics, but trust Jane more on matters of science).

Whenever participants encounter or create a piece of information that appears interesting (e.g., a news item, a paper, an email, a link to a website), they label this information and give it an interest score. The information is then automatically propagated to all others who trust the judgment of the sender in the domain of the label, with a degree of priority proportional to the interest score and the strength of the trust link. The receivers in turn can deem the item interesting, in which case it is propagated one (or maximum three) step further to the people that trust them.

To avoid being buried under an avalanche of messages, a single individual only receives a single copy of any item, however many trusted people may have forwarded it—although the priority can be increased with each additional person confirming the value of that item. If receivers do not like the information coming from a particular sender, they may reduce their trust in that person, so that messages from that sender get a lower priority, or eventually are no longer transmitted to them anymore. (This can be implemented by an automatic trust learning algorithm like in the Talking Nets simulation). Vice-versa, a person systematically propagating truly interesting information will see her trust score increase.

This appears like a straightforward application of the principles underlying our Talking Nets model to the automatic distribution of information over the Internet. It seems to have at least two major advantages over traditional ways of news distribution (mass media such as magazines or news services, and individual emails):

1) truly interesting messages are likely to propagate far and wide within their community of interest, independently of how general or specialized those interests are, making it unlikely that you miss out on anything important in your domain;

2) because of the automatic prioritizing, there is little danger of information overload by an avalanche of partly relevant messages: the least relevant messages are automatically relegated to the bottom of the heap, where they are unlikely to be read and therefore to be propagated further.

However, to really ascertain the value of such a service, it will need to be implemented and tested out with a sufficiently large community of interested people. Given the present state of web technology, the development of such a system seems relatively easy [**Rodriguez** et al., 2006]. Our experience with the Principia Cybernetica, KNOSOS and Smartocracy web communities moreover has convinced us that it is easy to gather a critical mass of volunteers willing to try out such a system, especially if the system is easy to use and offers a novel functionality.

Once the system is up and running, we can start monitoring its functioning, measuring variables such as the number of messages propagated, the number of steps they propagate, the distribution of the trust links, and the involvement levels of the participants. More specifically, we will try to determine the most successful messages, i.e. those reaching the largest audience, and compare them to the least successful. If our model is correct, the success rate should correlate with the messages' score on the different selection criteria such as novelty, consistency, utility, etc. These scores can be established by asking people to evaluate the messages on these different dimensions, as we have done in earlier research [**Chielens** & **Heylighen**, 2005]. We will also analyze the structure of the trust network, and in particular examine whether some participants emerge as "hubs", or rather "authorities", whose influence extends far and wide. These general results will of course also be compared with those of the simulations and the laboratory experiments.

6. Deliverables and Expected Results

In addition to the novel insights and improved conceptual framework, we expect this project to deliver the following more concrete "products".

6.1 Publications

At the end of the 5 year duration of this project, we expect to have published dozens of papers with the results of our research, both theoretical and empirical, in a variety of international, peer-refereed journals, as well as in proceedings of conferences, and as chapters in books. We plan in particular to submit the best of our papers to the very top-ranking journals, such as *Nature* (impact factor about 30), *Science* (30), *Behavioral and Brain Sciences* (7) and *Psychological Review* (7). (We are aware that this is an ambitious goal, but the last two journals are certainly within our reach given our recent publication record.)

Moreover, we plan to write at least two monographs:

1) A textbook for researchers and advanced students formally elaborating a conceptual framework for the modeling of distributed cognitive systems. Following a similar structure as this proposal, it will start with the most simple element (information, agents, links), and show how these can self-organize step by step to produce gradually more intelligent systems.

2) A practical handbook with exercises showing how to model distributed cognitive systems using our generic connectionist simulation environment.

In addition, this project should produce at least three PhD dissertations, investigating different computational and empirical aspects of our general project.

6.2 Simulation environments

Many of the connectionist simulations will be conducted with the aid of a software program, called FIT, developed by **Van Overwalle**. FIT implements the current "Talking Nets" model as described by **Van Overwalle** and **Heylighen** [2006], and can be freely downloaded from www.vub.ac.be/PESP/VanOverwalle.html#FIT. As the program is extended with the results of our research and with larger scale simulations, more advanced versions will also be made freely available to the research community.

Specifically, as part of this research project, we will integrate in the FIT software application other simulation approaches in a more general environment, combining the strengths of connectionist and multi-agent approaches, that can be used as a "virtual laboratory" for building models of distributed cognitive systems, and experimenting with them. In this way, other researchers and students will be able not only to replicate our results, but to devise their own models and explore their properties in a flexible and user-friendly manner.

6.3 Empirical data

Like the software we develop, we also plan to make the data gathered from our experiments and observations available via web, so that other researchers can use them to re-analyze and to test their own hypotheses.

6.4 Workshops, conferences and lectures

Like in the past, we will continue to regularly organize international meetings on the subject of distributed cognition and its specific aspects, so that our work can be discussed with other researchers in the domain, and receive input from their results. The talks presented at the more important meetings will be published in the form of proceedings. We will also present our ideas in seminars and lectures for local colleagues and PhD students, and include the most important insights in the undergraduate courses we teach.

6.5 Social relevance and potential applications

A model of distributed cognition as we envisage it here would offer a wealth of potential applications, with particular relevance to society at large. While our aim is in the first place to do basic research and reach a general understanding of the problem domain, we are likely to take a closer look at several of these application domains so as to confront our models with the practical issues that surround distributed cognition in reality, as contrasted with laboratory or theory.

To start with, a theory of distributed cognition has immediate applications in business, government, and other organizations, where problems need to be tackled by groups. It would help them to avoid the pitfalls of collective decision-making, such as *polarization* or *groupthink* [Janis, 1972], which stifle creativity and can sometimes lead to catastrophically bad decisions. In sum, it would foster the collective intelligence of the organization. More generally, it may help us to control for the cognitive biases and social prejudices whose ubiquity psychologists have amply demonstrated [Brauer et al., 2001; Klein et al., 2003; **Van Rooy**, **Van Overwalle** et al., 2003].

On a larger scale, understanding how knowledge and information are distributed throughout social systems would help to foster the economic and social development that new knowledge and better communication engenders [**Martens**, 1998; 2005]. In particular, such a theory should tell us how important new ideas can diffuse most efficiently, and conversely how the spread of false rumors, superstitions and "information parasites" can be held in check [**Heylighen**, 1998; **Chielens** & **Heylighen**, 2005].

Another important application is: under what conditions will two or more groups with different ideas or beliefs converge to a shared position? This problem describes the all too common situation in society where two cultures or subcultures live in close contact. In some cases the minority is assimilated in the majority (e.g., most initially European communities in the U.S.), in other cases there is strong and apparently growing polarization between groups (e.g., Islamic communities in Europe), leading to potentially violent conflicts. A first simulation with our model [**Van Overwalle** & **Heylighen**, 2006] suggests that the probability of convergence or assimilation increases with the amount of communication between the groups and with the degree of consistency between their initial positions, but a precise understanding of the circumstances that lead to such cultural integration/differentiation will require a much more in-depth investigation.

Insofar that our results provide concrete answers to such important social problems, we will try to make them known not only in specialized academic circles, but towards a broader public, including politicians, executives, and government administrations. This can be achieved for example via popularizing articles and books, contacts with journalists, press conferences, and appearances in the mass media (with which both of the promotors already have quite a bit of experience).

7. Project planning

A "GOA" research project, such as is this one, is scheduled to run for 5 years, i.e. from Jan. 1, 2007 to Dec. 31, 2011.

As detailed in section 5, we have divided the planned work according to general methodology (simulation, individual experiments, group experiments, Internet applications) into four independent approaches that can run in parallel. Moreover, given our previous experience and preparatory work, each subproject is ready to start at the beginning of the period: the simulation environment is ready to be run and further extended, the Internet applications have been outlined and merely need programming and debugging, while the infrastructure and basic designs for the laboratory experiments are available. Therefore, we would like to start the four work packages simultaneously, at the beginning of the 5 year period.

Moreover, the work packages have been designed to be flexible and open-ended: while we are likely to begin with the simplest and most clear-cut problems (e.g. replicating further existing findings (5.1.1), investigating the antecedents of individual trust(5.2.1), group communication (5.3.1) and designing the questions for a "collective IQ test" (5.4.1), our further investigations will be guided by the results of those first explorations. In particular, we plan for a continuous multidisciplinary interaction or feedback between the four approaches. For example, a failure to replicate a well-known result via simulation may inspire us to postulate an additional factor that needs to be tested via additional individual or group experiments. Similarly, an unexpectedly successful attempt to promote collective intelligence in the Internet environment may inspire us to test the same kind of set-up in the simulation or laboratory situation. Therefore, we have planned the project so that each work package will continue to run for as long as the funding allows (i.e. until the end of 2011).

The complete group (ECCO + SCL), under the supervision of the project promoters F. Heylighen and F. Van Overwalle, will provide input to the project in the sense of ideas for theoretical development, interpretation of empirical results, and general feedback. However, since the present members already have practically full-time responsibilities for other research projects (even when these have related subjects), for the bulk of the work (programming, setting up experiments and environments, collecting and analyzing data) we will need to employ four new research assistants, one for each subproject.

The two experimental subprojects (5.2 and 5.3) will require social scientists with experience in empirical research and statistics (preferably psychologists). The simulation project (5.1) will require someone with good programming skills and preferably experience with multi-agent simulations. The Internet project (5.4) can be carried out either by a social scientist with good ICT skills, or a computer scientist with experience in social software applications.

These research assistants will be paid initially by means of PhD scholarships. By the time they have gathered enough material to write down and defend their PhD thesis, which should normally happen in the fourth year (2010), their PhD grant can be converted into a PostDoc grant. This should provide them with another year to perhaps collect and process some additional data, and especially write up and present their work to the international scientific community in the form of high-level peer-refereed publications.

This additional year will also allow them to help us organize a large international conference on the subject of distributed cognition, with both invited and submitted papers from specialists around the world, during which the members of our team will present all the major results of the project to the academic community. Moreover, this PostDoc year will help them to establish contracts abroad, visiting other research centers, and thus possibly lay the foundations for a further academic career. Of course, if one or more of the PhD students we initially employ fails to perform satisfactorily, that person will be replaced by someone else as soon as possible, so that the new person may still get enough time to lay the basis for a PhD.

Requested Funding

The following budget (all costs in Euro) provides an estimate of the funding we will need over the 5 years to run the project, split up into the different cost categories. Scholarship costs are based on the official estimates of the VUB personnel service for a researcher from an EU country.

YEAR	2007	2008	2009	2010	2011	TOTAL
Purchase of books and journals	1500	1500	1500	1500	1500	7500
Travel and accommodation, to let team members visit conferences and research centers abroad	3000	3000	3000	3000	3000	15000
Organization of workshops and travel+ accommodation for visiting experts	4000	4000	4000	4000	4000	20000
Payment of participants in experiments (800 people x 7.5 euro/person)	6000	6000	6000	6000	6000	30000
Scientific Software licences (Statistics, experiment generator, development platforms...)	1000	1000	1000	1000	1000	5000
Additional hardware and upgrades	3000	3000	3000	3000	3000	15000
computer for research assistant 1	2000					2000
computer for research assistant 2	2000					2000
computer for research assistant 3	2000					2000
computer for research assistant 4	2000					2000
PhD/PostDoc scholarship 1	29218	30095	32679	33660	35000	160652
PhD/PostDoc scholarship 2	29218	30095	32679	33660	35000	160652
PhD/PostDoc scholarship 3	29218	30095	32679	33660	35000	160652
PhD/PostDoc scholarship 4	29218	30095	32679	33660	35000	160652
TOTALS	143372	138880	149216	153140	158500	743108

Relevant publications of the research team

The following is a selection of the most important publications of the research team that are relevant for the present proposal. Most of these can be found and downloaded via http://scholar.google.com.

Bollen D. (2004): Representation in situated models of cognition (ECCO working paper).

Bollen J. (2001) *A Cognitive Model of Adaptive Web Design and Navigation - A Shared Knowledge Perspective*, Free University of Brussels, Faculty of Psychology, PhD Dissertation.

Bollen J., **Heylighen** F. (1996) Algorithms for the Self-organisation of Distributed, Multi-user Networks, in: *Cybernetics and Systems '96* R. Trappl (ed.), (Austrian Society for Cybernetics), p. 911-916.

Bollen J., **Heylighen** F. (1998): A system to restructure hypertext networks into valid user models, *New Review of HyperMedia and Multimedia* 4, p. 189-213.

Bollen, J., **Heylighen** F., **Van Rooy** D. (1998): Improving Memetic Evolution in Hypertext and the WWW, in: *Proc. 16th Int. Congress on Cybernetics* (Association Internat. de Cybernétique, Namur), p. 449-454.

Bollen. J. (2000) Group User Models for Personalized Hyperlink Recommendations. In LNCS 1892 *International Conference on Adaptive Hypermedia and Adaptive Webbased Systems* (AH2000), pages 39-50, Trento, August . Springer Verlag.

Chielens K. & **Heylighen** F. (2005): Operationalization of Meme Selection Criteria: Methodologies to Empirically Test Memetic Prediction, in: *Proceedings of the Joint Symposium on Socially Inspired Computing*, *AISB 2005 Convention*, p. 14-20.

Chielens, K. (2003) *The Viral Aspects of Language: A Quantitative Research of Memetic Selection Criteria*. Unpublished Masters Thesis VUB.

Gershenson, C. (2001). *Artificial Societies of Intelligent Agents*. Unpublished BEng Thesis. Fundacion Arturo Rosenblueth, Mexico.

Gershenson, C. (2002a). Philosophical Ideas on the Simulation of Social Behaviour. *Journal of Artificial Societies and Social Simulation* vol. 5, no. 3.

Gershenson, C. (2002b). Behaviour-based Knowledge Systems: An Epigenetic Path from Behaviour to Knowledge. *Proceedings of the 2nd Workshop on Epigenetic Robotics*. Edinburgh.

Gershenson, C. (2003). Comparing Different Cognitive Paradigms with a Virtual Laboratory. IJCAI-03: *Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence*, pp. 1635-6. Morgan Kaufmann.

Gershenson, C. (2005) Self-organizing Traffic Lights. (ECCO working paper 2005-02)

Gershenson, C. (2004). Cognitive Paradigms: Which One is the Best? *Cognitive Systems Research*, 5(2):135-156, June 2004.

Gershenson, C., **Heylighen** F. (2003). When Can we Call a System Self-organizing?, In Banzhaf, W, T. Christaller, P. Dittrich, J. T. Kim and J. Ziegler (eds.), Advances in Artificial Life, 7th European Conference, ECAL 2003, (Springer, LNAI 2801.), p. 606-614.

Gershenson, C., **Heylighen** F. (2004a). How can we think the complex? in: Richardson, Kurt (ed.) *Managing the Complex Vol. 1: Philosophy, Theory and Application*. [in press]

Gershenson C., **Heylighen** F. (2004b): Protocol Requirements for Self-organizing Artifacts: Towards an Ambient Intelligence, in: *Proc. Int. Conf. on Complex Systems* (New England Institute of Complex Systems)

Gershenson, C., M. A. Porter, A. Probst, M. Marko and A. Das. (2002) A Study on the Relevance of Information in Discriminative and Non-Discriminative Media. *InterJournal of Complex Systems* , 533.

Gershenson, C., P. P. González and J. Negrete. (2000a) Action Selection Properties in a Software Simulated Agent, in Cairó et. al. (Eds.) *MICAI 2000: Advances in Artificial Intelligence*. Lecture Notes in Artificial Intelligence 1793, pp. 634-648. Springer-Verlag.

Gershenson, C., P. P. González and J. Negrete. (2000b) Thinking Adaptive: Towards a Behaviours Virtual Laboratory. In Meyer et. al. (eds.) *Simulation of Adaptive Behavior 2000 Proceedings Supplement*. Paris, France. ISAB press.

Heylighen A., **Heylighen** F., **Bollen** J. & Casaer M. (2005): A distributed model for tacit design knowledge exchange, *Proceedings of Social Intelligence Design 2005*, R. Fruchter (ed.).

Heylighen F. (1990a): Autonomy and Cognition as the Maintenance and Processing of Distinctions, in: *Self-Steering and Cognition in Complex Systems*, **Heylighen** F., Rosseel E., Demeyere F. (ed.), (Gordon and Breach, New York), p. 89-106.

Heylighen F. (1990b): *Representation and Change. A Metarepresentational Framework for the Foundations of Physical and Cognitive Science*, (Communication and Cognition, Gent), 200 p.

Heylighen F. (1990c): A New Transdisciplinary Paradigm for the Study of Complex Systems?, in: *Self-Steering and Cognition in Complex Systems*, **Heylighen** F., Rosseel E. & Demeyere F. (ed.), (Gordon and Breach, New York), p. 1-16.

Heylighen F. (1991): Cognitive Levels of Evolution: pre-rational to meta-rational, in: *The Cybernetics of Complex Systems - Self-organization, Evolution and Social Change*, F. Geyer (ed.), (Intersystems, Salinas, California), p. 75-91.

Heylighen F. (1991): Design of a Hypermedia Interface Translating between Associative and Formal Representations, *International Journal of Man-Machine Studies* 35, p. 491-515.

Heylighen F. (1991): Modelling Emergence, *World Futures: the Journal of General Evolutio*n 31 (Special Issue on Emergence, edited by G. Kampis), p. 89-104.

Heylighen F. (1991): Structuring Knowledge in a Network of Concepts, in : *Workbook of the 1st Principia Cybernetica Workshop*, **Heylighen** F. (ed.) (Principia Cybernetica, Brussels-New York), p. 52-58.

Heylighen F. (1992a) : 'Selfish' Memes and the Evolution of Cooperation, *Journal of Ideas* , Vol. 2, #4, pp 77-84.

Heylighen F. (1992b) : Evolution, Selfishness and Cooperation, *Journal of Ideas*, Vol 2, # 4, pp 70-76.

Heylighen F. (1992c): A Cognitive-Systemic Reconstruction of Maslow's Theory of Self-Actualization, *Behavioral Science* 37, p. 39-58.

Heylighen F. (1992d): Non-Rational Cognitive Processes as Changes of Distinctions, in: *New Perspectives on Cybernetics. Self-Organization, Autonomy and Connectionism*, G. Van de Vijver (ed.), (Synthese Library v. 220, Kluwer Academic, Dordrecht), p. 77-94.

Heylighen F. (1992e): Principles of Systems and Cybernetics: an evolutionary perspective, in: *Cybernetics and Systems '92*, R. Trappl (ed.), (World Science, Singapore), p. 3-10.

Heylighen F. (1993): Selection Criteria for the Evolution of Knowledge, in: *Proc. 13th Int. Congress on Cybernetics* (Association Internat. de Cybernétique, Namur), p. 524-528.

Heylighen F. (1994) Fitness as Default: the evolutionary basis for cognitive complexity reduction, in: Cybernetics and Systems '94, R. Trappl (ed.), (World Science, Singapore), p.1595-1602.

Heylighen F. (1995): (Meta)systems as constraints on variation, *World Futures: the Journal of General Evolution* .45, p. 59-85.

Heylighen F. (1997a): Objective, subjective and intersubjective selectors of knowledge, *Evolution and Cognition* 3:1, p. 63-67.

Heylighen F. (1997b): The Economy as a Distributed, Learning Control System, *Communication and Cognition- AI* 13, nos. 2-3, p. 207-224.

Heylighen F. (1998): What makes a meme successful? Selection criteria for cultural evolution, in: *Proc. 16th Int. Congress on Cybernetics* (Association Internat. de Cybernétique, Namur), p. 423-418.

Heylighen F. (1999a): Collective Intelligence and its Implementation on the Web: algorithms to develop a collective mental map, *Computational and Mathematical Theory of Organizations* 5(3), p. 253-280.

Heylighen F. (1999b): The Growth of Structural and Functional Complexity during Evolution, in: F. **Heylighen**, J. **Bollen** and A. Riegler (eds.) The Evolution of Complexity (Kluwer Academic, Dordrecht), p. 17-44.

Heylighen F. (2000): Foundations and Methodology for an Evolutionary World View: a review of the Principia Cybernetica Project, *Foundations of Science*, 5, p. 457-490.

Heylighen F. (2001a): Bootstrapping knowledge representations: from entailment meshes via semantic nets to learning webs, *Kybernetes* 30 (5/6), p. 691-722.

Heylighen F. (2001b): Mining Associative Meanings from the Web: from word disambiguation to the global brain, in: *Proceedings of the International Colloquium: Trends in Special Language and Language Technology*, R. Temmerman and M. Lutjeharms (eds.) (Standaard Editions, Antwerpen), p. 15-44.

Heylighen F. (2002): The Science of Self-organization and Adaptivity, in: Knowledge Management, Organizational Intelligence and Learning and Complexity, in: *The Encyclopedia of Life Support Systems*, (Eolss Publishers, Oxford).

Heylighen F. (2006a): "Mediator Evolution: a general scenario for the origin of dynamical hierarchies", in: D. Aerts, B.D'Hooghe & N. Note (eds.) Worldviews, Science and Us. (Singapore: World Scientific)

Heylighen F. (2006b): The anticipation-control theory of mind: an evolutionary-cybernetic perspective on cognition, consciousness and the brain (ECCO working paper 2006-01, to be submitted to *Behavioral and Brain Sciences*)

Heylighen F. (2006c): A Nodal Growth Algorithm for Concept Discovery (ECCO working paper 2006-02)

Heylighen F., **Bollen** J. (1996) The World-Wide Web as a Super-Brain: from metaphor to model, in: C*ybernetics and Systems '96* R. Trappl (ed.), (Austrian Society for Cybernetics).p. 917-922.

Heylighen F., **Bollen** J. (2002): Hebbian Algorithms for a Digital Library Recommendation System, in *Proceedings 2002 International Conference on Parallel Processing Workshops* (IEEE Computer Society Press)

Heylighen F., **Bollen** J., Riegler A. (ed.) (1999): *The Evolution of Complexity* (Kluwer Academic, Dordrecht).

Heylighen F., Campbell D.T. (1995): Selection of Organization at the Social Level: obstacles and facilitators of metasystem transitions, *World Futures: the Journal of General Evolution* 45, p. 181-212.

Heylighen F., Dewaele J-M. (2002): Variation in the contextuality of language: an empirical measure, *Foundations of Science* 6, p. 293-340

Heylighen F., **Gershenson** C. (2003): The Meaning of Self-organization in Computing, IEEE Intelligent Systems 18:4, p. 72-75.

Heylighen F., Heath M., F. **Van Overwalle** (2004): The Emergence of Distributed Cognition: a conceptual framework, *Proceedings of Collective Intentionality IV*, Siena (Italy).

Heylighen F., Joslyn C. & Turchin V. (eds.) (1993-2004): Principia Cybernetica Web (http://pespmc1.vub.ac.be/),

Heylighen F., Joslyn C. (1993): Electronic Networking for Philosophical Development in the Principia Cybernetica Project, *Informatica* 17, No. 3, p. 285-293.

Heylighen F., Joslyn C. (1995): Systems Theory, in: *The Cambridge Dictionary of Philosophy*, R. Audi (ed.) (Cambridge University Press, Cambridge), p.784-785.

Heylighen F., Joslyn C. (2001): Cybernetics and Second Order Cybernetics, in: R.A. Meyers (ed.), *Encyclopedia of Physical Science and Technology* (3rd ed.), Vol. 4 , (Academic Press, New York), p. 155-170.

Heylighen F., Rosseel E., Demeyere F. (eds.) (1990): *Self-Steering and Cognition in Complex Systems. Toward a New Cybernetics*, (Gordon and Breach Science Publishers, New York), 440 p.

Jordens, K., **Van Overwalle, F**. (2004). Connectionist Modeling of Attitudes and Cognitive Dissonance. In G. Haddock, G. Maio (Eds.) *Contemporary perspectives on the psychology of attitudes*. London: Psychology Press.

Jordens, K., & **Van Overwalle, F**. (2006) Cognitive dissonance and affect: An initial test of a connectionist account. *Psychologica Belgica*, in press.

Martens B. (1998): The relationship between knowledge and economic growth: an economic-cognitive approach *Proc. 23rd Flemish Economic Congress*, Leuven.

Martens B. (1999): The introduction of complexity concepts in economics: towards a new paradigm, in: F. **Heylighen**, J. **Bollen** and A. Riegler (eds.) *The Evolution of Complexity* (Kluwer Academic, Dordrecht),

Martens B. (1999): Towards a Generalised Coase Theorem: a theory of the emergence of social and institutional structures under imperfect information, in Barnett et al, (eds.): *Commerce, Complexity and Evolution*, Cambridge University Press.

Martens B. (2005): *The cognitive mechanics of economic development and institutional change*, (Routledge)

Martens B., P. Murrell, P. Seabright and U. Mummert (2002), *The institutional economics of foreign aid,* Cambridge University Press.

Rocha L. M., J. **Bollen** (2000). Biologically motivated distributed designs for adaptive knowledge management. In I.Cohen and L. Segel, editors, *Design Principles for the Immune System and other Distributed Autonomous Systems*, Oxford University Press. pp. 305-334.

Rodriguez, M. (2004), Advances Towards a General-Purpose Human-Collective Problem-Solving Engine, *Proc. International Conference on Systems, Man and Cybernetics*, IEEE SMC.

Rodriguez, M., Steinbock, D. (2004), A Social Network for Societal-Scale Decision-Making Systems, *Proc. North American Association for Computational Social and Organizational Science Conference*.

Rodriguez, M.A., Steinbock, D.J. (2006), "The Anatomy of a Large Scale Collective Decision Making System", [submitted], Los Alamos National Laboratory technical report: LA-UR-06-2139.

Rodriguez, M.A., **Bollen**, J., Van de Sompel, H. (2006), "The Convergence of Digital Libraries and the Peer-Review Process", Journal of Information Science, 32(2), pp. 149-159, April .

Rodriguez, M.A., Steinbock, D.J., Watkins, J.H., **Gershenson**, C., **Bollen**, J., Grey, V., deGraf, B. (2006), "Smartocracy: Social Networks for Collective Decision Making", [submitted], Los Alamos National Laboratory technical report: LA-UR-06-2244.

Rodriguez, M.A., Watkins, J.H., Steinbock, D.J., **Gershenson**, C. (2006), "Snurf: Social Networks for News Media Distribution", [submitted], Los Alamos National Laboratory technical report: LA-UR-06-2564, .

Timmermans, B., & Cleeremans, A. (2000). Rules versus statistics in biconditional grammar learning. *Proceedings of the 22nd Annual Meeting of the Cognitive Science Society*, 947-952. NJ: Erlbaum

Timmermans, B., **Van Duynslaeger,** M., & **Van Overwalle**, F. (2006) *Spontaneous circumstantial attributions*. Manuscript submitted for publication.

Timmermans, B., **Van Duynslaeger**, M., Vandewaerde, N. & **Van Overwalle, F**. (2006). *Your hidden smile & frown affects me: Subliminally manipulating one's self-esteem using facial expressions and traits*. Manuscript submitted for publication.

Timmermans, B., **Van Overwalle F.** (submitted) Spontaneous circumstantial attributions.

Vanhoomissen, T. & **Van Overwalle, F**. *Me or not Me as Source of Ingroup Favoritism and Outgroup Derogation: A Connectionist Perspective.* Manuscript submitted for publication

Vanhoomissen, T. & **Van Overwalle, F**., & De Haan, J. *Accentuation between Groups and Episodic Memory: A test of a Recurrent Connectionist Model.* Manuscript submitted for publication

Van Overwalle F. (1989). Structure of freshmen's causal attributions for exam performance. *Journal of Educational Psychology, 81*, 400-407.

Van Overwalle F. (1997a) Dispositional attributions require the joint methods of difference and agreement. *Personality and Social Psychology Bulletin, 23*, 974-980.

Van Overwalle F. (1997b) A test of the Joint model of causal attribution. *European Journal of Social Psychology, 27*, 221-236.

Van Overwalle F. (1998) Causal explanation as constraint satisfaction : A critique and a feedforward connectionist alternative. *Journal of Personality and Social Psychology, 74*, 312-328.

Van Overwalle F. (2003) Acquisition of dispositional attributions: Effects of sample size and covariation. *European Journal of Social Psychology*, 33, 515—533.

Van Overwalle F. (2004). Multiple Person Inferences: A View of a Connectionist Integration. In H. Bowman, C. Labiouse (Eds.), *Proceedings of the Eighth Neural Computation and Psychology Workshop* (Progress in Neural Processing). London, UK: World Scientific

Van Overwalle F. (under revision) *Discounting and augmentation of dispositional and causal attributions.*

Van Overwalle F., & Labiouse, C. (2004) A recurrent connectionist model of person impression formation. *Personality and Social Psychology Review*, 8, 28-61.

Van Overwalle F., Drenth, T., Marsman, G. (1999). Spontaneous trait inferences : Are they linked to the actor or to the action ? *Personality and Social Psychology Bulletin, 25*, 450-462.

Van Overwalle F., **Heylighen**, F. (1991). Invariantie-kenmerken bij antecedente condities en attributionele dimensies : Waarnemen van oorzaken, verwachtingen en emoties. [Invariance features in antecedent conditions and attributional dimensions: Perceiving causes, expectations and emotions] In J. van der Pligt, W. van der Kloot, A. van Knippenberg and M. Poppe (Eds.) *Fundamentele sociale psychologie: Deel 5* (pp. 44-60). Tilburg : Tilburg University Press.

Van Overwalle F., **Heylighen**, F. (1995) Relating covariation information to causal dimensions through principles of contrast and invariance. *European Journal of Social Psychology*, 25, 435-455.

Van Overwalle F., Heylighen F. (2006): “Talking Nets: A Multi-Agent Connectionist Approach to Communication and Trust between Individuals”, *Psychological Review* Vol. 114, issue 3 (in press).

Van Overwalle F., **Heylighen**, F., Casaer, S., Daniëls, M. (1992). Preattributional and attributional determinants of emotions and expectations. *European Journal of Social Psychology, 22*, 313-329.

Van Overwalle, F., Jordens, K. (2002). An adaptive connectionist model of cognitive dissonance. *Personality and Social Psychology Review*, 3, 204—231.

Van Overwalle, F., Labiouse, C. (2004) A recurrent connectionist model of person impression formation. *Personality and Social Psychology Review*, 8, 28—61.

Van Overwalle, F., Mervielde, I., De Schuyter, (1995). Structural modeling of the relationships between attributions, emotions and behavior of college freshmen. *Cognition and Emotio*n, 9, 59-85.

Van Overwalle F. & Siebler, F. (2005). A Connectionist Model of Attitude Formation and Change. *Personality and Social Psychology Review, 9*, 231–274..

Van Overwalle, F., Timmermans, B. (2001). Learning about an Absent Cause: Discounting and Augmentation of Positively and Independently Related Causes. French, R.M.,, Sougné, J.P. (Eds.) *Connectionist Models of Learning, Development and Evolution: Proceedings of the Sixth Neural Computation and Psychology Workshop, Liege, Belgium, 16-18 September 2000*. Springer Verlag.

Van Overwalle, F. & **Timmermans**, B. (2005) Discounting and the Role of the Relation between Causes. *European Journal of Social Psychology, 35*, 199—224.

Van Overwalle, F., **Van Rooy**, D. (1998) A Connectionist Approach to Causal Attribution. In S. J. Read, L. C. Miller (Eds.) *Connectionist and PDP models of Social Reasoning and Social Behavior* (pp. 143—171). Lawrence Erlbaum.

Van Overwalle, F., **Van Rooy**, D. (2001a). How one cause discounts or augments another: A connectionist account of causal competition. *Personality and Social Psychology Bulletin*, 27, 1613—1626.

Van Overwalle, F., **Van Rooy**, D. (2001b). When more observations are better than less : A connectionist account of the acquisition of causal strength. *European Journal of Social Psychology*, 31, 155-175.

Van Rooy, D. (2000): *A connectionist model of illusory correlation*. (PhD Thesis, Vrije Universiteit Brussel).

Van Rooy, D., **Vanhoomissen**, T., **Van Overwalle, F**. (submitted) Illusory correlation, sample size and memory: A connectionist approach.

Van Rooy, D., **Van Overwalle, F.**, **Vanhoomissen**, T., Labiouse, C., French, R. (2003). A recurrent connectionist model of group biases. *Psychological Review*, 110, 536-563.

Bibliography (publications by others)

§Ahuja MK, KM Carley (1999): Network structure in virtual organizations, *Organization Science*, 10 n.6, p.741-757

Aunger R. (ed.) (2001): *Darwinizing Culture: The Status of Memetics As a Science* (Oxford University Press)

Bangerter, A., & Heath, C. (2004). The Mozart effect: Tracking the evolution of a scientific legend. *British Journal of Social Psychology, 43*, 605-623.

Berger P. L., T. Luckmann: (1967) *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*, Anchor.

Bonabeau E., Dorigo M. and Theraulaz G. (1999) *Swarm intelligence: From natural to artificial systems*. Oxford University Press.

Brauer, M., Judd, C. M., & Jacquelin (2001). The communication of social stereotypes: The effects of group discussion and information distribution on stereotypic appraisals. *Journal of Personality and Social Psychology, 81*, 463—475.

Crutchfield J., Shalizi C., Tumer K & Wolpert D. (eds.) (2002): *Collective Cognition Workshop Proceedings: Mathematical Foundations of Distributed Intelligence* (http://www.santafe.edu/~dynlearn/colcog/, to be published in the Santa Fe Institute Studies in the Sciences of Complexity, Oxford University Press;)

Ebbesen, E. B. & Bowers, R. J. (1974). Proportion of risky to conservative arguments in a group discussion and choice shift. *Journal of Personality and Social Psychology, 29,* 316—327.

Epstein J.M. & R. Axtell (1996): Growing Artificial Societies: Social Science from the Bottom Up. MIT Press: Cambridge, MA.

Fiske, S. F. (2005). Social beings: A core motives approach to social psychology. Hoboken, JN: Wiley.

Foltz P.W. (1996) Latent Semantic Analysis for text-based research, *Behavior Research Methods, Instruments, & Computers*, 28, 197-202.

Garrod, S. (1998) How groups co-ordinate their concepts and terminology: Implications for Medical Informatics. *Methods of Information in Medicine* , 37, 471-477.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), Syntax and semantics: Speech acts (pp. 41—58). New York: Academic Press.

Heath C., Bell C. & Sternberg E. 2001. Emotional Selection in Memes: The Case of Urban Legends. *Journal of Personality and Social Psychology* 81(6):1028-1041.

Hutchins E (1995): *Cognition in the Wild* (MIT Press).

Isenberg, D. J. (1986). Group polarization: A critical review and meta-analysis. *Journal of Personality and Social Psychology, 50*, 1141—1151.

Janis, I. L. (1972) *Victims of groupthink*. (Boston: Houghton Mifflin).

Johnson N. (1998), "Collective Problem-Solving: functionality beyond the individual" (Los Alamos National Laboratory technical report: LA-UR-98-2227)

Kaplan, C.A. (2001): Collective intelligence: a new approach to stock price forecasting, in: 2001 IEEE International Conference on Systems, Man, and Cybernetics, 5, p: 2893-2898.

King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R. & Montague, R. (2005). Getting to Know You: Reputation and Trust in a Two-Person Economic Exchange. *Science, 308*, 78—83.

Klein, O. Jacobs, A., Gemoets, S. Licata, L. & Lambert, S. (2003). Hidden profiles and the consensualization of social stereotypes: how information distribution affects stereotype content and sharedness. *European Journal of Social Psychology, 33*, 755—777.

Krauss, R.M. & Weinheimer, S. (1964). Changes in reference phrases as a functio of frequency of usage in social interaction: A preliminary study. *Psychonomic Science, 1*, 113—114.

Larson, Jr, J. R., Christensen, C., Abbott, A. S. & Franz, T. M. (1996). Diagnosing groups: Charting the flow of information in medical decision-making teams. *Journal of Personality and Social Psychology, 71*, 315—330.

Larson, Jr., J. R., Foster-Fishman, P. G. & Franz, T. M. (1998). Leadership style and the discussion of shared and unshared information in decision-making groups. *Personality and Social Psychology Bulletin, 24*, 482—495.

Leavitt, H. J. (1951). Some effects of certain communication patterns on group performance. *Journal of Abnormal and Social Psychology, 46*, 38—50.

Lévy P. (1997): *Collective Intelligence*, Plenum.

Lyons, A. & Kashima, Y. (2003) How Are Stereotypes Maintained Through Communication? The Influence of Stereotype Sharedness. *Journal of Personality and Social Psychology, 85*, 989-1005.

Mackenzie, K. D. (1976). *A theory of group structure*. New York: Gordon & Breach.

Mackie, D.& Cooper, J. (1984) Attitude polarization: Effects of group membership. *Journal of Personality and Social Psychology, 46 (3),* 575—585.

McLeod, P., Plunkett, K. & Rolls, E. T. (1998). *Introduction to connectionist modeling of cognitive processes*. Oxford, UK: Oxford University Press.

McPherson M, L Smith-Lovin, JM Cook2001, BIRDS OF A FEATHER: Homophily in Social Networks, *Annual Review of Sociology,* 27, 1, p.415

Mitchell JC (1969): The Concept and Use of Social Networks, Bobbs-Merrill

Nadler A, I Liviatan (2006) Intergroup Reconciliation: Effects of Adversary's Expressions of Empathy, Responsibility, and and Recipients' Trust, *Personality and Social Psychology Bulletin 32*, 459.

Neisser U. (1976), Cognition and Reality: Principles and Implications of Cognitive Psychology. WH Freeman,

Nowak, A., Szamrej, J. & Latané, B. (1990). From private attitude to public opinion: A dynamic theory of social impact. *Psychological Review, 97*, 362—376.

Rumelhart D.E. & J.L. McClelland (editors) (1986): *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, volume 1. MIT Press.

Schober, M. F. & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology, 21*, 211—232.

Schul Y, R Mayo, E Burnstein (2004) Encoding under trust and distrust: the spontaneous activation of incongruent cognitions, *J Pers Soc Psychology* 86, p. 668.

Searle, J. (1995): *The Construction of Social Reality*, Free Press

Senge P. (1990) T*he Fifth Discipline: the Art and Practice of Learning Organizations*, Doubleday.

Shaw, M. E. (1964). Communication networks. In L. Berkowitz (Ed.) *Advances in Experimental Social Psychology, 1*, 111-147. New York: Academic Press.

Stasser, G. (1999). The uncertain role of unshared information in collective choice. In L. L. Thompson, J. M. Levine & D. M. Messick (Eds.) *Shared Cognition in Organizations: The management of Knowledge* (pp 49—69). Mahwah, NJ: Erlbaum.

Steels L. (1998): Synthesising the origins of language and meaning using co-evolution, self-organisation and level formation, in Hurford et al. (eds): *Approaches to the evolution of language* (Cambridge University Press), p. 384-404.

Sun, R. (2001). Cognitive science meets multi-agent systems: A prolegomenon. *Philosophical psychology, 14,* 5-28.

Surowiecki, J. (2004). The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economics, societies, and nations. New York: Doubleday.

Susi, T. & Ziemke, T. (2001). Social Cognition, Artefacts, and Stigmergy: A Comparative Analysis of Theoretical Frameworks for the Understanding of Artefact-mediated Collaborative Activity. *Cognitive Systems Researc*h, 2(4), 273-290.

Watts DJ , PS Dodds, MEJ Newman (2002): Identity and Search in Social Networks, *Science* 296. no. 5571, pp. 1302 - 1305

Weiss G. (1999): Multiagent systems: a modern approach to distributed artificial intelligence, (Cambridge, Mass. : MIT Press).

Wittenbaum, G. M. & Bowman, J. M. (2003). A social validation explanation for mutual enhancement. *Journal of Experimental Social Psychology,* 40, 169—184.

5 suggesties van externe experten voor de screening van het projectvoorstel

Dr. Bruce Edmonds
Centre for Policy Modelling
Manchester Metropolitan University
Aytoun Building, Aytoun Street, Manchester M1 3GH, UK.
Tel. +161 247 6479 Fax. +161 247 6802
bruce@cfpm.org
http://bruce.edmonds.name/

Prof. Dr. Luis Mateus Rocha
Cognitive Science Program
School of Informatics, Indiana University
1900 East Tenth Street, Bloomington IN 47408,USA
Fax: 812-856-3015, e-mail: rocha@indiana.edu
http://informatics.indiana.edu/rocha/lr_form.html

Prof. Dr. Mark H. Bickhard
Department of Psychology, Lehigh University
17 Memorial Drive East
Bethlehem, PA 18015, USA
mhb0@lehigh.edu
http://www.lehigh.edu/~mhb0/mhb0.html

Prof. Dr. Stuart Umpleby,
Research Program in Social and Organizational Learning
The George Washington University
2101 F Street NW, Suite 201, Washington, DC 20052 USA
Fax: +1- 202-994-3081
umpleby@gwu.edu
http://www.gwu.edu/~umpleby/bio.html

Prof. Dr. Yoshi Kashima
Department of Psychology, School of Behavioural Science
12th Floor, Redmond Barry Building
The University of Melbourne, Victoria 3010, Australia
Telephone: (+61 3) 8344 6312
y.kashima@psych.unimelb.edu.au
http://www.psych.unimelb.edu.au/staff/kashima.html