VRIJE UNIVERSITEIT BRUSSEL
Taal en Letterkunde: Germaanse Talen
Academiejaar 2002-2003

The Viral Aspects of Language.

A Quantitative Research of Memetic Selection Criteria.

Eindverhandeling ingediend tot het verwerven van
de graad van licentiaat in de Germaanse Taal- en Letterkunde.

Klaas Chielens

Promotor: Prof. Francis Heylighen

"The important thing in science is not so much to obtain new facts as to discover new ways of thinking about them."

(William Lawrence Bragg, Beyond Reductionism, 1968)

Acknowledgements

Writing a seniors thesis is rarely the accomplishment of only one person. In the process of writing and doing the research as well as in the daily motivation other people play an important role.

First of all, my wife, Theresa. She was always the first one to proofread and she was my greatest critic as well as my most passionate supporter. She helped me through periods of disillusionment and despair.

My parents and family have always supported me through my time as a student, even though my grades were often disappointing. It is to them that I dedicate this project, as token of my gratitude to their patience, love and support.

Francis Heylighen, research professor at the V.U.B. and one of the authorities in the field of memetics provided me with support and helpful information on the methodology to do this research. By being available for an exchange of thoughts but refusing to give comments on what I wrote he stimulated me even more and made the experience and self-fulfillment which one gets from finishing a project like this even richer.

Proofreaders of the first few chapters were Bart Blanquart, who was the first person ever to mention memetics to me, and Peter Mc Aulay.

Finally I would like to thank Nausikaä Glorieux. When I was first diagnosed with ADHD, three years ago, she was the one who guided me through a difficult period of accepting my own limitations. It is thanks to her motivation and assistance that since then I have been able to learn to understand ADHD and to learn to cope with it. Without her help I would never have graduated.

Table of Contents

1. Introduction

The field of memetics is still a theoretical field which has provided some valuable insights for different sciences but its usefulness remains limited as long as the theoretical insights can not be put to practical use. In this research an attempt will be made to measure memetic selection criteria and to see if the theory can stand up to the mathematical results.

The transition from theoretical criteria to criteria which are useful for quantitative research will bring up problems concerning the objectivity of the criteria and the limitations of only being able to do research on some of the criteria. It will be one of the goals of this project to find out what possible pitfalls can show up that have not been accounted for in the theoretical foundations of memetics. This may not only provide an insight on how memetics can be tested but it will also provide new arguments and aspects to the theoretical study of memetics.

By measuring memetic selection criteria it will theoretically be possible to make a ranking of the criteria that were used in this survey as to importance of the criteria towards the spreading of the researched memes, the virus hoaxes. The practical outcome of the study however will focus more on the correlations between the spreading of the virus hoaxes compared to the criteria and the different problems that are encountered when administering a survey like the one used.

Because of the size of this research project, a senior's thesis, there are limits to what exactly can be calculated or what the result will be. After all, it is not possible to measure all criteria for all possible virus hoaxes within the limits of this research. It is, therefore, not the goal of this research to provide a complete view of the importance of different selection criteria but to provide possible explanations for the correlations of the chosen criteria with the degree of fitness of a hoax.

The structure of the paper provides a theoretical outline at first before going in to the specific focus of the research.

An introduction to the field of memetics, which is essential if one wants to conduct a practical test of these theoretical concepts, can be found in the first chapter.

The second chapter concentrates on the proliferation behavior of memes and takes a closer look at the phases of memetic selection and the criteria which memes have to go through in order to spread successfully.

After these theoretical chapters a description of the methodology of the research is given and the selection of the corpus that will be used is explained. Finally the results of the quantitative research are analyzed and compared to this thesis' hypothesis.

The final chapter gives a general overview on the problems that were encountered during this project, recapitulates the research results briefly and gives some suggestions for further research in the field of memetics linked with the study of linguistics.

2. Introduction to Memetics

2.1. Introduction

When Richard Dawkins coined the term "meme" in The Selfish Gene, he gave food for thought to the scientific community of the time. In the eleventh chapter of his book, Dawkins briefly introduces a "unit of cultural transmission" (189-201). Many other scientists since then have commented on and reacted to this theory. A controversy for one, a welcome new light in the dark world of science for the other, memetics has slowly emerged from Dawkins' book and now has truly become an interdisciplinary field which has found proponents (and opponents) in sociology, biology, psychology and other fields.

2.2. Towards a definition of memes

Since Dawkins first spoke of memes in The Selfish Gene, much time has passed and several scientists have ventured into the realm of memetics. It has proven to be a problem for scientists to find a standing definition of '*a meme*'.

Dawkins only gave a very simple definition of what a meme is. Most authors quote him by limiting his meme definition to "a unit of cultural transmission". It can be interesting, however, to also add some words from the following paragraph of Dawkins' book to this early definition; more specifically the claim that memes "propagate themselves in the meme pool by leaping from brain to brain via a process which, in the broad sense, can be called imitation" (Selfish Gene 192). As this truly is the first definition, it is important to add this very specific factor of imitation which is usually not included when Dawkins' definition is quoted.

Other scientists have altered the definition; either adding more information and thus expanding what exactly can be called a meme or by limiting the definition to fit within a specific field of research. According to Wilkins, a meme is "The least unit of socio-cultural information relative to a selection process that has favourable or unfavourable selection bias that exceeds its endogenous tendency to change" (What's in a meme). Heylighen describes memes as "an information pattern, held in an individual's memory, which is capable of being copied to another individual's memory" (Memetics).

Susan Blackmore comes to the conclusion that it is nearly impossible to agree on a definition of memetics (Meme Machine 53). Indeed, it seems that every scientist has his or her own definition of what exactly can be considered a meme. In this study a definition is used that is somewhere in the middle of the three aforementioned ones. A meme is an information pattern which is capable of being copied to another individual's memory, mostly by means of imitation (though other techniques are possible as well) and which is subject to a selection process.

2.3. Memes as replicators

One of the major changes throughout the emergence of life on this planet was the formation of a special molecule. Whereas the molecules that were already formed at that time were more or less coincidental bindings of free-floating building blocks that were available in the primeval soup, at a certain point in time a molecule was formed which had a most noteworthy capacity. This new molecule was formed in a way that it "had the extraordinary property of being able to create copies of itself" (Selfish Gene 15). This evolution gave way to the creation of more replicators which was the final breakthrough to life on this planet.

Long before Dawkins, other scientists had uttered the idea of cultural transmission. Sociobiology had already acknowledged man as a product of not only biology but also of its cultural environment. Earlier than that James Baldwin had used the term “social heredity” to describe the learning process of mankind (Blackmore 24) and in 1880 William James published an article in which he talked about the “remarkable parallel (…) between the facts of social evolution on the one hand, and of zoological evolution as expounded by Mr. Darwin on the other” (1).

Dawkins, however, saw memes as replicators and compared them to the better known replicators: genes. When studying genetics, one makes a difference between the genotype, which is "(…) the sum total of all the genes present in an individual" and the phenotype, which is the way we perceive the organism. The phenotype is the result of interactions between the genotype and the environment (Curtis and Barnes, G10, G16), so different organisms can have the same genotype but different phenotypes. However, it is only the genotype that will be passed on when the gene spreads (through replication/ reproduction). Memes also have the ability to adapt themselves by interaction with the environment (host), and in this way change their phenotypic properties (Dennett, DDI 355).

For a gene, a genotype will only show in what we know as the phenotype of this particular genotype while still passing on its original genotype, no matter what changes are made to the phenotype. A meme however, will not be passed on in its original ‘genotypic’ state but will instead be copied in a phenotypic state which will, for the new host, become the genotype of this meme.

2.4. The Meme-Gene confusion

The idea of having a new replicator, the meme, grew from Dawkins work on *The Selfish Gene*. He wanted to prove that not only the well known replicators (i.e. genes) existed, but that there was a whole range of replicators which roughly followed the behavior of the selfish gene, focusing mainly on its own reproduction. That is why he added the chapter about memes to his book (Selfish Gene 322). Memes, however, do not behave in exactly the same way as genes. They can both be considered *selfish* replicators, which means that, from their own view, they only work to propagate themselves, regardless of other replicators. Some scientists have taken the naming of memes as an indication of a similar operational behavior, even though Dawkins claims it was just meant as wordplay on 'genes'.

For Dawkins, memes were, in the first place, just another example of what he called the selfish replicators. He did not invent them or describe them to give way to a new field, memetics, but to show that there are other replicators besides genes. He looks at memes as other selfish replicators, which might lead him to new insights into the behavior of genetic replicators -- and hence genetic natural selection (Extended Phenotype 112).

Genes are a part of the exact sciences, where strict rules apply for almost every occasion (though quantum theory seems to put limits to the exactness). Memes are more likely to be used in domains where such exactness is not feasible, such as anthropology, sociology, psychology or linguistics. The difference in fields makes way for another method to deal with these replicators. Dawkins indicates that he himself is not really sure whether or not the field of memetics can use the same Darwinian theories as the field of genetics can but he wants to indicate that the way of thinking about cultural phenomena and their spreading, compared to the genetic-Darwinian way of thinking can bring forth interesting results (Selfish Gene 322).

Dennett, however, describes memes as units that are affected by natural selection and evolution. The conditions for a natural selection process are:

- Variation: a continuing abundance of different elements.
- Heredity or replication: the elements have the capacity to create copies or replicas of themselves.
- Differential "fitness": the number of copies of an element that are created in a given time varies, depending on the interaction between the features of that element (whatever it is that makes it different from other elements) and features of the environment in which it persists. (Darwin's Dangerous Idea 343)

The mutation process or *variation* of memes and genes is quite different. Genes mutate 'by accident'; memes can mutate quite consciously or under influence of other memes. The meme carrier can take an active role in the mutation of the meme before spreading it (or *replicating* it) as a new mutation. This deliberate mutation is part of the variation (Bjarneskans, Grønnevik, and Sandberg).

The time-span needed for the effects of a mutation to become clear is another difference between memes and genes. An unsuccessful mutation will in almost all cases be immediately visible in a genetic mutation (because the mutation itself will not often be able to duplicate or copy itself), whereas, due to the selection criteria of memes, an unsuccessful mutation may be duplicated without errors until it hits the criteria of a host that does not accept the mutation. This notion of being biased as far as acceptance is considered is what Dennett calls *differential fitness* (Darwin's Dangerous Idea 343).

2.5. Linked Memes or Memeplexes

Many scientists have argued that memes do not always work alone; just like genes group themselves in gene complexes, memes group themselves in so-called meme complexes or memeplexes. Like genes, they work together and influence each other. Memeplexes work together in ways that they will build in protections for each other within the memeplex.

By taking a very open definition of memes as a starting point, it is possible to see every meme as a memeplex because, in some way, our own views have been influenced by our own memes and are thus unavoidably linked to them. This relates closely to the phenomenological problem that all the research is, per definition, not entirely objective, as there is already a frame of reference in which the research is carried out. The German philosopher, Kant, said that our thinking was limited by our perception; because of this, we can never observe the true nature of something ("ding an sich") but only the nature of the object as it is perceived, shaped by our frame of reference.

Both Dawkins and Blackmore have described how religions can be defined as memeplexes. It is essential to see the importance of a memeplex such as religion. Such memeplexes do not only find shelter in the mind of a new host, but they will change the perceptions and life of their new host

The purpose of religion may seem awkward or even unintelligible, but to the host the memeplex of religion creates a paradigm through which he or she can solve philosophical questions and feel content in knowing that these questions can be solved. The built-in defense mechanisms against other explanations will furthermore protect the host (and the memeplex) from being subjected to changes of this basic belief system. Aside from protecting the host from hostile meme-intrusions, religions also include a factor of 'conversion'. All major world-religions have a religious task to spread the religion and convert non-believers. Next to that, they all have their own holy scriptures which hasten spreading and make sure the memes can survive over time (Blackmore, Meme Machine 187-194).

2.6. Memes and their hosts

When taking a closer look at meme hosts one can easily make a distinction between two types. Meme hosts are divided into two groups, based on their ability to reflect upon the meme they are carrying. In his memetic lexicon, Glenn Grant calls these two different carriers vectors and hosts.

Vectors can not reflect upon the meme they carry. Examples of vectors are books, walls on which someone has left a message and almost every physical inanimate object. Humans can be vectors as well but they have the capacity to turn into hosts when they start contemplating the meme. The human vectors are therefore 'inactive hosts'.

Hosts are the opposite of vectors. They consider the meme, think about the meme and have what Grant calls a basic understanding of the meme. This requires a mechanism of reflection and thought.

There are two basic kinds of hosts, active and inactive. The first one, the active host, is what really makes the distinction between hosts and vectors. The active host is interpreting and thinking about the meme of which it is a carrier. The inactive host is a vector which has the potential of becoming a host at any time, because it has the capability to reflect upon the meme although it is, in its current situation, not doing so.

2.7. The Life-Cycle of a Meme

A meme is virtually immortal. The Life-span of a meme depends on several criteria. In the definition that is used in this research, the aspect of selection is also considered. The selection process of memes is what will determine their longevity. A meme which is frequently selected for replication will find more hosts and will have a larger 'offspring' than a meme which resides within one host and is only very rarely selected for replication and transmission. Seth Godin calls this measure of how fast a meme can spread its "velocity" (33).

When the field of marketing became interested in memetics, they tried to define the criteria to create a good meme in order to use these criteria to help create "Marketing Epidemics" and improve marketing strategies. According to Godin, "an ideavirus adores a vacuum" (28). It will be easier for a meme to be spread if there is not already a certain idea that embodies the same information that is contained in the meme. He gives the example of a search engine such as Yahoo! that could easily fit in to the new internet and fill up the vacuum there was in the lack of a structured search engine; the same goes for Amazon, eBay and Hotmail (Godin 63). The names have become buzz-words and are now well known to almost all internet users.

Next to the idea of the vacuum in the market, he also proposes two new criteria for memes; the velocity and the smoothness factor. The velocity refers to the speed the meme spreads, the smoothness is the factor that describes how easy it is for a meme to spread. In the case of viral marketing the memes propagate themselves, in other cases, especially in web-related memes; there is an easy way to spread the meme to other hosts. Websites use techniques such as buttons that allow you to "send this page to a friend" or will add in a self referring message in emails sent through the specific website (e.g. Hotmail) (Godin 33/64).

Once a meme has found a host, there are different selection processes at work. Heylighen described four selection phases for memes: 'Assimilation, Retention, Expression and Transmission' (1998). During each of these stages there is a striving of multiple memes to be the one that will be selected and within each of these stages memes must try to stay in the host instead of being pushed out of the way by another meme. These phases will be discussed more in depth in the chapter on memetic selection.

2.8. Examples of Memes

Many authors (e.g. Dawkins, Dennett and others) have summed up several examples of memes. Both Dennett & Dawkins give the example of Beethoven's symphonies. In this specific case it is often enough to take only a couple of bars, for example the first few bars of Beethoven's fifth symphony (Dennett 344).

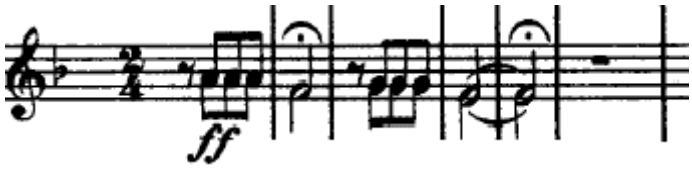


Figure 1: First six bars of Beethoven's fifth Symphony.

This example shows the importance of perception. At the end of the previous paragraph you may have imagined the first measures of this symphony. It can be assumed that you would not have done so if the only reference given was the figure of the first six bars. If a host already has the knowledge of modern western musical notation, he or she will find it easy to recognize the tune. A host who lacks this knowledge, or memeplex that allows the decryption and encryption of music, will not be able to decode the code and will therefore be insusceptible and will not be able to assimilate the meme in this form. This example also shows that if one would hear the first few notes, this may be enough to trigger the rest of the symphony or at least the rest of the first movement.

Other examples of memes are what Dennett calls '*distinct memorable units*', with examples such as the alphabet, chess, impressionism, building arches etc. (Darwin's Dangerous Idea 344). The notion of being 'memorable' limits memes in this definition. A random sequence of letters is not very memorable. When this random sequence of letters takes the shape of a self contained token with a distinct meaning to it, it becomes memorable and can then be considered a meme. So 'dll' may not mean anything to some people and may not be something that will be passed on, but to an experienced computer user and programmer, it will stand for 'dynamic link library' and will have a specific meaning. This shows that the frame of reference of the host is very important when considering meme examples.

The distinct meaning of the particular meme does not necessarily have to be decodable by conventional means. Symbols and logos often carry no direct link with their meaning. Nike's wavelike symbol says nothing about the company 'Nike' but is nevertheless so popular that the company has started using it even without the brand name.

Sometimes memes only proliferate within a small community. They become tokens of mutual understanding and represent a shared background or interest. Band logos, street gang tags and territory markers, religious symbols and the hacker alphabet all belong to a certain category of hosts. Another example is the number 42; to most people this is the number between 41 and 43 but to adepts of Douglas Adams' Hitchhikers Guide to the Galaxy, it is the answer to the question of life, the universe and everything.

An example of a graphical meme is the symbolic representation of the initials of J.R.R. Tolkien, author of the Lord of the Rings Trilogy and creator of the world of Middle Earth. These initials are an artistic rendition of the letters and represent not only the name of the author but specifically the name of the author linked to the novels which he wrote and not to his scholarly work.

Figure 2: Artistic rendition of Tolkien's Initials

2.9. Language Viruses

Language makes up for a large portion of memetic interaction. Conversations, publications and social interactions depend on language and are ideal ways to pass memes on from one host to another. This is why some authors have called memes 'language viruses'. Human communication can therefore be seen not only as functioning for the traditional Jacobson interpretation of communication and his functions of language[1], but can also be seen as a method and a way of spreading memes. Therefore, some scientists claim that the development of language is also caused by the necessity and driving force of memes in the early human evolution (Blackmore, The Power 57-59). Although these language viruses may seem harmless, the results can be very dangerous.

A typical example of a potentially dangerous language virus is a viral sentence. Viral sentences contain a copy instruction; these sentences are, for example, 'Copy me' or 'say me' (Hofstadter 52). These examples are not very successful, except for the fact that Hofstadter has copied these from Waltons and it has been copied in this research and in other papers from Hofstadter's article. Still, this little reproduction is nothing compared to the effects of a well-known viral sentence, or meme, that is highly successful.

[1] Jacobson defined 6 functions of language, Emotive, Conative, Phatic, Poetic, Referential and Metalinguistic.

The evolution of the internet has given viral sentences and viral messages an enormous boost. With the new information technology available to them, these memes have spread very easily all over the internet and there is hardly a mailbox untouched by this phenomenon. Chain letters and hoaxes do not seem to have any purpose at all. Their negative effects outweigh their positive effects (if they have any at all). They clog up bandwidth, mailboxes and valuable time. Nevertheless, they are spread on every day, even though internet users are told in many ways not to spread these stories.

Taking a closer look at electronic junk mail, it is possible to categorize it on different levels. There is one large section which is of no importance to this research and that is the direct marketing mails, also called 'unsolicited bulk e-mail' (UBE) or spam. They generally contain no copy instruction but are a collection of memes that marketers are trying to spread in order to sell their products or services. Other types, such as virus hoaxes and digitized chain letters or pleadings for help/postcards/prayers which do have a memetic value, contain a specific copy-instruction and demand from the receiver that he or she passes on the message contained in the email.

Virus hoaxes are false virus warnings. They urge the recipient to take actions that are often harmful to the working of the system by exploiting the fear of such harm coming to the system. Results of these actions can be that computers fail to do certain commands or disconnect their network (Sophos).

Whereas a virus hoax tries to convey a certain warning (a false warning but nevertheless a warning), a chain letter's sole purpose is to be forwarded to other people; often without any modification at all. Sometimes these chain letters will contain petitions against various subjects and will call upon the social awareness of the receiver to add their names to the support of a general cause. This kind of a petition will not have the slightest meaning because of, among other reasons, its exponential spreading factor and the uncontrollability.

The same reasoning applies to pleadings for help, postcards or prayers. Although they may very well have originated from a compassionate mind and call upon the empathy of the receiver, they remain a potential danger to the network infrastructure of companies due to the overload they can create.

Not all examples of memetic language use need to be perceived as dangerous. Considering the rather routine behavior of phatic communication, in which certain phrases are constantly used, it is possible to define these phrases as highly memetic. The fact that these phrases are used by a large number of people and that they are actually copied and taught to non native speakers of the language they appear in, classifies them as a culturally transmittable item, or, in our terms, a meme.

2.10. Suicide Memes

It is unlikely (but not impossible) that a memeplex which includes one single 'suicide' meme will be very successful. A memeplex like this automatically reduces its own longevity and threatens its own life-cycle. It is not entirely impossible for a suicide meme to survive the selection procedure.

It has happened before that cults have resulted to mass suicides in name of their religion. A relatively recent example, the *Heaven's Gate* cult, professed alien abduction when the comet Hale-Bopp was at the point where its course was the closest to the earth. However, they do not refer to the act of suicide as a crime or even as suicide. The cult even strongly objects to suicide. Their deaths are to be seen as a transgression to *'The next level'* (Heavens Gate Website).

The Islamic martyrdom is an example of a suicide meme that succeeds in being selected, although it will, when fully succeeding, kill the meme-host and therefore this individual meme. The survival of the meme-carrier has no direct link with the survival of the gene-carrier in this example (Dawkins, Extended Phenotype 110). The meme will have succeeded, however, through press attention, to spread along and to reach a high degree of publicity.

Not only cults and religions can spread suicide memes; suicide itself has occurred throughout history and there has been much research on the viral aspects of suicide. Newspaper reports about suicide tend to trigger more suicides and when Goethe published Die Leiden des jungen Werthers in 1774 a wave of suicides spread through Europe, causing the book to be banned in some countries.

2.11. Death of a meme

Although this introduction to memetics claimed that "a meme is virtually immortal", this proposition needs to be put in perspective. There are several factors that have to be taken into account. Death causes of memes are often tied to their host or carrier. Whenever a host dies, all memes within that host are dead. This varies from non-human carriers in that human carriers will often contain a meme but since it is not the genotype but the phenotype of a meme that is spread through human interaction, the meme dies with the human and only the specific phenotype, as passed on by this particular host, lives on. For non-human carriers, or vectors, this is not true as they do pass on the meme and not the phenotypic rendition of the meme. One can argue that the words, as recorded on the vector, are already a phenotype of the original ideas of the author but still it will be this exact phenotype and genotype which will be passed on every time.

Paul Bouissac gives three other factors: changes in the environment, intrinsic fragility and elimination by other strains of memes.

The environment of memes is not limited to textbooks, music or billboards. A large number of memes reside within the human brain. Because of this it is possible that certain conditions in the brain, such as Alzheimer's disease, can limit or completely halt the further spreading of memes (Bouissac 4). For other, non-human sources, the removal of the carrier of the meme, whether it be by destruction, voluntary removal or external factors making the carrier unreadable (such as loss of access to certain carriers; e.g. data-tapes for which we no longer have hard- or software to access them) will also possibly lead to the death of the meme. Of course, this is only valid when this one host is the only carrier of the meme or if this host is the last possible decoder of the meme, for example the last speaker of an indigenous language.

One of the intrinsic fragilities of memes has to do with the complexity of the memes. When a meme is very complex and very specialized, it becomes harder to pass on an exact copy of the meme (Bouissac 4). It is, however, not necessary for a meme to be copied in an exact way. Very frequently only the essence of a meme is passed on, e.g. the main points of a story, which will pass on the meme without giving an exact representation. As long as the meme is still fertile after this transmission, the life span of the meme will not be in danger. However, certain memes are very susceptible to these mutations (e.g. a scientific formula, a recipe …) and will possibly not survive if the meme is not an exact copy.

The last cause of death is the attack of other strains of memes. Just like the original replicators had to struggle to get the necessary building blocks for their reproduction, memes have to battle for the limited brain-resources of man. They can win this battle in two ways, the first one is to "neutralize the information value of their competitors", the second way is "to instruct their own hosts to destroy the hosts of competing memes" (Bouissac 5). An example of this behavior is the instruction in several religions to not only convert non-believers but also to attack anyone who opposes the spreading of the religion. The medieval crusades or recent terrorist' attacks from Muslim fundamentalists against Western targets are an illustration of this meme-induced behavior.

3. Spreading of Memes

3.1. Introduction

In an ideal situation a replicator will have an exponentially increasing number of carriers which are each hosts or vectors of this replicator. The hosts will help it to spread until every possible host is infected while the vectors will often allow other potential hosts to pick up the meme. The ideal situation in which every possible host is infected does not exist. The memes mutate and cause faults in replication which leads to a multitude of variations of the original replicator, up to a point where the original replicator will hardly be recognizable.

It is often seen that certain mutations are more successful in spreading then others. This is due to the fact that there are certain selection mechanisms which each of the potential hosts will impose on the meme before adopting it. Certain carriers will therefore accept a meme while others will not. Figure three shows these selection criteria at work, the black dots have not selected the meme for further propagation, the grey dots, on the contrary, are hosts which are actively spreading the meme.

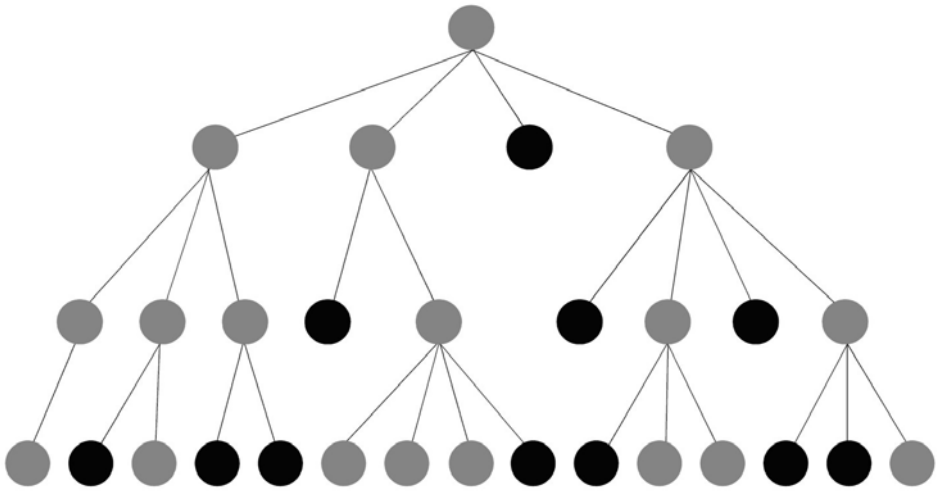


Figure 3: Selection Criteria at work.

Memes have to go through a 'battle for the brain'. The limited availability of meme-carriers and the large amount of memes makes it so that there is a competition among memes (Dennett, Exploitation of Imagination). Because of this battle, certain memes will be passed on while others will either remain in a dormant state or will die.

Memes have to go through several stages before they are actually adopted by their new hosts. At every stage there is a selective pressure and a struggle with other memes to be selected and to pass on to the next phase. Every phase means selection and therefore selection criteria.

These selection criteria vary from host to host. A criterion that is important for host Y may have less importance to host Z. A simple example of this is a grammatical error in a meme.

Consider a meme which propagates a certain product 'X'. This product calls upon self-esteem and promises something of great value to the carrier. A meme like this one can be found in certain frauds such as the Nigerian letter (of which an example can be found on http://www.sophos.com/virusinfo/hoaxes/nigerian.html). This letter offers money in exchange for the help in establishing an account or helping to transfer a large sum.

If respondent A receives this scam, he or she may act favorably on it because of the promised reward; this respondent will then have acted as a new carrier of this meme and will, in due time, find his own bank account half empty.

Respondent B however may notice that the letter contains some grammatical mistakes and may therefore conclude that, if this person really was who he or she claims to be, he or she would have a better knowledge of basic grammar rules and therefore respondent B may recognize this fraud for what it is and not react upon it in a favorable way.

If a meme wants to be propagated over several different hosts, then this meme will have to cater to different sets of selection criteria. If not, it will sooner or later have a very small number of carriers and eventually die or stop spreading.

3.2. Stages of Meme Replication

3.2.1. Introduction

When a meme is selected and afterward re-propagated by the new host, it has to go through several different phases before reaching the end-goal of replication and redistribution. Heylighen describes these four stages and calls them *assimilation*, *retention*, *expression* and *transmission*. These stages are only valid for human interaction. A lot of the criteria depend on a subjective feeling from the host. A vector can, by its definition, not reflect upon the meme so will not have subjective feelings towards it.

3.2.2. Assimilation

The assimilation process looks for a meme-carrier. This is the phase of the initial infection of the new host. There are three factors which are important in this stage: *the meme must be noticed, understood and accepted*. Noticed means that the meme has attracted the attention of the new host. The host has to select new memes; memes which succeed to attract the attention will be more likely to find a new host and therefore be more likely to be propagated.

Understanding the meme means that the potential host, after noticing the meme, can actually adapt it to its own knowledge. A meme may catch the attention of a potential host and therefore succeed in being noticed but when the potential host does not understand the meme, the chances of actual incorporation become much smaller. An example of this is a very flashy slogan with colorful and bright letters but in a language that is not one the host understands. This meme may be very successful with speakers of this language but will not succeed in being spread among non-speakers.

The last part of the assimilation phase is *accepting*. Even when a meme presents itself to a host in a way that immediately catches the host's attention and is understandable without any problem, then the last selection will decide whether or not the host accepts this meme. Here is where the meme will be checked against the already present knowledge and memeplexes of the host. When the new meme does not fit into the already existing frame of reference of the host, it is not likely to be accepted (Heylighen, What makes 1-2).

3.2.3. Retention

The second phase deals with how long the meme can stay in the memory of the host. The input of different memes is so large and differentiated that there will be a selection as to which memes will be retained and which memes will be forgotten. If a meme does not succeed in being selected in this retention phase, it can not be called a '*distinct memorable unit*' as Dennett called memes (Darwin's Dangerous Idea 344).

This is again a very subjective stage of replication. A host will only retain memes which are interesting or important. The phase of retention is therefore highly subjective. A way for a meme to encourage the retention is to make use of repetition (Heylighen, What makes 2). If a host hears a certain slogan frequently, it will retain the message, whether it is interested in it or not. Examples from this are catchphrases used in commercials, company slogans, radio jingles and popular songs.

Viral marketers use exactly this type of repetition to try to propagate their product. It is however wrong to consider all viral marketing to be based on commercial principles. Memes are often propagated out of social habits, group membership and ordinary daily routine.

An example of this is the religious duty to spread a religion. Through the use of prayers, incantations and mantras the retention of the religious memes is made easier for the potential host and therefore it has a higher potential chance of succeeding in finding the needed number of hosts to ensure its longevity.

Some memes have a short but very intense lifespan. Occasionally a meme develops a cult-status which leads to enormous spreading within specific sub-cultures such as computer and internet enthusiasts. In no time a meme like this will spread all over the internet. An example of this was the spreading of the sentence "All your base are belong to us". This sentence is originally derived from a translation of a Japanese computer game (Zero Wing) into English. From the summer of 2000 until the end of 2001 this meme spread over the internet in such a way that it was noticed by several magazines and newspapers. Big software companies such as Microsoft and Hewlett-Packard, which normally do not lend themselves to humor on their corporate website, incorporated "all your base are belong to us" in their websites, thus reinforcing the meme. By the end of 2001 there seemed to be an over saturation and the meme became infertile.

The state of the meme now is not to be considered 'dead' but rather 'dormant'. Several websites still carry the meme and the dormant state may become active again as sudden as it started in 2000 (Murray).

3.2.4. Expression and Transmission

Not everything which is retained in memory is also spread again. Heylighen describes this stage as the transition of a meme from a "memory pattern" into "a physical shape that can be perceived by others" (Heylighen, What makes 2). Important in this stage is the ability of being perceived by a potential new host. This requires the host to choose a valid representation of the meme. As shown in the example of the fifth symphony of Beethoven (2.8) not all ways of expressing a meme are as successful as others. The way of expressing a meme has to be chosen by the host but this is not necessarily a conscious action. Heylighen gives the example of memetic gestures which the host may show without even realizing it.

When the right way of expression has been chosen, the next phase is the transmission of the meme. The transmission gives the meme a "physical carrier or medium which is sufficiently stable to transmit the expression". During the transmission stage the meme is actually sent from one host to one or more potential hosts. Even at this point the meme is subject to selection. Selection at this point can either occur through the destruction of the meme's vehicle or by interference (e.g. noise during broadcasts) (Heylighen, What makes 2-3).

The more a meme is transmitted, the more chance there is of it reaching at least a few hosts. Multiple transmission can be separated into either mass transmission, which means that one host reaches several other hosts at once (for example broadcasting, mailing lists, etc.) and repetition. To be able to use repetition as a way of multiple transmission it is necessary for the meme to succeed in being retained in the memory of the host (see 3.2.3). The shorter the time that a meme can stay in the memory of the host, the more likely it will be that there will only be a few transmissions of this meme.

3.3. Social and Cultural criteria for meme spreading

3.3.1. Introduction

Several scientists have touched upon memetic selection criteria. Castelfranchi focuses on the social and cultural mechanisms of cultural transmission whereas Heylighen is more interested in the criteria that influence both the memes and the hosts.

Culture is to a large extent a bundle of rules and regulations. It is one of the factors that separates humans from other species. To historians culture started as soon as hominids began to use tools. Tool making can be seen as a cultural phenomenon or even a meme. When an individual profits from the tool it created, it is likely that others will copy this behavior and from this behavior early human interaction emerged. In Humanity's Descent, Rick Potts analyzes tool making. The necessary knowledge for the creation of tools consists of several factors which he gathers in the '*culture Monolith*'. For him culture is made out of the abilities for transmission, memory, reiteration, innovation, selection, symbolic coding and institutions (181-204).

These abilities are not so much different from what a host needs to be able to transmit a meme. Transmission, memory, reiteration and selection are part of the stages of memetic replication (see 3.2). Innovation and selection are processes which are necessary to be able to speak about replicators (cf. Dennett's criteria) and symbolic coding makes up for language or the physical shape in which a meme is being spread.

When there was no longer only one tool or habit but multiple ones, it became, according to prehistorians and anthropologists, necessary to communicate in a different way then only by imitation; according to some scientists, this was the initial stimulant for human language.

When primatologists noticed that some primates copied behavior from others and even learned through instruction, the need for a new definition of culture arose. Culture was divided in proto-culture and culture. This way human culture could easily be distinguished from the cultural behavior of other primates. To classify as proto-culture, a species needs to be able to copy behavior, culture required not only copying and spreading but also social symbolic representations. (Potts 181-204).

To be able to speak about culture, it is necessary that hosts can in some way communicate the specific aspects of culture. This can be done by cultural transmission or the replication and transmission of cultural conventions. Castelfranchi distinguishes three mechanisms for cultural transmission: instrumental adoption, normative adoption and identity-based adoption.

3.3.2. Instrumental Adoption

Instrumental adoption of memes helps an individual to solve a certain problem. When the meme helps to solve the problem it is positively reinforced (Castelfranchi 6-7). This is what Heylighen calls the utility factor of a meme (see 3.4.3.4). This does not imply that the meme really does solve this problem in a logical way. A rain dance may not really help to bring rain to areas which are suffering from a drought but if the host thinks that this will help, and if the local healer or medicine man has a good barometer to know when he needs to induce the rain dance, it will seem logical to the host that both the action and the result are related.

3.3.3. Normative Adoption

Through education and upbringing, an individual learns the rules of the society (cf. Pott's institutions) he or she lives in. A part of culture is this set of expectations and prescriptions. Castelfranchi distinguishes between memes and norms. He defines norms as meta-memes which are "memes for regulating the conformity to and the spreading of behaviors, goals, ideas; they provide Models and prescribe who should or can do what" (8). This is the limiting function of culture which filters certain memes and gives them the notion of social acceptability. If an individual accepts a certain meme, this is then not only the acceptation of this meme as belonging to the culture, but also accepting the norm which has granted this meme to be copied.

Normative adoption of memes is the adoption of a meme because the agent wants to conform to the expectations of the society in which it lives. The instruction of the norm does not need to be direct. Through observation an agent will recognize which memes and norms are already parts of the culture (Castelfranchi, 7-8).

3.3.4. Identity-based Adoption

Social and cultural identity play an important role in human interaction. There is a basic urge to identify oneself and to belong to a certain group. This social identity mechanism can be the reason why a host adopts certain memes. This requires the ability of this host to be able to recognize the memes (speech pattern, gestures, etc.) of the group to which it desires to belong.

A theoretic example of this behavior is the use of certain language features that belong to a higher social status by members of a lower social status. By doing so, they are identifying themselves with the higher social class and thus reinforcing the specific language behavior of this class. This can then lead to a situation where the higher social class recognizes that the particular feature has become a characteristic of the language of the lower social class and trigger a reverse mechanism (which Castelfranchi calls Social Identity Hostility (11)) by which the higher social class will distinguish itself from the lower class.

3.3.5. Drawbacks of the Socio-Cultural criteria

Castelfranchi's criteria are very valuable for anthropological, sociological and psychological research. For a quantitative research however they are not very useful as it is hard to objectively measure someone's urge to belong to a certain social category or the cultural pressure of a behavior. To be able to do a quantitative research it is necessary to use criteria with a limited definition which do not intermingle as much as Castelfranchi's criteria.

3.4. Heylighen's Selection Criteria

3.4.1. Introduction

In contrast to Castelfranchi, Heylighen does not solely focus on the host's point of view. He distinguishes four general criteria families: objective, subjective, intersubjective and meme centered. These four families each focus on a different aspect of memetic selection and the memes on their own (What makes).

The objective criteria are independent of the memes or the potential hosts. He distinguishes three criteria: distinctiveness, invariance and controllability. Although these criteria group together as objective, they do not have an influence on the same stages of memetic spreading.

The subjective criteria are highly dependent on the potential host. It is these criteria that are subject to change and are different for almost every meme-carrier. They form the basic selection criteria that are applied when a host encounters a meme. They are novelty, simplicity, coherence and utility.

Intersubjective criteria have an impact on the behavior of several individuals together. They deal with more general factors that are often set within (sub)cultures. Authority, formality, conformity, expressivity and publicity fall into this group.

The last group of criteria is the meme-centered ones. These criteria reflect on nothing but the meme itself. From the meme's view, they are the important criteria that will directly deal with enforcing the meme and urging it to be spread by the carrier. These criteria are self-justification, self-reinforcement, intolerance and proselytism.

3.4.2. Objective Selection Criteria

3.4.2.1. Distinctiveness: The more a meme can be seen as a unique entity, clearly different from another entity, the more it is likely to stand out. When a meme stands out, the probability of being noticed is much higher then when it blends in with other memes. This is called distinctiveness.

3.4.2.2. Invariance: If a meme can be noticed by more than one of the major senses (sight & hearing), it has more chances of finding a host. Kelley breaks invariance up in three parts, invariance over modalities, time, and persons. The invariance over modality reflects on the possibility to notice a meme with more than one sense. Invariance over time means that the longer a meme can be perceived, the more likely it is to be a good meme and not to be erased because of other memes replacing the niche that this meme had filled. Invariance over persons means that when the meme is noticed by different people it adds to the credibility of the meme (*qtd. in* Heylighen, Selectors of Knowledge 3).

3.4.2.3. Controllability: Heylighen describes this criterion as 'reacting differently to different actions performed on it'. Adding controllability provides 'methods of experimentation and of preparation-detection, which characterize science' (Selectors 4).

3.4.3. Subjective Selection Criteria

3.4.3.1. Novelty: Novelty will catch attention. In advertising products are always new and the buzzword 'new' is sure to attract some attention. Grabbing attention from a possible host is very important to a meme; failing to do so will make the possibility of being selected for retention and thus for finding a new host unlikely.

3.4.3.2. Simplicity: Although this criterion is focused at the meme itself, it is still a subjective one. The grade of simplicity is based on the knowledge of the potential host, the perceiver. When a meme can be adopted easily, it will find less resistance to become adopted as it will require a minimal effort from the host.

3.4.3.3. Coherence: When a meme can relate to beliefs, ideas and knowledge of the potential host, and does not contradict these, it will be easier to be accepted by the host. The meme, once adopted, will become part of an already available memeplex. It will then not only reinforce itself but it will also support the memeplex it has become a part of. When a meme needs to fit into already existing beliefs, it will generate obstruction from the memes that are already hosted and chances of survival amidst these opposed memes will diminish.

3.4.3.4. Utility: Applying directly to the functionality it can offer to the host, the utility norm reflects the usability of the meme. Does the meme add any knowledge to the host that will prove to be useful? This system of reward-after-retention will mutually benefit the host and the meme itself.

3.4.4. Intersubjective Selection Criteria

3.4.4.1. Authority: When the host of the meme is considered a trusted source, the potential new host will be more likely to accept a meme because this meme is backed up by the subjective feeling of safety. Castelfranchi calls this criterion the *reliability* and *credibility* of the meme (6.2). Examples of this can be found in teaching; for example, where students will be more likely to question ideas brought up by their peers than theories explained by their instructor. The social function of the carrier, in this example the instructor, will imply that due to its function, the carrier will only provide correct information. This criterion is dependent on time and space. The same meme-carrier can have a different level of authority depending on the situation.

3.4.4.2. Formality: The degree of formality in the intersubjective criteria is somehow misleading. The formality as used by Heylighen & Dewaele is not an equivalent of the grammatical idea of formality. Formality here is used to measure the possibility of misinterpretation. Higher formality means that there is a smaller chance of misinterpreting the meme and therefore a bigger chance that the meme is properly copied (qtd. in Heylighen, Selectors 5).

3.4.4.3. Conformity: When a meme is popular and *goes around*, it will be selected by a large number of hosts. Marketing methods that use viral marketing use this strategy in order to make sure that the customer has heard what the (publicity) company is trying to say or sell. The meme will receive a certain status and to obtain the status linked to the meme, the meme-carrier needs to incorporate this meme in its memory (retention) and spread this particular meme. This leads to an exponential spreading of the meme, until it will lose its status.

3.4.4.4. Expressivity: If a meme can easily be expressed, it will facilitate the spreading of this particular meme. If the host needs to make a considerable effort to spread the meme, the spreading of this meme will be much more unlikely.

3.4.4.5. Publicity: A meme which is easy to spread will be spread more often, in analogy with the expressivity criterion.

3.4.5. Meme-centered Criteria

3.4.5.1. Self-justification: The meme-centered criteria are solely dependent on the meme. Self-justification expresses the "*degree to which the components of a meme mutually support each other*" (Heylighen, What makes 4).

3.4.5.2. Self-reinforcement: If a meme contains an instruction towards the host to be repeated, it will reinforce itself with the host. Reinforcement can be found e.g. in prayer. (Heylighen, What makes 4).

3.4.5.3. Intolerance: Intolerant memes have a defense mechanism against potentially dangerous memes. Once an intolerant meme is in place, it will make it harder for rivals to find a place within this meme-carrier.

3.4.5.4. Proselytism: A high degree of proselytism will force or encourage the host to spread the meme more. The higher the proselytism factor, the more chance that the meme will find extra hosts.

4. Research Methodology and Corpus

4.1. Introduction

As mentioned earlier, one of the great steps for memetics to take is to find a way to put all theories into practical usage. A theory can be useful to think about a certain subject in a new way but one needs to be able to put the new ideas to the test. This research will try to find out if memetic spreading, and more specifically the importance of several selection criteria, can be tested.

To narrow down the wide field of memetics, this project will focus on virus hoaxes only. This way it is possible to create and select a small corpus of which all elements can be compared to each other. There are several specialized libraries of virus hoaxes which make it easier to select them and which reduce the time needed to gather the data.

The selection criteria are based on the ones as described by Francis Heylighen in What makes a meme successful? Selection Criteria for Cultural Evolution. These criteria (as described in 3.4) will be tested through a survey by asking respondents to scale each criterion.

4.2. Goal of the Research

Based on the theory of memetic spreading, this research will look at several specifically selected criteria to find out which criteria are most helpful to the hoax to encourage its spreading. By making a survey which asks the participants to rank several criteria on a scale of one to five, it will be possible to make a ranking of these criteria by correlating their importance to the fitness of the hoaxes.

It is important to emphasize that the outcome of the research will not give a ranking of importance of these criteria for all memes. Virus hoaxes are only one of several types of memes and criteria which are important for hoaxes may not be important to the other kinds of memes. This research will therefore not be able to make a generalized ranking of all criteria for all memes but will only look at a limited number of criteria tested on a limited set of virus hoaxes.

The result of the survey will then be compared to the results of an expert in memetics. This specialist has rated the six criteria of the same hoaxes as the users. This was done using the same methods as the users have used (an online survey) so there was no bias towards the authority.

4.3. Quantitative Research

When contemplating a new field of research such as memetics, a theoretical foundation is needed before one can attempt to test the hypotheses set out by the theoreticians. When theories have been formed they can be tested by doing quantitative research.

The reception of a meme by a possible host is a very subjective factor in the process of meme-spreading. Several potential meme-carriers will have different opinions or different preferences. With this survey it is hoped to be able to collect a sample of data on the six criteria that will be large enough to determine the preferences of an average population. If the sample is large enough, the average result will show a reliable representation for the population as a whole, regardless of individual differences.

There are different methods to carry out a quantitative research project such as interviews and surveys. Because this particular topic is closely linked to the Internet, an online survey was chosen over live interviews. The advantage of a survey over interviews is that there is less chance of a deliberately false answer, based on social expectations. Using computer display makes it possible to give a representation of the hoax as it would appear in a participant's mailbox, including the capitalization and grammatical or spelling errors. As there is no time-pressure in this survey, the participants have all the time they need to re-read the questions and the hoaxes.

By administering an online survey it is possible to gather a large data sample. This is more time consuming if it is done by taking interviews and the latter requires more time for the digitalization of the results of interviews as well.

4.4. Choosing the hoaxes

4.4.1. What are hoaxes?

A hoax is a prank, a message without any real information contained in it. Hoaxes try to deceive people and can take numerous forms. Most hoaxes have no intentional harm but some types of hoaxes can have negative results.

Negative hoaxes can result in panic, scares or even damage to certain systems. Examples of hoaxes like this are the scares about drug-filled candy or about possible negative effects of certain kinds of food (for example the 'aspartame scare' which warned about possible multiple-sclerosis symptoms after consuming the artificial sweetener aspartame). Virus hoaxes also fall into this category.

Examples of innocent hoaxes are: April fools stories, practical jokes and urban legends. Hoaxes sometimes lead their own lives, such as the famous Nessie hoax, which remained alive for 60 years until it was finally unmasked in 1994.

In April 1934 newspapers in the U.K. reported of a monster in Loch Ness and had photographic evidence of this creature. The picture, commonly known as "the surgeon's photograph", became one of the most famous pictures ever taken at Loch Ness. It started a whole new area of research around the lake and brought many tourists to the Scottish highlands. Scientific research with sonar's and mini-submarines took place in the lake, movies were based on the monster, several scientists had theories about 'Nessie' and a fortune was made selling various items to Loch Ness fans.

In November 1993 Christian Spurling, grandson of the original photographer, revealed how the pictures were forged. Not a sea-monster but a toy submarine with some modifications was what was really on the picture.

This hoax was relatively innocent. One could remark that quite some money was spent on needless scientific research but on the other hand there has been a lot of revenue from the hoax as well. Other hoaxes such as chain letters and virus hoaxes are not always this innocent.

4.4.2. Virus hoaxes

A virus hoax is a message warning people about a so-called new virus. A few hoaxes are known to warn about viruses which directly affect humans but most virus hoaxes concentrate on computer viruses which could damage or steal useful information from one's computer. These messages spread all over the internet thanks to the means of electronic communication.

The features of virus hoaxes, according to Sophos, are that they:

- "falsely claim to describe an extremely dangerous virus
- use pseudo-technical language to make impressive-sounding (but impossible) claims
- falsely claim that the report was issued or confirmed by a well-known company

- ask you to forward it to all your friends and colleagues“ (jdbgmgr.exe)

In these messages people are often warned about certain emails which could contain a virus. Other virus hoaxes ask people to delete certain files from their hard drives. Deleting these files, which are often important for the security or stability of the computer system, can cause failures in the operating system and data-loss.

Another problem of these hoaxes is the amount of data-traffic that they can create. If a successful hoax is able to enter a certain company it can lead to a general slow-down of the network, using bandwidth that would otherwise be used for company purposes. This, in return, can lead to an overload of other systems within the company and can result in a loss of data or valuable resources.

These doom scenarios may seem over-dramatized but specialized helpdesks spend a large percentage of their time helping users to restore files that have been deleted because of virus hoaxes or explaining other innocent or fake warnings. The result of this is a reduction of the time which can be used for real problems and the frustration of many helpdesk agents.

Possibly the most dangerous result of hoaxes is that computer users may become so much aware of the fact that virus hoaxes are being spread that they will ignore real threats thinking that they too are hoaxes. Creators of computer viruses are aware of this. They also use the hoaxes by naming some of their viruses after them, hoping that this will mislead the users into thinking that there is no threat in a certain file where there actually may be one.

4.4.3. Selection of the virus hoaxes

The survey focuses on six different hoaxes. These hoaxes were selected using the lists of virus hoaxes as provided by various security software companies. For this research the websites of Sophos (www.sophos.com), Symantec (www.symantec.com), Mc-Afee (www.mcafee.com) and F-Secure (www.f-secure) were used.

Initially, eleven hoaxes were chosen based on information about their current spreading (as available on the aforementioned websites). These hoaxes were subjected to research on three major search engines: AltaVista (www.altavista.com), Alltheweb (www.alltheweb.com) and Google (both web & Usenet search) (www.google.com). To be able to determine the spreading (and therefore the success) of the hoaxes, two different methods were used.

Since hoaxes spread quickly over a wide range, many different versions of one hoax can be found. This sort of mutation and even recombination of several hoaxes shows the ability of hoaxes to have different phenotypes while actually they are related to the same original hoax. One of the problems this research faced was that as different hoaxes appear, it would be hard to check their spreading success. This is why several versions of one hoax were compared, as they can be found on the website of the software companies, and one reappearing string of words was used, as shown in table 1.

The result of this is that some versions of one hoax are lost, but by using this technique it was possible to focus not on one phenotype of a hoax but on several at once. This does not mean that the phenotypes that are not used in this research are less successful and therefore it is important to keep in mind that not all appearances of one hoax message are counted when comparing the final results. Nevertheless, the survey will give a good overview of the fitness of the used versions of the hoaxes if not of the overall fitness of one hoax due to the fact that these mutations exist.

From the combined results of all search engines the average of the natural logarithm was taken and it was this final result that was used to make up the ranking of the hoaxes. The results of the eleven chosen hoaxes (table 1) were then compared to find six hoaxes which had enough variation to test several grades of success in spreading. A correlation needed to be found so it was decided that it would be best to make three groups of hoaxes, a high spreading one, a middle spreading and a group of hoaxes that was not very successful but still successful enough to be used in quantitative research without having to classify the results as 'not trustworthy due to a too small sample'.

To avoid interference of two or more hoaxes over each other in the survey, the hoaxes were also filtered based on their content and methodology. For example the virus hoaxes sulfnbk.exe and jdbgmgr.exe both scored within the top three results of virus hoaxes found online but they contain the same type of warning; the first one targeting sulfnbk.exe, the second one demanding to delete jdbgmgr.exe. It is likely that if these two hoaxes would both have been included in the same survey, they would have a strong influence on each other as far as criteria such as novelty are concerned.

name	search query	Google	AltaVista	Alltheweb	Usenet
Good Times	“to the prospects of this newest creation”	448	274	401	963
	+”good times” +virus	28800	15445	168725	14500
Virtual card for you	“the virus destroys sector zero”	730	418	519	684
	“virtual card for you” +virus	4960	3651	27327	1640
Jdbgmgr.exe	“a teddy bear icon”	1670	844	623	1100
	+jdbgmgr +virus	17100	13612	101883	11000
Sulfnbk.exe	“will become active on June 1”	528	265	291	434
	+”sulfnbk” +virus	14100	10389	60082	19000
WTC Survivor	“DO NOT OPEN” + "WTC Survivor"	539	213	264	97
	+”WTC survivor” +virus	4400	8291	28279	166
Perrin	"upgrade internet2”	410	251	231	172
	+perrin +virus	11200	5559	43038	1860

Buddylst	“an apparently harmless Budweiser screensaver”	37	23	31	124
	+buddylst +virus	2130	1280	1765	2760
Wobbler	“with a file called “California””	315	154	185	247
	+wobbler +virus	3510	2231	19805	2220
Life is Beautiful	“life is beautiful.pps”	494	197	407	128
	+"Life is Beautiful" +virus	2460	557	33326	1180
Penpal Greetings	“destroying all of the data present”	581	408	484	1260
	+”penpal greetings” +virus	3030	2142	2815	4790
California IBM	“an e-mail titled “California IBM”	366	27	43	315
	+”California IBM” +virus	1680	405	1189	1570

Table 1: Search Results (03/07/2003)

4.5. The six used hoaxes

4.5.1. Introduction

This is the text of the hoaxes as they were used in the survey. It is important to notice that they are only versions of a certain hoax but it was necessary to limit to one version per hoax, as the research here is too small to count the differentiation in fitness between several versions of one hoax.

4.5.2. Jdbgmgr.exe

Jdbgmgr.exe was the most successful hoax used in this research. When the fitness of the hoaxes was measured, jdbgmgr.exe had spread all over the internet and had the highest proliferation of all hoaxes. Jdbgmgr.exe was first reported in April 2002.

I found the little bear in my machine because of that I am sending this message in order for you to find it in your machine. The procedure is very simple:

The objective of this e-mail is to warn all Hotmail users about a new virus that is spreading by MSN Messenger. The name of this virus is jdbgmgr.exe and it is sent automatically by the Messenger and by the address book too. The virus is not detected by McAfee or Norton and it stays quiet for 14 days before damaging the system.

The virus can be cleaned before it deletes the files from your system. In order to eliminate it, it is just necessary to do the following steps:

1. Go to Start, click "Search"

2.- In the "Files or Folders option" write the name jdbgmgr.exe

---

3.- Be sure that you are searching in the drive "C"

4.- Click "find now" 5.- If the virus is there (it has a little bear-like icon with the name of jdbgmgr.exe DO NOT OPEN IT FOR ANY REASON

6.- Right click and delete it (it will go to the Recycle bin)

7.- Go to the recycle bin and delete it or empty the recycle bin.

IF YOU FIND THE VIRUS IN ALL OF YOUR SYSTEMS SEND THIS MESSAGE TO ALL OF YOUR CONTACTS LOCATED IN YOUR ADDRESS BOOK BEFORE IT CAN CAUSE ANY DAMAGE.

4.5.3. Good Times

The oldest hoax message used was 'Good Times'. Good Times was first reported by CIAC (Computer Incident Advisory Capability) in 1994. Sophos distinguishing features of virus hoaxes can be found in this example: a) the threat of an extremely dangerous virus, as can be seen in paragraph one; b) the pseudo technical language as can be found in the fourth paragraph and c) the reproduction pressure in the last paragraph.

Thought you might like to know...

Apparently, a new computer virus has been engineered by a user of America Online that is unparalleled in its destructive capability. Other, more well-known viruses such as Stoned, Airwolf, and Michaelangelo pale in comparison to the prospects of this newest creation by a warped mentality.

What makes this virus so terrifying is the fact that no program needs to be exchanged for a new computer to be infected. It can be spread through the existing e-mail systems of the InterNet.

Luckily, there is one sure means of detecting what is now known as the "Good Times" virus. It always travels to new computers the same way - in a text e-mail message with the subject line reading simply "Good Times". Avoiding infection is easy once the file has been received - not reading it.

The act of loading the file into the mail server's ASCII buffer causes the "Good Times" mainline program to initialize and execute.

The program is highly intelligent - it will send copies of itself to everyone whose e-mail address is contained in a received-mail file or a sent-mail file, if it can find one. It will then proceed to trash the computer it is running on. The bottom line here is - if you receive a file with the subject line "Good TImes", delete it immediately! Do not read it! Rest assured that whoever's name was on the "From:" line was surely struck by the virus. Warn your friends and local system users of this newest threat to the InterNet! It could save them a lot of time and money.

4.5.4. Penpal Greetings

The first notion of Penpal Greetings came in March 1999. The hoax warns about an email with the subject ‘penpal greetings’ and about data loss. It has a whole paragraph dedicated to nothing but replication pressure.

If anyone receives mail entitled: PENPAL GREETINGS! please delete it WITHOUT reading it.

This is a warning for all internet users - there is a dangerous virus propogating across the internet through an e-mail message entitled "PENPAL GREETINGS!". DO NOT DOWNLOAD ANY MESSAGE ENTITLED "PENPAL GREETINGS!" This message appears to be a friendly letter asking you if you are interested in a penpal, but by the time you read this letter, it is too late. The "trojan horse" virus will have already infected the boot sector of your hard drive, destroying all of the data present. It is a self-replicating virus, and once the message is read, it will AUTOMATICALLY forward itself to anyone who's e-mail address is present in YOUR mailbox!

This virus will DESTROY your hard drive, and holds the potential to DESTROY the hard drive of anyone whose mail is in your inbox, and who's mail is in their inbox, and so on. If this virus remains unchecked, it has the potential to do a great deal of DAMAGE to computer networks worldwide!!!!

Please, delete the message entitled "PENPAL GREETINGS!" as soon as you see it! And pass this message along to all of your friends and relatives, and the other readers of the newsgroups and mailing lists which you are on, so that they are not hurt by this dangerous virus!!!!

4.5.5. WTC Survivor

WTC Survivor was first found in November 2001, a few months after the September 11 attacks on the World Trade Center, to which the hoax' title refers.

BIGGGG TROUBLE !!!! DO NOT OPEN "WTC Survivor" It is a virus that will erase your whole "C" drive. It will come to you in the form of an E-Mail from a familiar person. I repeat, a friend sent it to

me, but called and warned me before I opened it. He was not so lucky and now he can't even start his computer!

Forward this to everyone in your address book. I would rather receive this 25 times than not at all. So, if you receive an email called "WTC Survivor", do not open it. Delete it right away! This virus removes all dynamic link libraries (.dll files) from your computer.

PLEASE FORWARD THIS MESSAGE

4.5.6. California IBM

The California IBM hoax is first mentioned in hoax libraries in June 2000. Although this is the smallest hoax of the six hoaxes on which this survey focuses, an example of nearly all the features that Sophos describes is present: the claim of a false virus, the so-called confirmation of a well-known company, Microsoft in this case, and the pressure to pass this email to others.

Another new virus has been discovered. It arrives in an e-mail titled "California IBM". Microsoft has announced that it is very bad, worse than "Love Letter". There is no remedy or cure. It will consume all the information in the hard drive, and will destroy Netscape Navigator and Microsoft Internet Explorer. Do not open anything with this title, and pass this message on to your e-mail contacts. Right now not many people know about this, so please pass it on as quickly as possible.

Thank you.

4.5.7. Buddylst

According to CIAC, the Buddylst hoax was first found in October 1999. Several versions exist of this file of which the used one is the one most commonly found in the specialized hoax databases.

READ THIS!!!!!

This information came from Microsoft yesterday morning. Please pass it on to anyone you know who has access to the Internet. You may receive an apparently harmless Budweiser screensaver, entitled BUDDYLST.SIP.

If you do - DO NOT OPEN IT UNDER ANY CIRCUMSTANCES, but delete it immediately.

Once opened, you will lose EVERYTHING on your PC. Your hard disk will be completely destroyed and the person who sent you the message will have access to your name and password via the Internet.

As far as we know, the virus was circulated yesterday morning. It's a new virus, and extremely dangerous. Please copy this information and e-mail it to everyone in your address book.

We need to do all we can to block this virus.

AOL has confirmed how dangerous it is, and there is no Antivirus programme as yet which is capable of destroying it.

Please take all the necessary precautions, and pass this information on to your friends, acquaintances and work colleagues.

4.6. Selection of the Criteria

4.6.1. Introduction

The criteria that were chosen to be included in the survey needed to be easily comprehendible by the average survey-taker. From Heylighen's list, six criteria were chosen. These criteria were described as novelty, simplicity, danger, benefit, authority and replication pressure. In the introduction to the survey each of these criteria were described to give the survey-taker an insight on the meaning of each criterion.

The descriptions of the criteria were repeated with every question in the survey, with the extra information of the values that could be entered for the criterion (a scale from one to five).

> For example: "How easy is it to understand this message? Is it hard to grasp or is it pretty clear and simple? (1: Very Hard / 5: Very Easy)"

4.6.2. Novelty

The criterion '*novelty*' was renamed in the online survey as '*originality*' this was done in order to avoid a misunderstanding or confusion with the concept 'new'. The online description of originality was:

> Is this an original idea or have you seen messages like this before? Does the content of the message strike you as being unexpected or novel? (1: Not Original / 5: Very Highly Original)

4.6.3. Simplicity

Simplicity was not changed. Although dependant of the previous knowledge of the survey-taker this is a criterion that can easily be scored. The different background of the participants is compensated by the fact that each survey taker will score all hoaxes with the same previous knowledge. Simplicity was described as:

How easy is it to understand this message? Is it hard to grasp or is it pretty clear and simple? (1: Very Hard / 5: Very Easy)

4.6.4. Danger and Benefit

The two criteria danger and benefit are both sub-criteria from what Heylighen called utility. By separating these two it can be verified whether or not one of these two is more important.

The danger criterion measures the gravity of the danger it warns you for; it reflects the threat a certain message contains. (1: Low Danger / 5: High Danger)

The benefit norm reflects how much advantage one could get out of the message. Would the information in the message seem useful? (1: no benefit / 5: very high benefit)

4.6.5. Authority

The authority selection criterion should be one of the least important criteria in this survey. All of these hoaxes will arrive by email from a known source, so it is questionable if the receiver of the email will try to find an authority source within the meme, other than the sender of the email. Nevertheless, for this research it is useful to check for the correlation between the authority of the hoax itself as well although it has to be kept in mind that the original authority (from the sender of the email) is not present in this survey.

Authority reflects the degree to which this message appears to be backed up by a trusted and reliable source or expert. (1: Not trusted (no source) / 5: Very Trustworthy)

4.6.6. Replication Pressure

The criterion of replication pressure or proselytism can logically be considered one of the most important ones in the spreading of the hoax. It is this criterion which will try to influence the receiver of the hoax to spread it. However, it is likely that a too high replication pressure will not lead to a higher replication rate but to a ridicule of the hoax. A hoax that consists only of 'please pass me on' messages may not be passed on due to the lack of content because people simply do not take it seriously (Hofstadter 52-53).

The degree of replication pressure indicates in how far this message urges you to spread it further and communicate it to other people. (1: No demand / 5: Very high demand)

4.7. Creating the Survey

4.7.1. Technical Aspects

An online survey is automatically limited to the availability of survey methods for the internet. Several commercial packages and specialized websites offer ready made tools for commercial surveys. For this survey a program called PHPESP was chosen.

PHPESP is a free program that is developed by members of the Sourceforge community. PHPESP is written using the programming language PHP and a MYSQL database server to create online surveys. The open source of the program and the fact that several people are working on it at once means that the development of the survey was not always easy as updates and new versions of the software were released while the survey was created.

However, this also meant that errors and problems with the survey program were always dealt with very quickly. The open source factor of this survey package is especially noteworthy as professional survey creation packages are often extremely expensive, or as a review of survey software packages recently wrote, "Survey applications and services are often so expensive, however, that they cost more than the combined mail and Web infrastructure required to manage them" (Rapoza).

From the example hoaxes as provided by the virus information databases from several software houses (see 4.4.3) the format that most frequently appeared (i.e. that appeared in more than one virus information center) was chosen.

Each of the hoaxes was to be valued on the six criteria. The survey taker was asked to grade each of the criteria on a scale from one to five. There was no default value given to the criteria so a choice had to be made before the participant could continue with the following hoax.

4.7.2. The Surveys

When creating a survey one has to be aware of possible fluctuations in attention span of the survey taker. Questions which are asked earlier in the survey tend to get more attention than questions at the end of the survey.

To make up for this variation, three different versions of the survey were created. The six hoaxes that were selected for examination (see 4.5) were divided into three groups of two hoaxes each. The criteria for these groups were that they could not contain two hoaxes from the same success-category of proliferation. The three proliferation groups were high, middle and low proliferation, based on the results of the searches as described in 4.4.3. The first pair consisted of a low scoring hoax, California IBM and a hoax which scores average on fitness, Penpal Greetings. The second group is made up of one high scoring, Jdbgmgr.exe and one low scoring hoax, Buddylst. The third combination has an average scoring hoax, WTC Survivor combined with a high scoring hoax, Good Times.

These groups of hoaxes were placed in three different subsequent orders to ensure that each group was placed in the first part of the survey, the middle part of the survey and the end of the survey. Because of the way the pairs were organized a low scoring (California IBM), a middle scoring (WTC Survivor) and a high scoring hoax (Jdbgmgr.exe) were at the beginning of the survey and that a low scoring one (Buddylst), a hoax with an average fitness (Penpal Greetings) and a high scoring hoax (Good Times) were at the end of each group in one of the surveys.

The exact spreading of the hoaxes is shown in table 2.

Group 1	Group 2	Group 3
California IBM Penpal Greetings	Jdbgmgr.exe Buddylst	WTC Survivor Good Times
Jdbgmgr.exe Buddylst	WTC Survivor Good Times	California IBM Penpal Greetings
WTC Survivor Good Times	California IBM Penpal Greetings	Jdbgmgr.exe Buddylst

Table 2: Groups of questions

4.8. Statistical Analysis

The results of the surveys are gathered per hoax and per question. The answers to the 36 different questions were imported into a spreadsheet program, Excel. By means of the data analysis module it was possible to make a statistical analysis and to compare the answers with the results from the search engine test in order to make a correlation between those two figures.

When correlation is measured it is important not only to pay attention to the positive correlation but to also keep the negative correlation in mind. A positive correlation will, in this research, indicate that the meme's fitness can be related to the criterion, a negative correlation will mean that the spreading degree of the meme is also dependent of this criterion but in an opposite way.

If the study would show that a high simplicity rating would correlate with a high fitness level of the meme, this would mean that the simpler the structure of a hoax is, the more chance this hoax has of being spread. If there would be a strong negative correlation this would mean that a hoax needs a complex language structure to have a higher spreading potential.

5. Research Results

5.1. Thesis Hypothesis

5.1.1. Introduction

The hypothesis which this study would like to confirm is that the mentioned selection criteria will have a clear correlation with the success rate or fitness of the hoaxes. Criteria such as the ones that are mentioned in the research are thought to influence the spreading of the hoax.

With the results of the study it would be possible to make a limited formula to predict the success rate of a previously unknown hoax by scaling it on the researched criteria. This formula could then look like this:

High originality + average simplicity + low danger + high benefit + high authority + average proselytism = high fitness

The initial hypothesis was based on the results of the survey as filled out by an authority in the field of memetics. The expert used for this research is Francis Heylighen. He is the author of the original definitions of the criteria.

5.1.2. Expert Results

The results of the survey as done by the expert in the field of memetics did not confirm the importance of all criteria. Instead three criteria stood out from the rest. Novelty, with a positive correlation of +0.5 and, even more so, benefit with a positive correlation of +0.8 seem to be the most important criteria. Simplicity has the lowest correlation with -0.6 to the fitness of the meme in both Search Engines and Usenet newsgroups and -0.53 when the correlation is measured with only the search engines. This means that if a hoax is perceived as being very simple, it will not reach a high degree of spreading.

	Search Engine	All Searches
Novelty	0.57	0.50
Simplicity	-0.54	-0.61
Danger	-0.47	-0.40
Benefit	0.82	0.88
Authority	-0.55	-0.50
Proselytism	-0.34	-0.43

Table 3: Expert Correlation with the six chosen hoaxes

From the expert's results it can be concluded that a hoax will be successful if it has a high degree of both novelty and benefit and if it has a complex structure. The correlation therefore has to be: +novel + beneficial –simple.

As the expert did not only score the six hoaxes that were used in the public survey but also the other 5 hoaxes, it was possible to see if a change in the number of hoaxes would give significant changes in the correlation. It is to be expected that a small difference should occur if more data are added and the results of such a correlation test show that there is indeed a deviation from the result of the six chosen hoaxes but the general trends (the positive correlation of both novelty and benefit and negative correlation of the other criteria) remain the same.

	Search Engines	All Searches
Novelty	0.29	0.25
Simplicity	-0.52	-0.52
Danger	-0.34	-0.31
Benefit	0.68	0.76
Authority	-0.09	-0.13
Proselytism	-0.20	-0.21

Table 4: Expert Correlation with 11 Hoaxes

The results in table 4 show that the overall correlations are weaker if more hoaxes are chosen. This is not only true for the negative correlations, which in the case of authority becomes almost insignificant, but also in the positive correlations, where the score for benefit drops considerably.

5.1.3. Research Goals

The survey's first goal will be to find out if it is possible to confirm the results of the authority. This means the importance of mainly the factors Novelty, Simplicity and Benefit. The second goal is to find out why certain criteria are more important then others and what might have influenced the participants to score a certain hoax. A third focus of this survey was to find out if there is a significant difference between the importance of danger and benefit which are both part of the utility criterion.

5.2. Hoax Results

The hoaxes fitness was measured with the technique described in the previous chapter. As the results of the searches were too widespread due to the proliferation behavior of hoaxes, the natural logarithm was taken of the final results. The ranking of the hoaxes by logarithmic can be found in table 5.

	Average All Searches
JDBM	16.92
Good Times	16.53
Penpal Greetings	14.45
WTC Survivor	13.64
California IBM	11.65
Buddylst	11.30

Table 5: Spreading of the Hoaxes (All Searches)

The average of all search engines was roughly the same as the average of all searches done (including newsgroups). There is however a small difference; Penpal Greetings showed to be more successful if the search engine tests were left out and WTC-survivor was less successful.

A possible cause for this difference is the limited availability of sources for the newsgroup search. There was no indication of how far back the archives of the newsgroups went or which newsgroups were included in the database. This is why the correlations of the hoaxes with the criteria are done with the search engines only. It is, however, interesting to keep the result of the searches with the newsgroups included in mind.

	Average Search Engine
JDBM	17.13
Good Times	16.55
WTC Survivor	14.96
Penpal Greetings	14.06
California IBM	11.16
Buddylst	10.82

Table 6: Spreading of the Hoaxes (Search Engines only)

5.3. Survey Results

5.3.1. Participation

It is impossible for a survey on a topic like this to question everyone who is actively involved in the spreading of the hoaxes. This is why the focus lays on a specific audience. The survey takers were all volunteers of the V.U.B. student community. By using this group, several important factors could be measured. The age category of the participants, which will for every student be somewhere between 18 and 25 years (on average) and language background are controlled. The educational background will differ slightly according to both the field of the participant and the progress which each individual survey taker has made in this field. However the educational level is limited to participants who have successfully finished high school.

It is important to note that the majority of V.U.B. students are not native speakers of English; this limitation also guarantees that there is no strong influence of native speakers over non-native speakers. It would, however, be interesting to do further research and to compare the results of students with English as a foreign language to students whose native language is English.

From a total of 280 students who started filling out the survey, 195 people actually finished it. This means that roughly 70% (69.6) of all survey takers finished the survey and answered all 36 questions. For the final average scores, only the results of those who completed the survey are used to avoid filtering out the order bias by counting the unfinished answers.

One of the reasons why people did not finish the survey may be the language difficulty. The participants are non-native English speakers whereas the survey is only available in English. Nevertheless the number of respondents is a large sample size and should provide an insight on the importance of the different criteria.

5.3.2. Data Significance

When gathering data it is always necessary to make sure that the data are not only correct but also trustworthy. For online surveys one does not know whether or not the participants were answering truly or whether they made up the answers just for fun. To be able to judge this, all the answers on each question were made into one average score. Based on this average score the standard deviation of the answer was calculated. The hypothesis was that if people would fill out the survey at random, the standard deviation would be large whereas if the participants would answer truthfully, this deviation would be smaller. The average standard deviation for all criteria is 1.17.

The three different versions of the survey were used to check the validity of the results. A t-test was performed between each version for each question in the survey to find out if there were any signs of significant differences in the results of these questions. In 91% of these tests no significant difference was observed between the groups. In the 9% where there was a significant difference, this difference did not always occur in the same survey or the same question. Because of this the differences are not scientifically important to this research and can, possibly, be explained.

While looking for the reason why a few scores showed a significant difference, it was noticed that the questions which appear sooner within a survey have a slightly deviating score from the same hoaxes in other surveys. This can be explained by the fact that these hoaxes can not be scored in comparison with the earlier hoaxes. However it is never a deviation between the three groups but only when two of the different groups are compared to each other. By taking the average of the total results of the three different groups these aberrations in the scores are nullified.

The idea of creating these three groups was to make up for a decline in attention given to the hoaxes. However this also makes it so the criteria are differently scored in the beginning of the test. The fact that the first hoaxes encountered in the survey are not always the same creates a different order of comparison.

5.4. Correlations

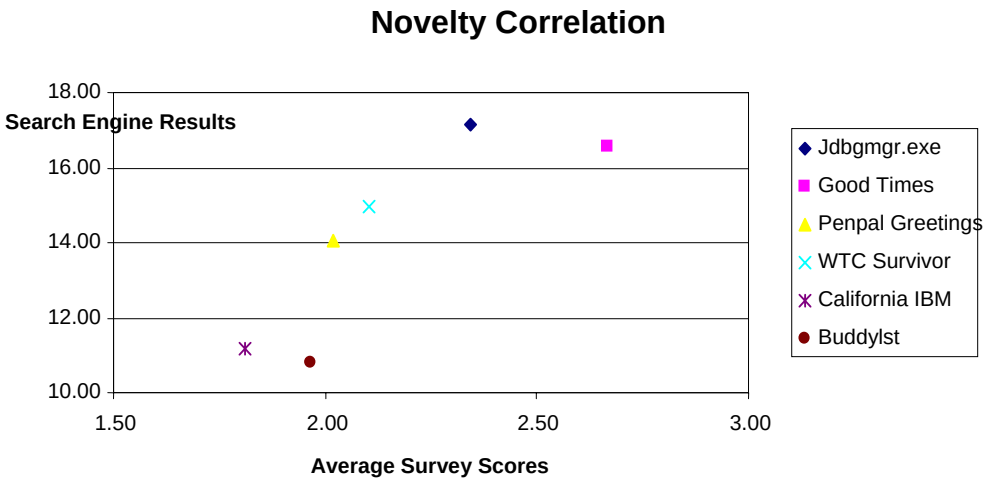
5.4.1. Introduction

The best way to describe correlation is by calculating the correlation from the numerical data that were gathered and by using a visual displaying of the correlation. The technique used for the visualization of the correlation is a scattergraph. By using this method of visualization it is easy to see both positive, negative and zero correlations.

The correlations are not only displayed and calculated but they are also analyzed by means of a comparison with the authority results and, where possible by other methods. For every criterion a possible explanation of the result is presented.

5.4.2. Novelty

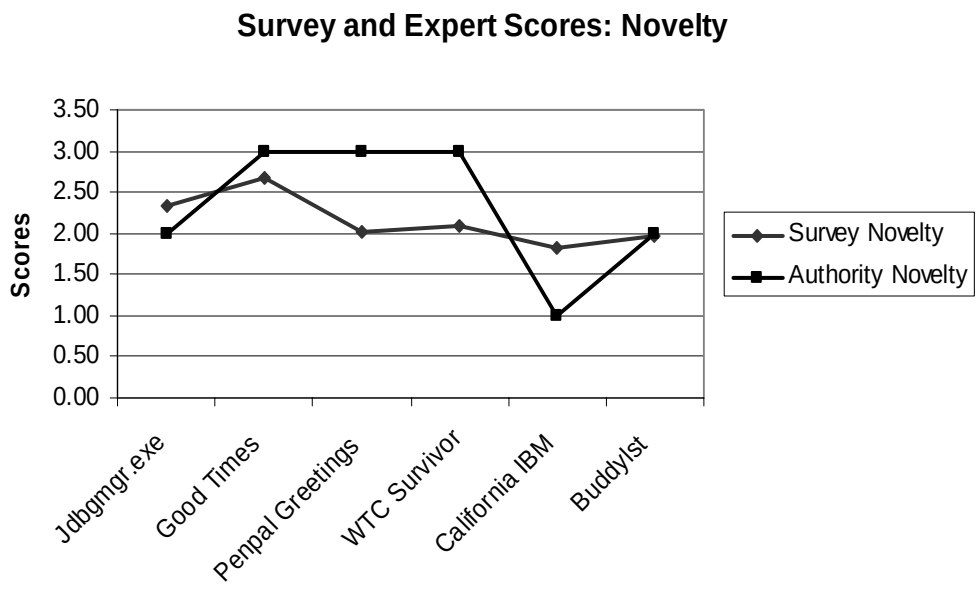
The correlation between Novelty and the degree of spreading is a strongly positive one. One of the explanations for this correlation is that many virus hoaxes can be debunked because of the fact that most of them have the same structure and/or content. When a new type of hoax emerges it is not immediately recognized as a fake message whereas a hoax that uses the same appearance and methodology of older hoaxes can be discovered quickly because the pattern is already known to the users.



Graph 1: Novelty Correlation

The first graph shows the correlation between the average user score and the search engine results. This shows that indeed the high scoring hoaxes (Jdbgmgr.exe and Good Times) score high on novelty whereas the hoaxes with lower fitness (California IBM & Buddylst) have low scores on novelty.

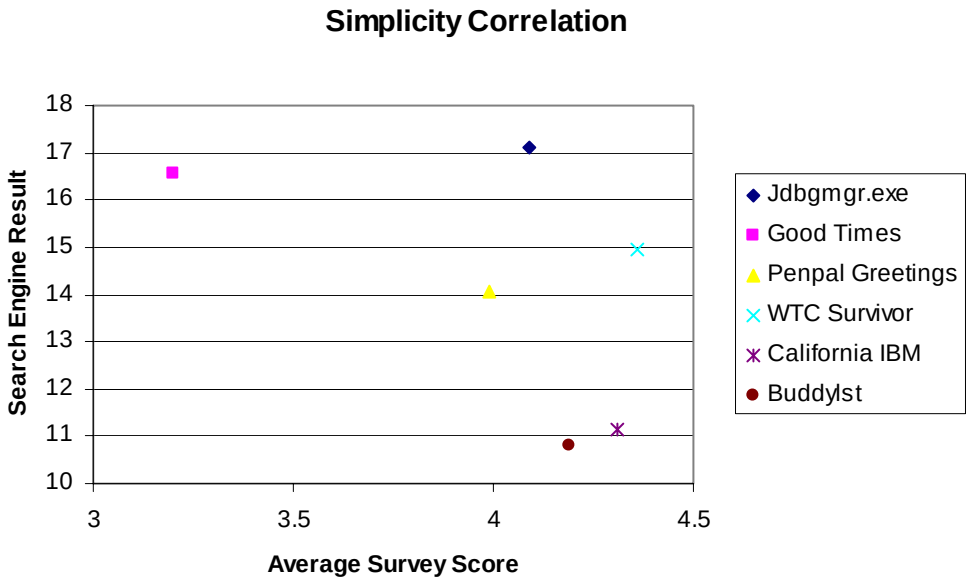
When the scores as given by the memetics expert are compared to the user scores on novelty, it shows that both have a positive correlation. The authority in the field of memetics comes to a positive correlation of 0.57 whereas the correlation of the survey participants is 0.83. The difference of 0.26 is mainly because of the differences between the scores for Penpal Greetings and WTC Survivor which are both scored at 3 by the expert but only have averages of 2.02 and 2.10 in the users survey, as can been seen in the second graph, comparing both users scores and the scores of the expert. The main reason for this difference is that novelty depends on the background of the person scoring the hoaxes. The authority score is based on the background of one person and is therefore very subjective whereas the results of the users are an average of the 195 participants.



Graph 2: Survey and Expert Scores: Novelty

5.4.3. Simplicity

The initial hypothesis was that a high simplicity would help the spreading of the hoax, as it would then provide easy understanding to the host as well as an easier adoption in the mind of the potential host. This research has shown that a high simplicity does not necessarily lead to a high degree of fitness.



Graph 3: Simplicity Correlation

The hoaxes with the highest simplicity are WTC Survivor and California IBM. These hoaxes are from the middle and lower group when considering the proliferation of the hoaxes. The hoax with the highest degree of spreading, Jdbgmgr.exe scored only 4.09 on average for simplicity; this is still a high degree of simplicity but it was ranked fourth when compared to the other hoaxes. Correlation versus the search engine results is negative 0.51 whereas the correlation verses the average of all searches is negative 0.63.

A possible explanation for the negative correlation between the survey results and the degree of simplicity is that the participants may find it easier to see through a simple hoax. When the hoax scores low on simplicity, this means it is harder to understand. A hoax which warns for a virus in just a few words may lack the credibility of a hoax which gives a full explanation of what might happen. The Good Times hoax for example uses no special capitalization, uses full sentences to warn about the possible dangers of this so called virus and has strong cohesion within the hoax, using in-text references; both cataphoric and anaphoric.

There are different ways to measure the simplicity of the hoaxes. One of the ways to measure the simplicity is by asking respondents to score the simplicity based on how easy they think it is to understand the message of the hoax. Another way is to calculate a readability score. An example of a technique to measure readability is the Flesch-Kincaid grade level formula. This calculates the difficulty of a text by scoring it to a certain education grade that is needed to comprehend the text. The grades are based on the American educational system. By using this technique it is possible to check the results of the survey against an established scoring technique. This is a different way of verifying the results of the survey and to show the credibility and significance of the data.

To measure the Flesch-Kincaid grade level of a text it is necessary to count two different variables. The first variable is the average sentence length; the second variable is the average word length. The exact formula to measure the grade level is:

$$(.39 \times ASL) + (11.8 \times ASW) - 15.59$$

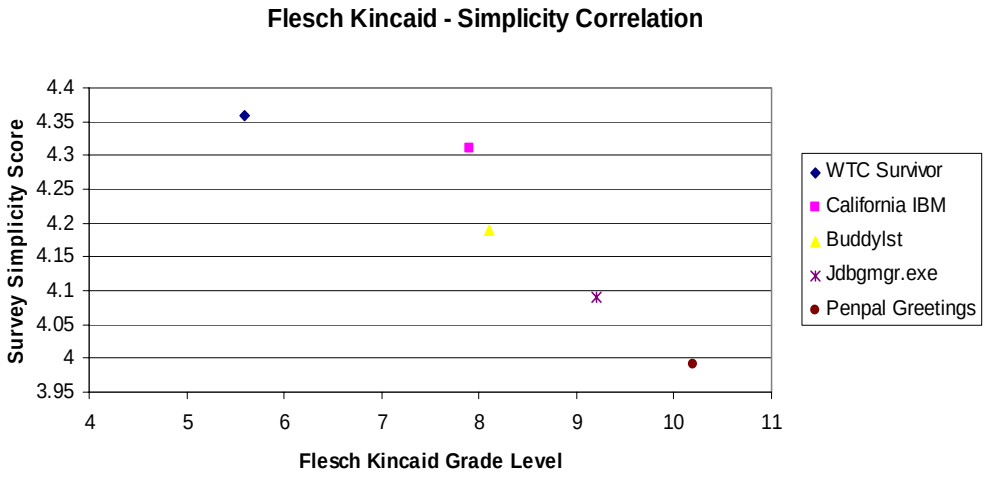
If we calculate the readability of the hoaxes according to this formula we see that there is a negative correlation between what the respondents score as most simple and the hoaxes that score lowest on the grade level.

	Grade level	Survey Score
WTC Survivor	5.6	4.36
California IBM	7.9	4.31
Buddylst	8.1	4.19
Good Times	8.9	3.20
Jdbgmgr.exe	9.2	4.09
Penpal Greetings	10.2	3.99

Table 7: Flesch Kincaid - Simplicity Correlation

Good Times is the only hoax which really stands out in this table as being 'oddly scored'. This can be related to the length of the hoax. With a total of 15 lines (measured on a screen resolution of 1024x768 with the survey window on full screen) it was the hoax with the longest continuing text. Jdbgmgr.exe had one line of text more then Good Times but its simplicity was helped by the way this hoax is structured and gives a numbered list of actions that need to be taken for the removal of the jdbgmgr.exe file.

When Good Times is left out of the comparison between the grade level and the survey scores, the negative correlation becomes much clearer (see figure below).



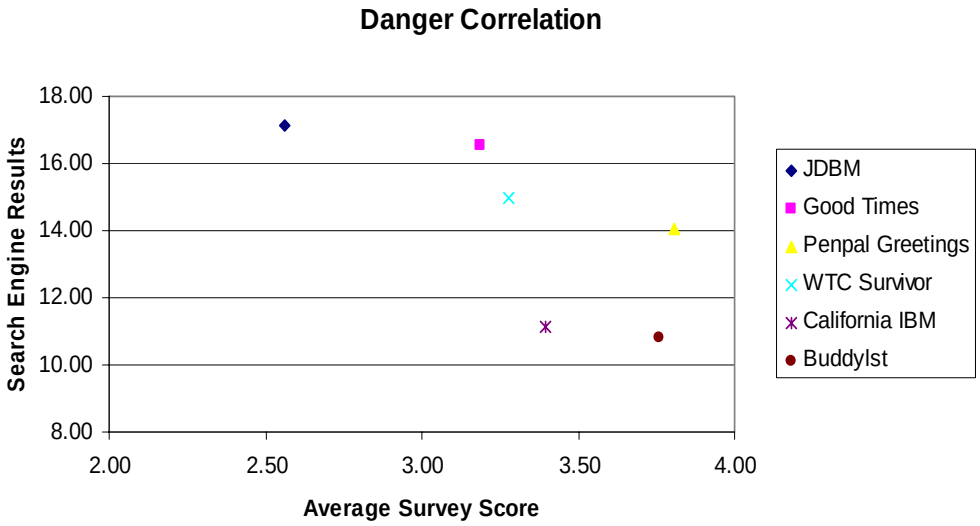
Graph 4: Flesch Kincaid-Simplicity Correlation (without Good Times)

Readability scores like these are calculated by only concentrating on surface characteristics of a text, such as word and sentence length. However, readability relates to other factors as well. By only measuring these lengths, the formula pays no attention to criteria such as cohesion and coherence, which help to make a text understandable. This is demonstrated by the fact that a hoax such as Good Times scores low on simplicity while it is a very coherent hoax.

The correlation between the degree of spreading and the simplicity of the hoax as scored by both the authority and the participants on the survey is much alike. The expert's correlation is negative 0.537 whereas the participants' score is negative 0.518.

5.4.4. Danger

All of the hoaxes in this test contained a danger factor. However, the expected result that there would be a positive correlation between the hoaxes with the highest proliferation and the hoaxes which scored the highest on the danger criterion was not endorsed by this research. The scattergraph shows a negative correlation. The negative correlation between the search engine results and the survey average is negative 0.72. When compared to all searches (including the newsgroups) the degree of correlation is a negative 0.69.



Graph 5: Danger Correlation

The hoaxes which scored best on the danger criterion are Penpal Greetings and Buddylst. The threats in Penpal Greetings message are 'a Trojan virus', 'destruction of data', 'destroying data with anyone who is in your mailbox' and even 'do a great deal of damage world wide'. Buddylst names 'destroying the hard disk' and 'spreading private information'.

The difference between the threats contained in Buddylst and Penpal Greetings versus the other hoaxes is that the first ones focus on the danger. This is done by capitalizing certain words in phrases that express the explicit danger. For example in the fifth paragraph of Penpal Greetings we find a capitalization of 'destroy' and 'damage', two words which express an immediate danger.

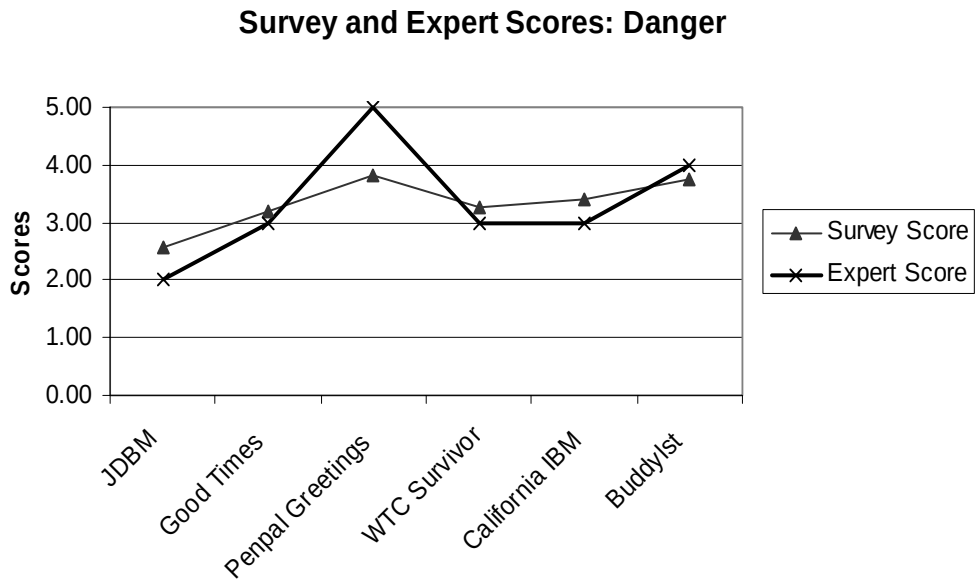
> This virus will DESTROY your hard drive, and holds the potential to DESTROY the hard drive of anyone whose mail is in your inbox, and who's mail is in their inbox, and so on. If this virus remains unchecked, it has the potential to do a great deal of DAMAGE to computer networks worldwide!!!!

The lowest danger is found in jdbgmgr.exe. This is explained more in the next criterion, benefit. The other low scoring danger (scoring .6 higher then jdbgmgr.exe) is the Good Times hoax. The threats contained in Good Times are 'destructive capability' and 'trashing the computer'. It also mentions to be more dangerous than other viruses such as 'Stoned, Airwolf and Michelangelo'. The inclusion of the names of these viruses is an indication of the age of the hoax. Good Times first appeared around 1994 in the databases of virus-companies. This means that it was spread around that time. Both Michelangelo and Stoned had received a lot of media attention thanks to their proliferation and the threat they imposed on the digital community of that time (Social Impact). The danger threat imposed in this first paragraph would therefore have been a lot higher if this survey was done in the early nineties.

The importance of this time limited reference in Good Times shows that it is hard to make snapshots of the situation of hoaxes. By the time this research is finished, the situation will have changed, some hoaxes will be more popular and others will diminish.

When we look at the correlation between danger and the spreading of the hoaxes we see that the high scores for danger are linked to both a hoax from the average spreading group and one of the low spreading. These two figures are the main reason for the strong negative correlation. This means that if we do not take these two exceptions into account, there is no correlation between the spreading of the hoaxes and the degree of danger within a hoax.

The difference in correlation total between the score given by the expert and the survey takers' score is 0.3. The negative correlation is stronger within the results of the survey then in the authority's results. In the experts score we also see that Buddylst and Penpal Greetings score the highest (see graph below). The maximum score of Penpal Greetings is the main factor of difference between the two results. The same general trend (a negative correlation) is noticeable in both results as the hoaxes are in the order of most spread to least spread.



Graph 6: Survey and Expert Scores: Danger

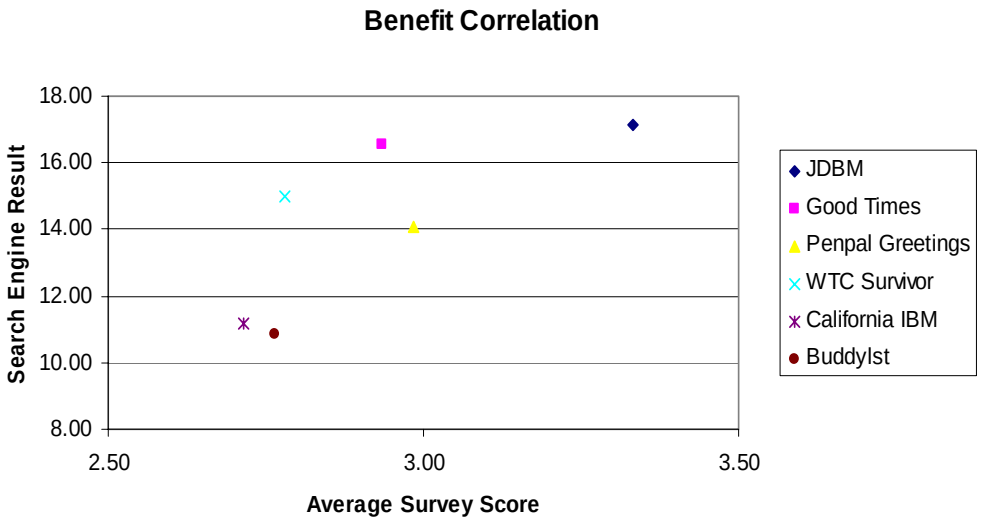
5.4.5. Benefit

Benefit is another criterion that has a high correlation with the degree of spreading of the hoaxes. A detailed look at the individual benefit scores for every hoax shows that jdgbmgr.exe tops the list of benefit whereas the other hoaxes follow behind but all have a very close score, ranging between 2.74 for California IBM and 3.33 for Jdbgmgr.exe (see table 8). Nevertheless this indicates that there is a correlation between the fitness of the hoaxes and the benefit as scored by the survey takers.

High-Low	Scores
Jdbgmgr.exe	3.33
Penpal	2.98
Good Times	2.94
WTC survivor	2.78
Buddylst	2.77
California-IBM	2.74

Table 8: Individual Benefit Scores per Hoax

The graphical representation of the correlation of benefit shows the high positive correlation (graph 7). Compared to the Search Engine Results this correlation is positive 0.75. If the results of the newsgroup search are taken into account as well, the correlation goes up to positive 0.82.

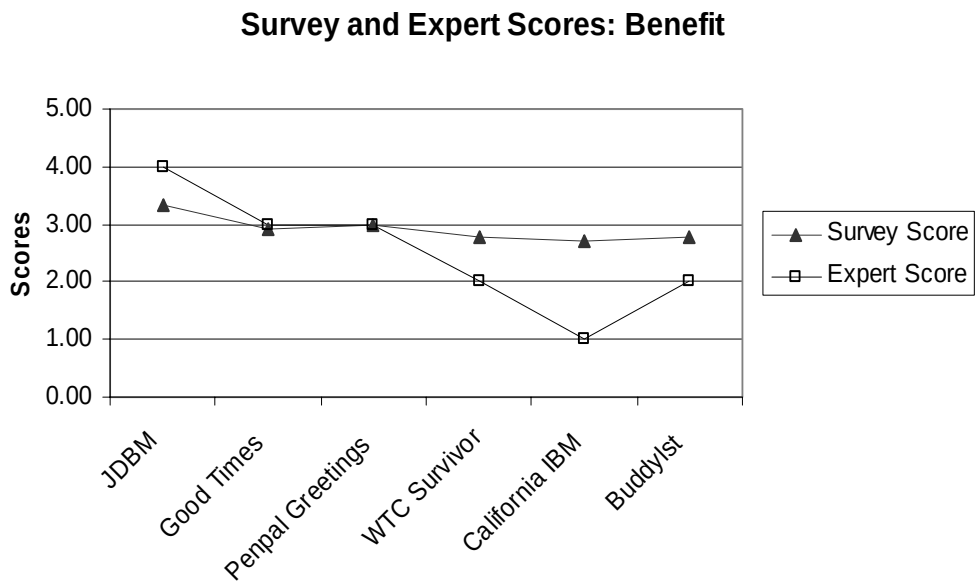


Graph 7: Benefit Correlation

The hoax which has the highest individual benefit score (cf. table 8) is the Jdbgmgr.exe hoax. This hoax promises an immediate solution to a threat, thereby neutralizing the danger immediately. The Jdbgmgr.exe hoax has both the highest benefit and the lowest danger.

Giving a solution to a possible danger helps the hoax to spread as it is linked to the objective criterion of controllability. By offering a direct solution the receiver of the hoax has an alternative way of confirming the false threat, more specifically by carrying out the instructions and finding that, indeed, the specified file exists on his or her hard drive. This adds not only to the benefit score of this hoax but also to the internal *self-centered* authority score of the hoax (see 5.4.7). Unless the receiver of this hoax refers to hoax databases or has prior knowledge of the existence of this hoax or a variant (cf. novelty) he or she will not know that this is a false warning and find an immediate positive reinforcement when the file is found in the system.

The other hoaxes each work according to the same system; they warn for an email message which may possibly come or, in the case of Buddylst warn about a certain file which one may receive by email. This is the way most virus hoaxes work. They warn in advance for a possible threat. Their benefit is therefore linked to their danger. They promise an answer to a solution as well but it is a solution which will only work in the future. If ones system would be infected with the virus of which the hoax speaks, it would be too late. The hoax is a warning in advance whereas the hoaxes such as jdbgmgr.exe give solutions to an already existing problem.



Graph 8: Survey and Expert Scores: Benefit

The main difference between the benefit scores of the survey and the benefit scores of the authority is the individual benefit score of California IBM. The expert scored this hoax' benefit on one whereas the average of the survey takers score was 2.71. The total correlation score of the two sources is not much different. The expert's correlation between benefit and the degree of spreading is 0.81; the survey correlation is 0.74.

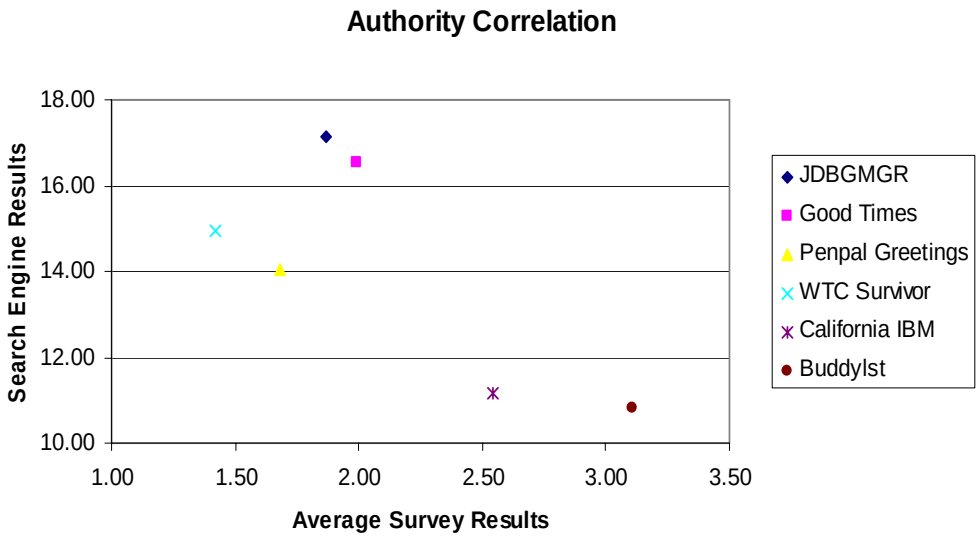
5.4.6. Utility

Danger and benefit are both part of the utility criterion. In this survey they were split up to be able to see if there is a significant difference between the two sub-criteria. The results of this survey confirm the significant difference between the two criteria; benefit having a strong positive correlation whereas danger has a negative correlation. When we average both criteria we find a small negative correlation for utility in general.

One of the reasons why benefit may score higher is a social acceptance of benefit over the social acceptance of danger. When passing on a danger, one imposes an indirect threat on the receiver of the email; passing on a benefit is passing on an award, a bonus. It is therefore more likely that the receiver will appreciate the reward, the benefit than the danger.

5.4.7. Authority

Authority also has a negative correlation with the spreading success of the hoaxes if we look at the results of the survey. As mentioned in 4.6.5 this research did not expect to find a strong correlation. This is also explained in depth in the difference between internal and external authority.



Graph 9: Authority Correlation

The hoax with the highest authority rating is Buddylst. Buddylst mentions two sources of authority, Microsoft and America Online (AOL). The second highest scoring hoax is California IBM which only mentions Microsoft. Good Times, the third runner up in the ranking of authority mentions America Online and Jdbgmgr.exe mentions that the virus will not be detected by McAfee or Norton. Those last two are notable because, although they mention the companies, they do not claim to be endorsed by those companies whereas the first two hoaxes explicitly mention that they are backed up by Microsoft and AOL. This shows that there is more than one kind of internal authority within these hoaxes. WTC survivor also claims an authority source but in that case it is not a company but 'a friend' who assumingly 'called and warned' the sender of the hoax of the impending doom that was threatening his system.

Authority can be split up into both internal and external authority. The internal authority of the hoax is the authority information which is available within the hoax. In these cases the companies that are mentioned.

Within the range of internal authority it is possible to make a distinction between three types of authority. A first type is found in Buddylst and California IBM, where the companies are used to provide a backup to the hoax and in WTC Survivor where this role is played by an anonymous person, the friend. The second kind is found in Jdbmgr.exe and Good Times where authority sources are mentioned that do not give the extra weight of endorsement but that give the hoax a certain degree of trustworthiness.

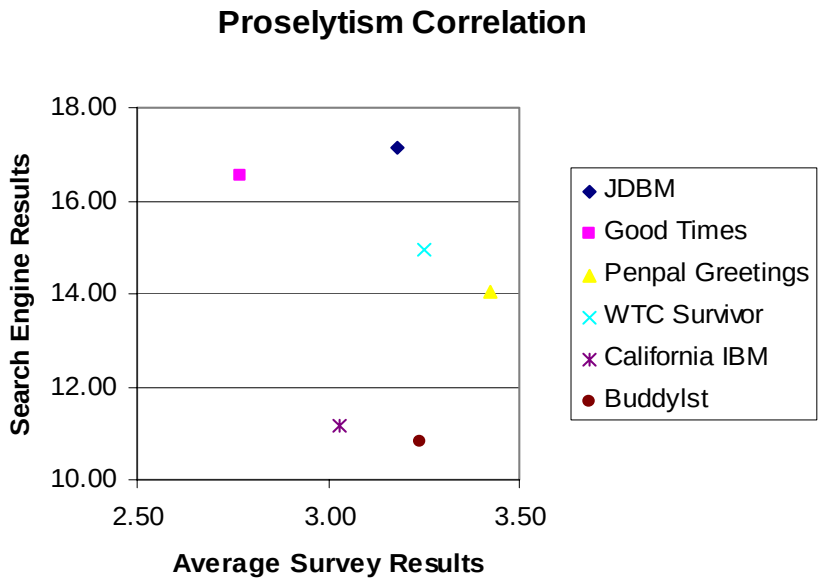
A special kind of internal authority is what could be called a 'meta-authority' or 'self-centered authority'. This is the internal authority that can be found in the Jdbgmgr.exe hoax when one considers its controllability. Because of the confirmation of the information contained in the hoax, the hoax will obtain a degree of authority although no authority was ever mentioned in the hoax. In this case, the hoax itself becomes an authority.

External authority in this type of hoax is what could be called the authority which the receiver of this hoax gives to the sender of the email containing the hoax or to the publisher if the hoax is spread in Usenet forums or on websites. Each of these communication channels has its own degree of authority. Email reaches a potential host in a surrounding of intimacy, in a personal, private environment. Usenet postings are wider than that but are still aimed at a limited public, those who read the specific newsgroup to which this message is posted. A website is the most general of these communication methods and will therefore have the least authority unless the maker of the website is already an authority figure to the potential host.

The only authority that was measured in this research is the internal authority. The external authority may play a role in the spreading and success of the hoax but this can not easily be measured through a survey or an interview as it is a very subjective criterion which is different for every person. A possible way to measure this is by asking the participants by whom they would have to have received a certain hoax which would give it enough credibility to send the hoax on to other people. However, an open question such as this limits the role of the other criteria as it directly focuses on the external authority.

5.4.8. Proselytism

Surprisingly the results of the survey showed that the more a hoax demands to be spread, the less it is actually spread. This negative correlation is rather low however; negative 0.24 on the search engine results and negative 0.28 when considering all the searches. In both searches this is a near zero correlation, meaning that there is no significant correlation between the replication pressure as measured in this test and the degree of spreading.



Graph 10: Proselytism Correlation

It is important to look at how exactly the hoax asks to be spread to measure the proselytism of this hoax. The highest scoring hoax for proselytism is Penpal Greetings. Although by sentence count Penpal Greetings only demands replication in one sentence, it is still ranking the highest. This can be explained because Penpal Greetings is not only asking to be passed on in email but also in other forums.

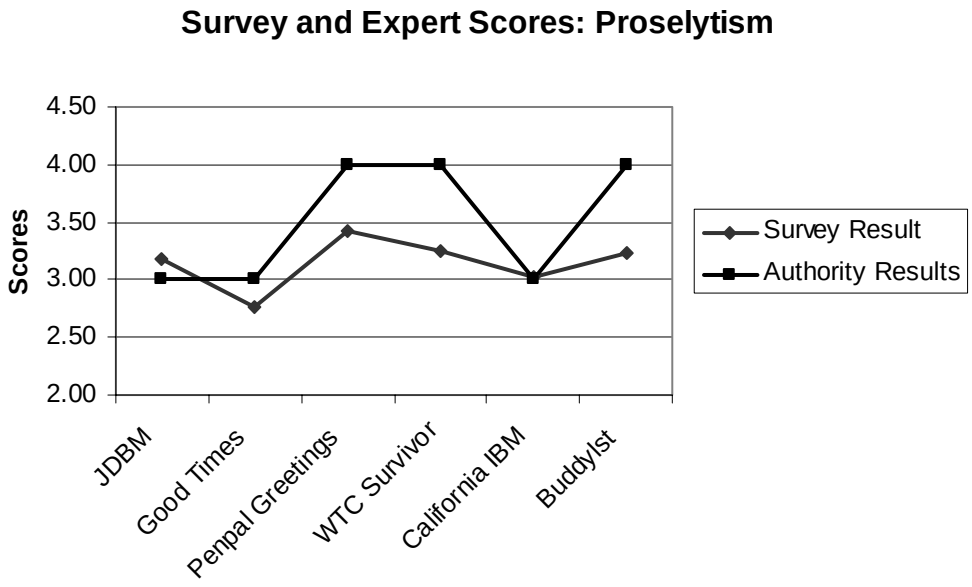
> And pass this message along to all of your friends and relatives, and the other readers of the newsgroups and mailing lists which you are on(…)

The other hoaxes in this research use the same method to demand the spreading. If we count the number of times they urge the receiver of the hoax to send it on, Buddylst tops the ranking with three times: "please pass it on to anyone you know who has access to the internet", "please copy this information and e-mail it to everyone in your address book" and "(…) pass this information on to your friends, acquaintances and work colleagues". Nevertheless, Buddylst is only the third one in rank when it comes to proselytism. It comes very close to the second hoax in rank however, with a score that only differs 0.01. The second highest score, with an average of 3.25, is WTC Survivor. This hoax only demands replication twice but it does this once in a phrase within a paragraph and once at the end of the message as a single line, all capitalized and thus putting extra weight on the demand. We can see this as well in the fourth hoax in this ranking, jdbgmgr.exe. Here the demand is conditional: "if you find the virus(…) send this message to all of your contacts located in your address book(…) ". The sentence in the hoax is capitalized.

	Demands	Send to	Caps	Score
Penpal Greetings	1	Address book newsgroups mailing lists	No	3.42
WTC Survivor	2	Address book Unspecified	No Yes	3.25
Buddylst	3	Anyone you know Address book Friends, Acquaintances Work colleagues	No	3.24
Jdbgmgr.exe	1	Address book	Yes	3.18
California IBM	2	E-mail contacts Unspecified	No	3.03
Good Times	1	Friends and local system users	No	2.77

Table 9: Proselytism Features

Table 9 shows that capitalizing the demand does indeed make a difference in the scoring of the proselytism. Whenever a demand is capitalized it has a higher score then the ones where it is not capitalized as can be seen between WTC-Survivor and California IBM which both have two demands but WTC Survivor has one of these demands capitalized. This can also be seen when comparing Jdbgmgr.exe with Good Times.



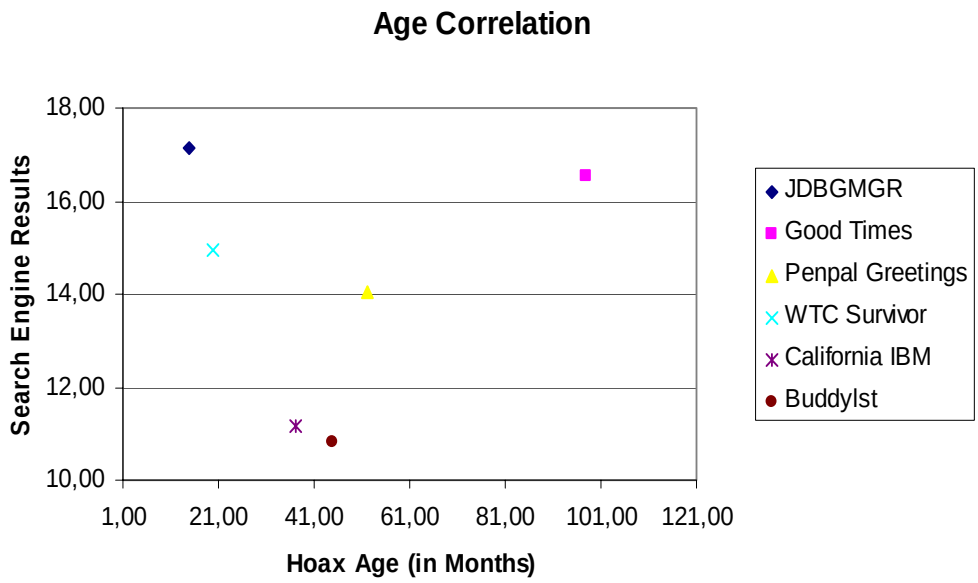
Graph 11: Survey and Expert Scores: Proselytism

Here again we find a relatively close correlation between the score of the expert and the survey scores. Both correlations are very low however; the authority's correlation is minus 0.34, the survey scores have a negative correlation of 0.24. For both the survey results and the expert results this is the lowest scoring correlation of all the criteria.

5.4.9. Hoax age

One possible explanation for the degree of spreading of hoaxes may be the age of the hoax. The longer a hoax has been around, the more chances it has to be passed on infinitely. However, older hoaxes will most likely get weaker over time as they will loose the battle for hosts with newer hoaxes which prove to have a higher fitness, for example because of the novelty that they possess.

The age of the hoaxes was determined by looking through the databases of the companies and organizations which have hoax databases. Especially Mc-Afee and CIAC keep track of the date when a hoax was added to their database.



Graph 12: Hoax Age Correlation

The slight positive correlation which can be found in Good Times, Penpal Greetings, California IBM and Buddylst is mainly unbalanced because of the results of Jdbgmgr.exe and WTC Survivor. These two virus hoaxes are the youngest hoaxes in the research but are among the most widespread hoaxes in this project. This means that these hoaxes, and especially jdbgmgr.exe, have threatened the spreading of other hoaxes and seem to have taken over some of the attention the others were receiving.

Because this research is only a snapshot, a description and analyses of the situation at a certain time, it is expected that this situation will change. The proliferation behavior of hoaxes is not one that works very gradually but can be seen as a sudden boom when the hoax first appears which then slowly fades out. Because of this Jdbgmgr.exe is still showing strongly whereas older hoaxes such as Buddylst or California IBM have known their booms a long time ago and now the remnants of this are showing.

When novelty and age of a hoax are compared it would be expected that the newest hoaxes score the highest on this criterion. However, Good Times is the oldest hoax but has one of the highest novelty scores (cf. 5.4.2). If the age category of the survey takers is also taken into account it becomes clear that to the participants Good Times is indeed novel as it is safe to assume that the participants will not have been in contact with the Good Times hoax when it was first released. If the Good Times virus would be updated (for example with references to newer viruses) and spread again, chances are that it would proliferate well. This would be an example of the revival of a meme from its dormant state to an active state.

5.5. Conclusion

The hypothesis of this thesis was that all selection criteria will have a clear correlation with the degree of fitness of the different hoaxes. After considering the results of the survey, it is found that indeed there is a correlation, although it is not the positive correlation that was expected. Only two of the six criteria have a strong positive correlation with the success rate of the hoaxes. The other criteria are found to have a negative correlation, varying from very light (cf. proselytism) to a very strong negative correlation (e.g. simplicity). To validate these results and to make sure that they have not been caused because of a poor understanding of the selection criteria, the results were compared to those of an authority figure in the field of memetics. We see that this expert's scores have the same tendency and the same positive and negative correlations. Therefore the results can not be explained with this reasoning.

This research has shown that there is a wide variety in the ways the hoaxes use the criteria. Every hoax has found its own niche in the brains of several new hosts and on the Internet. Therefore all hoaxes here have already gone through the selection procedures that have been measured. The popularity of these hoaxes is different and the analysis of every single criterion that was measured in this survey has shed a new light on the ways in which hoaxes have adapted to the criteria in order to become as good or 'as fit' as possible.

The wide variance in techniques shows how exactly the hoaxes became successful. The importance of finding these techniques is arguably even greater than the failure to make a classification of the criteria. After all, it has become clear now that the criteria are very important but that the way in which the hoaxes use the different criteria has even more influence on their success rate.

A minor goal of this project was to find out if the utility criterion could be split up in both a danger and a benefit criterion. Although initially both criteria were expected to have a positive correlation with the degree of fitness, the results of both the expert and the survey showed that there was a great difference. Possible explanations for this phenomenon were given and the different degree of importance between the two criteria was confirmed.

6. Overall Conclusion

Initially all criteria were considered equally important. Heylighen's theory of memetic selection criteria claims that all criteria will have a positive correlation with the degree of spreading. When focusing on virus hoaxes this research has shown that not all of the criteria are equally important and that some criteria may not be important or may have a negative correlation with the fitness of these kinds of memes.

The importance of subjective criteria and previously acquired knowledge is very important for scoring hoaxes. Furthermore this survey analysis has shown that factors which were not included in the original theory about memetic selection criteria play an important role as well. These factors are explicitly linked to the medium of communicating the hoaxes, electronic communication and to the method of representation of the meme, written text.

Both the medium and the method of representation limit the sole importance of the memetic selection criteria and offer other explanations for the degree of fitness of a hoax. The structure of a hoax as well as stylistic criteria such as the capitalization of words and sentences have a strong influence on the correlation between certain criteria and the hoax' popularity.

Nevertheless, this analysis has shown that it is possible to do a quantitative research on selection criteria when one focuses on a small subcategory of memes; in this case the virus hoaxes; and when one keeps in mind that factors other than only the selection criteria can interfere with or strengthen the importance of these criteria.

By using an authority result as a reference it was possible to make sure that the different criteria were explained according to the original definitions and that both the expert in the field and the audience compared the hoaxes to this same definition.

A number of questions remain open however. This research has focused on a very limited public; university students who speak English as a second or foreign language. It would be very interesting to compare their results to the results of a native speaking audience to see whether or not they interpret the criteria differently and to see whether or not the criteria have a different weight or order of importance according to or dependent of the mother tongue of the participants.

Another question which remains is the status of the other criteria. The size of this research made it impossible to expand on the number of criteria or the number of hoaxes. A further research project could try and measure the other criteria as well, spread out over more hoaxes and more participants. It would take quite some time to set up such a research project but it would be a major breakthrough in the research of the theory of memetics.

The particular evolution of hoaxes and memes could be studied using the different ways in which the hoaxes use the variety of methods to alter their score on certain criteria. This way one could compare the different stages in the phylogeny of a certain hoax and follow the evolution of this specific hoax to a point where it maximizes its fitness.

Linguistic principles about the functions of language, socio-linguistic ideas about why people communicate and reception analysis of different kinds of information can offer a number of new insights in memetic spreading. Theories about language evolution and selection mechanisms which can be compared to genetic drift can offer new insights in various linguistic fields; the theory about meme-gene co-evolution can maybe provide answers, or at least provoke discussions, among those who think about the origin of language and the beginning of speech. It is time to cross the borders between the two fields and to take a genuine interest in the research done in both domains. This project has tried to take a first step in this direction.

7. Bibliography

Bjarneskans, Hendrik, Grønnevik, Bjarne, and Sandberg, Anders. The Lifecycle of Memes. 1997. Transhumanist Resources. 1 Mar. 2003 <http://www.aleph.se/Trans/Cultural/Memetics/memecycle.html>.

Blackmore, Susan. The Meme Machine. Oxford: Oxford University Press, 1999.

Bouissac, Paul. "Why Do Memes Die?" Semiotics 1992. Ed. J. Deely. Lanham: University Press of America, 1993. 183-191: 5 Dec. 2002 <http://users.lycaeum.org/~sputnik/Memetics/MEMES.TXT>.

British Columbia Scientific Cryptozoology Club. BCSCC Website. 10 Jul. 2003 <http://www.cryptosafari.com/bcscc/nessie.htm>.

Castelfranchi, Cristiano. "Towards a Cognitive Memetics: Socio-Cognitive Mechanisms for Memes Selection and Spreading." Journal of Memetics – Evolutionary Models of Information Transmission 5, 2001: 5 Dec. 2002 <http://jom-emit.cfpm.org/2001/vol5/castelfranchi_c.html>

Dawkins, Richard. The Extended Phenotype. Rev. ed. Oxford, Oxford University Press, 1999.

---. The Selfish Gene. 2nd ed. Oxford, Oxford University Press, 1989.

Dennet, Daniel C. Consciousness Explained. London, Penguin Books, 1991.

---. Darwins Dangerous Idea. London, Penguin Books, 1995.

---. "Memes and the Exploitation of Imagination." Journal of Aesthetics and Art Criticism 48 (1990): 127-35.

Diaz, Arnold. "Internet Hoax Scares Diabetics." ABCNEWS.com 10 Feb. 1999. 10 Jul. 2003 <http://abcnews.go.com/onair/2020/diaz990210.html>.

Edmonds, Bruce. Letter: "Three Challenges for the Survival of Memetics." Journal of Memetics - Evolutionary Models of Information Transmission 6, 2002: 28 Nov. 2002 <http://jom-emit.cfpm.org/2002/vol6/edmonds_b_letter.html>.

Godin, Seth. Unleashing the Ideavirus. London, Simon & Schuster, 2000.

Grant, Glen. "Memetic Lexicon." Principia Cybernetica Web (Principia Cybernetica, Brussels). Ed. F. Heylighen, C. Joslyn and V. Turchin. 1990. 21 Jan. 2003 <http://pespmc1.vub.ac.be/MEMLEX.html>

Heavens Gate Website. 5 Dec. 2002 <http://www.heavensgate.com/>.

Heylighen, Francis and Dewaele, Jean-Marc. Formality of Language: definition, measurement and behavioral determinants. Internal Report, Center "Leo Apostel", Free University of Brussels, 1999. 21 Jan. 2003 <http://pespmc1.vub.ac.be/papers/Formality.pdf>.

Heylighen, Francis. "Memetics." Principia Cybernetica Web. Ed. F. Heylighen, C. Joslyn & V. Turchin. Brussels: Principia Cybernetica. Nov 2001. 3 Mar 2002 <http://pespmc1.vub.ac.be/MEMES.html>.

---. "Objective, subjective and intersubjective selectors of knowledge" Evolution and Cognition. 3,1 1997. 63-67. 21 Jan. 2003 <http://pespmc1.vub.ac.be/papers/knowledgeselectors.html>.

---. "What makes a meme successful? Selection Criteria for Cultural Evolution." Proc. 16th Int. Congress on Cybernetics. Namur: Association Internat. de Cybernétique. 1998: 05 Dec. 2002 <http://pespmc1.vub.ac.be/Papers/MemeticsNamur.html>.

Hofstadter, Douglas. Metamagical Themas: Questing for the Essence of Mind and Pattern. New York: Basic Books, 1996.

James, William. "Great Men and their Environment." Atlantic Monthly. 46. 1880: 12 Dec. 2002 <http://www.emory.edu/EDUCATION/mfp/jgreatmen.html>.

Lynch, Aaron. Thought Contagion: How Belief Spreads Through Society. New York: Basic Books, 1996.

Murray, Ian. All Your Base Are Belong To Us: The History. 07 July 2003 < http://www.tvsian.com/hubert/article.php?sid=1>.

Potts, Rick. Humanity's Descent: The Consequences of Ecological Instability. New York: Avon Books, 1996.

Rapoza, Jim. App Handles Surveys for Less. 2 Jun. 2003. eWeek. 9 Jul. 2003 <http://www.eweek.com/article2/0,3959,1111892,00.asp>.

"Social Impact." Virus: a retrospective. Stanford U. 23 Jul. 2003 <http://www-cse.stanford.edu/classes/cs201/projects-00-01 /viruses/social.html>.

Sophos describes hoaxes and scares. Sophos. 5 Dec. 2002 <http://www.sophos.com/virusinfo/hoaxes/>.

Sophos hoaxes and scares description: Jdbgmgr.exe. 15 Jul. 2003 <http://www.sophos.com/virusinfo/hoaxes/jdbgmgr.html>

Wilkins, John S. "What's in a Meme? Reflections From the Perspective of the History and Philosophy of Evolutionary Biology." Journal of Memetics. 2,1, 1998: 5 Dec. 2002 <http://jom-emit.cfpm.org/1998/vol2/wilkins_js.html>.